Attention is drawn to the fact that the copyright of this thesis rests with its author.

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without the author's prior written consent.

D 19167/77

Sawbridge, P H

pp 445.

WHAT DID HE SAY THAT FOR? Some contextual effects on the process of understanding a sentence.

Paul H. Sawbridge
University of Stirling
September 1976

Ph.D.

Ph.D. awarded ~~March~~ 1977 (February)
Conferred 1977 (March)

ACKNOWLEDGMENTS

I would like to acknowledge the assistance of the following:

Ranald Macdonald	
Paul Jackson	
Graham Hitch	for help with statistics
Angus Annan	
Paul Shanahan	
Paul O'Riordan	for technical help
Robin Campbell	
Terry Bowe	for general advice and criticism
Robin Campbell	
Paul Shanahan	for constant encouragement

Robin Campbell had the difficult job of supervising the work reported here.

I am most grateful to my wife, Rachel. Without her this work would never have been completed.

This research was financed by the Social Science Research Council.

CONTENTS

	Page
Acknowledgments	iii
List of tables	x
List of figures	xiii
Abstract	xvi
Chapter One. Introduction and literature review.	2
Chapter Two. The influence of context on the comprehension and verification of sentences describing simple spatial relations between objects.	50
Introduction	50
Why study topicalisation?	50
What picks out the topic?	52
Rationale and predictions for Experiment 1	61
Method	67
Results	71
Comprehension	72
Verification	85
Discussion	101
Comprehension data	101
Lexical marking	101
Syntactic marking and topic position	103
Summary of results not involving definiteness factors and suggested explanations	104
Results involving definiteness marking	107
Verification data	110
Chapter Three. Two experiments on the production of sentences describing simple spatial relationships.	113

	Page
Experiment 2	114
Method	114
Results	117
Discussion	141
Experiment 3	146
Introduction	146
Method	148
Results	151
Condition A	155
Condition B	157
Condition C	160
Condition D	163
Discussion	170
Chapter Four. The role of pronouns in cohesion, with special reference to the three term series problem.	176
Experiment 4: A very preliminary investigation of four methods of making anaphoric reference.	177
Introduction	177
Method	179
Results	181
Discussion	183
Experiment 5: The three term series problem re-examined.	186
General introduction to Experiment 5	186
The Literature: part one: Data	187
The Literature: part two: Models	194
The role of cohesion and systemic choices in the three term series problem	203
Method	209

Page

Results 215

The first premise 215

The second premise 215

The question 220

Total times 226

Errors 230

Discussion 231

The first premise 231

The second premise 232

The question 237

Conclusions 241

Chapter Five. Three experiments involving verification of transitive declarative sentences. 245

Transitive sentences 246

The Literature 247

Rationale of the present series of experiments 262

Experiment 6 268

Method 268

Results 274

Reaction times 275

Number of eye movements 280

First two fixations 280

Discussion 282

Experiment 7 285

Method 285

Results 285

Reaction times 288

Number of eye movements 288

First fixations 292

Discussion 293

Page

Experiment 8 296

Method 296

Results 300

Total reaction times 300

Adjusted reaction times 303

Number of fixations 303

First two fixations 308

Discussion 310

Experiments 6,7, and 8: general discussion 318

Chapter 6. An experiment involving answering Wh- questions. 324

Introduction 324

Linguistic analysis of a subset of Wh- questions 325

Predictions derived from the linguistic analysis 330

Method 332

Results 335

Discussion 357

Chapter 7. Conclusions. 365

Substantive conclusions 365

Methodological conclusions 373

Error rates 373

Fixations 374

Production frequency 375

Reaction times 375

References 382

Appendix A 403

Appendix B 409

Appendix C 413

Appendix D 423

LIST OF TABLES

Chapter Two — Page

1. Mean reaction times and error totals: text data. 73
2. Mean reaction times and error totals: no text data. 75
3. Comprehension data: analysis of variance, part one. 78
4. Comprehension data: analysis of variance, part two. 82
5. Verification data: analysis of variance. 86
6. Comprehension data: summary of major effects, first analysis. 88
7. Comprehension data: summary of major effects, second analysis. 95
8. Verification data: summary of major effects. 98

Chapter Three

1. Experiment 2: X^2 tables. 121
2. Experiment 2: results, article inserted data. 123
3. Experiment 2: article inserted data: effects. 126
4. Experiment 2: results, relational term or topic inserted data. 129
5. Experiment 2: relational term or topic inserted data: effects. 132
6. Experiment 2: results, blank sentence frame data. 135
7. Experiment 2: blank sentence frame data: effects. 138
8. Experiment 3: results, condition A (sketch oriented description starting with central object). 154
9. Experiment 3: results, condition B (sketch oriented description starting point unspecified). 158
10. Experiment 3: results, condition C (schematic description starting with central object). 161
11. Experiment 3: results, condition D (schematic description starting point unspecified). 161

Chapter Four

1. Six types of series problem. 189
2. Four studies of the three term series problem. 191
3. Clark's data on problems in which it is not possible to produce a strict ordering of the elements. 192

Page

4. A slightly revised version of Johnson-Laird's formulation of Clark's model: predicted operations. 196

5. An extremely crude method of comparing the models of Huttenlocher and Clark with the data from the four studies. 197

6. Johnson-Laird's formulation of Huttenlocher's model: predicted operations. 201

7. Experiment 5: adjectives used together with the mean RT to each of them in the first premise and their Lorge magazine count frequency. 213

8. Three term series problem experiment: means in msec. 216

9. Three term series problem experiment: analyses of variance. 217

10. Second premise: list of significant effects. 219

11. Questions: list of significant effects. 222

12. Total times: list of significant effects. 225

13. Error data: list of significant effects. 228

14. Predictions from Johnson-Laird's versions of the linguistic and image models for the second premise times. 233

15. Selected predictions from Clark's question answering model as formulated by Johnson-Laird. 236

Chapter Five

1. Reaction time data: Experiment 6. 274

2. Mean number of eye movements: Experiment 6. 277

3. Experiment 6: analyses of variance. 278

4. Experiment 6: means for major effects. 279

5. Experiment 7: mean reaction time data. 287

6. Experiment 7: mean number of eye movements. 289

7. Experiment 7: analyses of variance. 290

8. Experiment 7: means for major effects. 291

9. Experiment 8: mean reaction times, unadjusted. 299

10. Experiment 8: mean reaction times, adjusted. 302

11. Experiment 8: mean number of eye movements. 304

Page

12. Experiment 8: analyses of variance. 305

13. Experiment 8: means for major effects. 306

Chapter Six

1. Suggested list of correspondences between indicative and Wh-interrogative for the set considered. 329

2. Mean stimulus durations. 336

3. Mean onset-onset times. 337

4. Mean offset-onset times. 338

5. Mean number of fixations. 339

6. Analysis of variance: stimulus materials. 342

6A. Summary of significant effects: stimulus durations. 343

7. Analysis of variance: onset-onset times. 344

7A. Summary of significant effects: onset-onset times. 345

8. Analysis of variance: offset-onset times. 346

8A. Summary of significant effects: offset-onset times. 347

9. analysis of variance: fixations. 348

9A. Summary of significant effects: number of fixations. 349

10. Total number of errors. 355

LIST OF FIGURES

Chapter One — Page

1. A highly schematic summary of Clark's verification model. 7

Chapter Two

1. Experiment One: reaction times to each sentence type. 93

Chapter Three

1. Experiment Two: article inserted data, unmarked syntax. 124
2. Experiment Two: article inserted data, marked syntax. 125
3. Experiment Two: relational term or topic inserted data, unmarked syntax. 130
4. Experiment Two: relational term or topic inserted data, marked syntax. 131
5. Experiment Two: blank frame data,unmarked syntax. 136
6. Experiment Two: blank frame data, marked syntax. 137
7. Experiment Three: distribution of responses in the four conditions. 168

Chapter Four

1. Johnson-Laird's formulation of Clark's linguistic model. 195
2. Johnson-Laird's formulation of the image model. 200
3. Experiment 5: reaction times to the second premise. 218
4. Experiment 5: reaction times to the question. 221
5. Experiment 5: total reaction times to each problem. 224
6. Experiment 5: mean number of errors for each problem. 227

Chapter Five

1. Apparatus for the eye movement experiments. 266
2. Plan geometry for the eye movement situation. 267
3. Time course of Experiment6. 271
4. Results of Experiment 6. 276
5. Results of Experiment 7. 286
6. Time course of Experiment 8. 298

Page

7. Results of Experiment 8. 301

Chapter Six

1. Experiment 9: principal results. 340

2. Experiment 9: subsidiary results. 341

Abstract

Though it seems intuitively very probable that processing of sentences is likely to be affected by the environment - both linguistic and non-linguistic - in which the sentence is encountered, remarkably little work has been carried out to investigate the parameters of the process. While some work in recent years (for example that of Bransford and his co-workers) has amply demonstrated that what is remembered from a sentence is strongly influenced by other sentences which have to be remembered, few people have carried out investigations like those of Huttenlocher and her co-workers into how the situation can affect comprehension of a sentence. Psycholinguistic work has for the most part been directed towards the process of understanding single sentences shorn, so far as possible, of any ties with either other sentences or the real world. This unfortunate occurrence seems to have been largely due to the overwhelming influence of transformational grammar, which encourages one to believe that processing of sentences constituting part of a text is much the same as processing of single sentences. At any rate the core processes would appear to be the same on this account.

In the present work an alternative theory of grammar is utilised. This theory, called systemic grammar, places great emphasis on the information structure of a sentence and hence, implicitly, on the relation between sentences and the context in which the sentence is encountered. The present work concentrates pre-eminently on the influence of other sentences on the processing of a particular sentence - rather than on the influence of the accompanying situation. However, much is said on the subject of Huttenlocher's work.

The work starts with a highly selective review, heavily reliant on Clark (1974), of relevant literature. The review concludes that the bulk of the work reported in the literature lacks a coherent theory within which to conceptualise contextual influences on sentence processing. It is also suggested that a number of effects reported in the literature may be

heavily dependent on contextual influences. Systemic grammar is put forward as a theory which might enable us to conceptualise some of the factors relevant to understanding sentences in context. A brief outline of systemic theory is given. Following this, nine experiments are reported on a variety of traditional effects, but in all cases manipulating whether the target sentences occur as part of a text or not. In addition variation of such cohesive devices as the use of pronouns and the definite article also occurs.

The first experiment involves the systematic manipulation of definiteness marking, syntax, relational term and presence of text. A verification task is used in which the sentence precedes the picture with separate measures of comprehension (how long subjects choose to have the target sentence exposed) and verification (how long it takes them to respond when they see the picture). Interesting results occur in both sets of data but the main findings are : firstly, that the traditional lexical marking effect only occurs with single sentence presentation - when embedded in a longer text there is no difference between reaction times to the marked and unmarked words; secondly, although the marked syntactic form (in this case with the locative phrase before rather than after the copula) tends to be harder to understand this effect almost entirely disappears when the sentence is in a context and the topic of the paragraph is the theme (first noun) of the sentence - with unmarked syntax reactions are quicker if the topic is the second noun, but neither of these effects occur, of course, in the no text case where there is no topic; thirdly the marked lexical form was responded to faster if the two nominals were marked differently for definiteness, whereas the unmarked form tended to be responded to more rapidly if both nominals were similarly marked. The first and last results were explained as due to a "good reason" interpretation of marking in which marking is considered subordinate to topicalisation choices and the marked form does not convey additional information if it can be seen to have been chosen for that reason. The second result received a related explanation, though with a fuller analysis of the role of this marked syntactic form.

Experiments 2 and 3 investigate precisely the same phenomena as Experiment 1. The first of these requires subjects to write down a series of sentences of the same sort as those presented to subjects in Experiment 1, to describe a series of pictures presented to them. Various constraints are built into the task in order to encourage them to produce a wide variety of responses. The frequency of different forms is similar to what one might expect from Experiment 1 given the assumption that reaction times and production frequencies are inversely related. Similar effects were observed to those in Experiment 1 with the exception that definiteness marking was seen to be of much greater importance in this experiment. This last result was also replicated in Experiment 3 - a much less constrained study in which subjects described pictures orally. Classifying responses on the basis of a large number of criteria this study demonstrated that very few of the possible responses occurred. However many more occurred with unmarked than with marked syntax - supporting the interpretation of the marked option as not in itself more complex, but rather with more complex entry conditions (selected in a narrower set of circumstances). This study also demonstrated some interesting differences in the patterns of use of pronouns and the definite article.

The next two experiments follow up this last point by investigating differences between pronouns and other means of cross-referring in terms of reaction times. It is clear from these two experiments that pronouns do not simply speed up comprehension relative to other methods. The effect seems to depend upon several factors including the information structure of the sentence. The fifth experiment used the three term series problem to examine the use of pronouns, lexical marking and Huttenlocher's result that the second premise is easier to understand if the new item is first in that premise. Reaction times to the first premise, the second premise and the question were measured separately. Huttenlocher's effect was greatly enhanced by the use of a pronoun in the second premise to cross refer to the first premise. This was interpreted as being due to pronouns making clear the new and old information parts of the sentence and so enabling subjects to take advantage of the fact that their primary focus of interest when reading the second premise - namely the third object - is referred to by the more prominent theme, something which is more usually reserved for old information. A second factor influencing processing of sentences with pronouns in them is whether the pronoun in the second premise refers to the same object as the subject or object of the first premise. Subjects respond more rapidly if it is co-referential with the subject. Experiment 3 demonstrated that this is also the more common occurrence in free descriptions. Other results in this experiment provided more support for the interpretation of lexical marking in terms of a good reason principle : there being a strong effect of marking of the first premise (where it is hard to see any topicalisation reason for choosing it) but no straightforward effect of marking of the second premise. Furthermore marked questions do not take longer to process than unmarked - in fact, thanks probably to an interaction, they actually take significantly less time.

The next three experiments again involve verification but here the presentation of sentences was experimenter controlled and oral. Reaction times were again used but the measure taken was a complex comprehension/verification one. This measure was supplemented by a measure of the number of fixations subjects made in scanning the picture. This set of data was analysed in much the same way as the reaction time results. Experiments 6 and 7 involved successive presentation of sentence and picture (in that order), while Experiment 8 involved simultaneous presentation. On the whole the latter was more successful but this may have been because more complex pictures were used. The fixation data, though producing a number of apparently reliable results, did not produce results which bore any clear relationship to the reaction time data and evidence to the contrary reported by Hall (1975) is therefore called into question. However the reaction time data as a whole are not very clear in these three experiments. Only in Experiment 8 in which passives are shown to be easier to understand when the theme is previously mentioned, and actives when the theme is not previously mentioned (this is true, of course, only for the context condition) are there any very clear results. In Experiments 6 and 7 on the other hand, it does seem that passives are only harder to understand than actives if they are false, but Experiment 8 only shows a simple effect of truth value. Effects of context in Experiments 6 and 7 are not large - possibly this is due to the delay between presentation of the sentence and the taking of any measure.

The final experiment again used the text manipulation and like Experiment 8 presented sentence and picture simultaneously, but the sentence was a question which had to be answered rather than a statement to be verified. Questions differed in whether the noun preceded or succeeded the main verb, in voice, and in whether the noun was previously mentioned or not. On the whole results approximated quite closely to what one might expect from corresponding declaratives and a functional interpretation of the systemic options involved. A feature of both this experiment and the previous one is the use of two sets of reaction time data : data from the onset of the question to the onset of the answer and data from the offset of the question to the onset of the answer. On the first analysis passives take consistently longer to process, but on the second they are, if anything, processed faster.

A final chapter summarises some of the major results and compares both the experimental methods and the measures used in the various experiments. On the whole the conclusion is that sentence by sentence presentation for subject-controlled durations is the most satisfactory method. The gross measure of number of fixations is not seen as a useful one, though it is suggested that with simultaneous presentation of sentence and picture a moment by moment comparison of the sentence with what the subject is fixating may be of interest. The main substantive contributions of the present work are seen as:

(1) further evidence that the canonical form view of sentence processing is unhelpful,

(2) a good deal of support for the"good reason" approach to both lexical and grammatical marking which explains the greater difficulty of marked forms as due, not to the fact that they are themselves more complex, but to the fact that the reasons for selecting them (entry conditions) are more complex,

(3) some preliminary evidence of the effects on RT of a handful of cohesive devices among them the use of definiteness marking, pronouns, lexical marking, the passive voice and certain other marked syntactic configurations.

Chapter 1 : Introduction and Literature Review [1]

There has been much interest in recent years in verification tasks, that is, tasks in which the subject has to compare an input sentence with other information. The information may have been presented previously (Just, 1974; Clark and Chase, 1972), it may be presented subsequently (Clark and Chase, 1972; Carpenter and Just, 1975) or it may be derivable from information presented previously (Trabasso, 1971). There do not as yet appear to have been any studies in which the two sets of material are presented simultaneously for the simple reason that the universal preference for written materials of necessity requires the subject to look at either the sentence or the other material first. This is true even when they are both present at the same time : indeed any possible peripheral pick up of information is ignored in model building (e.g. Clark and Chase, 1972).

The original interest in verification tasks clearly derived (as Johnson-Laird, 1974, has pointed out) from a desire to test for psychological correlates of the kind of structures which so-called "transformational-generative grammar" (Chomsky, 1957, 1965, 1970; Jackendoff, 1972) predicts underly sentences. In investigating the process of understanding a linguistic structure one needs to be sure that the structure has been understood by the subject : that the task is not susceptible to any simple strategies not requiring that the sentence be processed to any deep level. Earlier tests of transformational grammar (hereafter "TG") had tried various memorial techniques (Savin and Perchonock, 1965; Marks and Miller, 1964). These however are subject to the "echo-box" criticism : perhaps the subject does not really need to understand to respond appropriately (Fillenbaum, 1973). The ability to assess the truth value of a sentence is

1. This introduction is intentionally general in nature. It is intended merely to give an overview of the topics which will be tackled in more depth later. More thorough reviews of each topic are given in the introduction to the chapter in which experimental work on that topic is presented. (see Contents).

often thought of as the prime criterion for semantic processing to be said to have occurred. Indeed it has often been said that if we could produce a complete theory of truth we would be able to "capture" all that is required for a semantics of, say, English (Davidson, 1970). This is a rather narrow view of meaning, though, in that such obviously non-truth-conditional factors as differences in "register" (Halliday, M^{c}Intosh and Strevens, 1965) are clearly meaningful. In addition many people now believe that accounts of meaning should incorporate Gricean conversational conventions (Grice, 1969; Clark and Haviland, 1976). Even if one ignores the interpersonal aspects of the "meaning potential" of a language, there is still a lot more to be considered besides truth. Part of the purpose of this thesis is to show what else there is, and how we might investigate it. The assumption is made throughout that these other processes are additional to truth testing and do not interact with it. In the long run this is unlikely to prove tenable since these other factors are responsible to some extent for the assignment of reference to some noun phrases, for example those with pronouns in them. For the moment, though it is expedient to make a non-interactive assumption.

Even if such an assumption is made one still cannot be sure that the verification task, as used in most of the experiments to be reviewed, is an adequate measure of comprehension. In many cases it happens that if a sentence is false a similar sentence with the positions of the two terms reversed, is true. (For brevity's sake this will be called the "reverse" in what follows). This seems unlikely to be the commonest case in everyday life where, amongst other things, reference failure or disagreement over the strength of an assertion (e.g. "I said he was behind John - I didn't say he was following him") may intrude. Subjects may well capitalise on the equivalence of the falsity of a sentence and the truth of its reverse, to the extent of not performing a thorough analysis of the sentence. This is not to suggest that all everyday sentence comprehension is carried to any great depth. Indeed it seems a priori unlikely that we process sentences very fully in what Malinowski has called "phatic communion". But given that we take the view that language is essentially a means of conveying information about the state of the world (a model which is, as already noted, at least partly inadequate) we might do better to examine, in the first instance anyway, only cases where this is the prime role of the communication. Of course, this is not possible in the strictest sense as long as one is doing an experiment since the testing situation carries with it its own special features - features which can sometimes become a major problem. But one can at least try - subjecting one's conclusions to the usual *ceteris paribus* clauses.

The verification task, although it does carry with it the possibility of special artifacts in the context of many of the experiments in the literature, does at least bear similarities to an everyday occurrence. What is more it seems reasonable to suppose that in verifying a sentence in everyday life we do need to have carried the processing of the sentence to a comparatively deep level. It is therefore not really surprising that verification has become the most commonly used task in experimental psycholinguistics. What is perhaps much more surprising is that the original notion of using the verification task as a vehicle for investigating the comprehension of sentences (as exemplified in the work of Gough, 1966) and a method of making sure that subjects have processed the sentence to a reasonable depth, has come to be replaced in recent years by an interest in the verification process *per se*. If the comparison stage is affected, as surely it must be, by the kind of artificial equivalences noted above, surely the interest of any model of the process in an experimental setting must be rather limited. However investigators in this area seem to have taken the view that if an elegant model of the process can be developed, the question of its generality can be assessed later. It is to a selective review of work in this area that we now turn.

As Johnson-Laird (1974) has noted, the development of this area has been largely due to the independent work of Clark and Trabasso. Though their models differ in some respects they are sufficiently similar to merit a single treatment. Since Clark has attempted in a comprehensive review paper to integrate the results of work on a number of separate problems (Clark, 1974) and since he has worked with a wider range of problems than Trabasso, I will here concentrate on Clark's work and work arising directly from it. (The reader is referred to Glucksberg, Trabasso and Wald, 1973, Garrod and Trabasso, 1973, and Trabasso, Rollins and Shaughnessy, 1971, for details of the work of Trabasso and his collaborators). In his review paper Clark divides the field into four main areas : negatives, locatives, comparatives and voice phenomena. Experiments will only be presented in the last three of these four areas in this thesis, but since the model covers all four areas in essentially the same fashion, negatives are also covered in the present review.

Clark starts with what he calls the "deep structure assumption" which is that "linguistic deep structure" accurately represents what people know once they have comprehended a sentence (1974, p.1293). He eschews any discussion of the controversy over the nature of linguistic deep structures, such as whether they should be logical in form (Lakoff, 1972) or contain performatives (Ross, 1972) or be formed of case roles with or without compulsory transformations (Fillmore, 1968; Halliday, 1967, Hudson, 1971) or be more like "standard theory" base structures (Chomsky, 1965) or else those of Chomsky's earlier theory (Chomsky, 1957). These are all genuine alternatives for many linguists, despite Chomsky's (1971) claim that many of the different proposals may be only "notational variants" of the standard theory. Instead Clark opts for a notation in which sentences are represented as a series of simple "propositions" which can be embedded one inside the other. For example "John is not happy" would be represented as

((John is happy) is false).

Setting aside any possible debate about scope problems here ("John is not happy" may not be equivalent to ""John is happy" is false", depending on how one treats reference failure), it is not clear at what level we are to models differ in some respects they are sufficiently similar to merit a single treatment. Since Clark has attempted in a comprehensive review paper to integrate the results of work on a number of separate problems (Clark, 1974) and since he has worked with a wider range of problems than Trabasso, I will here concentrate on Clark's work and work arising directly from it. (The reader is referred to Glucksberg, Trabasso and Wald, 1973, Garrod and Trabasso, 1973, and Trabasso, Rollins and Shaughnessy, 1971, for details of the work of Trabasso and his collaborators). In his review paper Clark divides the field into four main areas : negatives, locatives, comparatives and voice phenomena. Experiments will only be presented in the last three of these four areas in this thesis, but since the model covers all four areas in essentially the same fashion, negatives are also covered in the present review.

Clark starts with what he calls the "deep structure assumption" which is that "linguistic deep structure" accurately represents what people know once they have comprehended a sentence (1974, p.1293). He eschews any discussion of the controversy over the nature of linguistic deep structures, such as whether they should be logical in form (Lakoff, 1972) or contain performatives (Ross, 1972) or be formed of case roles with or without compulsory transformations (Fillmore, 1968; Halliday, 1967, Hudson, 1971) or be more like "standard theory" base structures (Chomsky, 1965) or else those of Chomsky's earlier theory (Chomsky, 1957). These are all genuine alternatives for many linguists, despite Chomsky's (1971) claim that many of the different proposals may be only "notational variants" of the standard theory. Instead Clark opts for a notation in which sentences are represented as a series of simple "propositions" which can be embedded one inside the other. For example "John is not happy" would be represented as

((John is happy) is false).

Setting aside any possible debate about scope problems here ("John is not happy" may not be equivalent to ""John is happy" is false", depending on how one treats reference failure), it is not clear at what level we are to take this analysis. Clark acknowledges that the notation is adopted primarily for illustrative purposes but, granting that, there seems to be no reason to have a system which has as its base elementary one and two term simple sentences unless one believes these have a real existence at some level. In fact Clark seems to believe that some analogue of them does play a role in the process of comparing sentences against pictures : an analogue to the extent, at least, of having ordered subject and predicates (for the Principle of the Primacy of Functional Relations to work - see below). Johnson-Laird (1974) asserts that the evidence for this is "sketchy and indirect" (p.147), but it would be more accurate to say that it is difficult to see what would count as evidence either for or against this assumption. Let us simply note that this notation is (a) clearly not like any of the types of linguistic deep structure referred to above (with the possible exception of Lakoff's) so that Clark's claim quoted at the beginning of this paragraph is very difficult to interpret (b) unjustified except in the context of its role in the overall model of sentence/picture comparison.

The crux of the model is the assertion that the process of verification is divided into four discrete stages:

1. the representation of the sentence
2. the representation of the picture
3. the comparison of the two representations
4. the production of a response.

(Stages 1 and 2 can be reversed leading to different predictions (see Clark and Chase, 1972)). Clark in facts claims (1974, p.1295) that similar stages are also involved in question answering and instruction following. What is being claimed here is that sentences and pictures are coded in the same kind of format, successively; that the two codes are then compared in an ordered series of mental operations each of which contributes additively to the response latency; that sentence encoding, picture encoding, comparing and responding are serially ordered and their component latencies are additive. Of major interest here is the comparison process which Clark claims can be

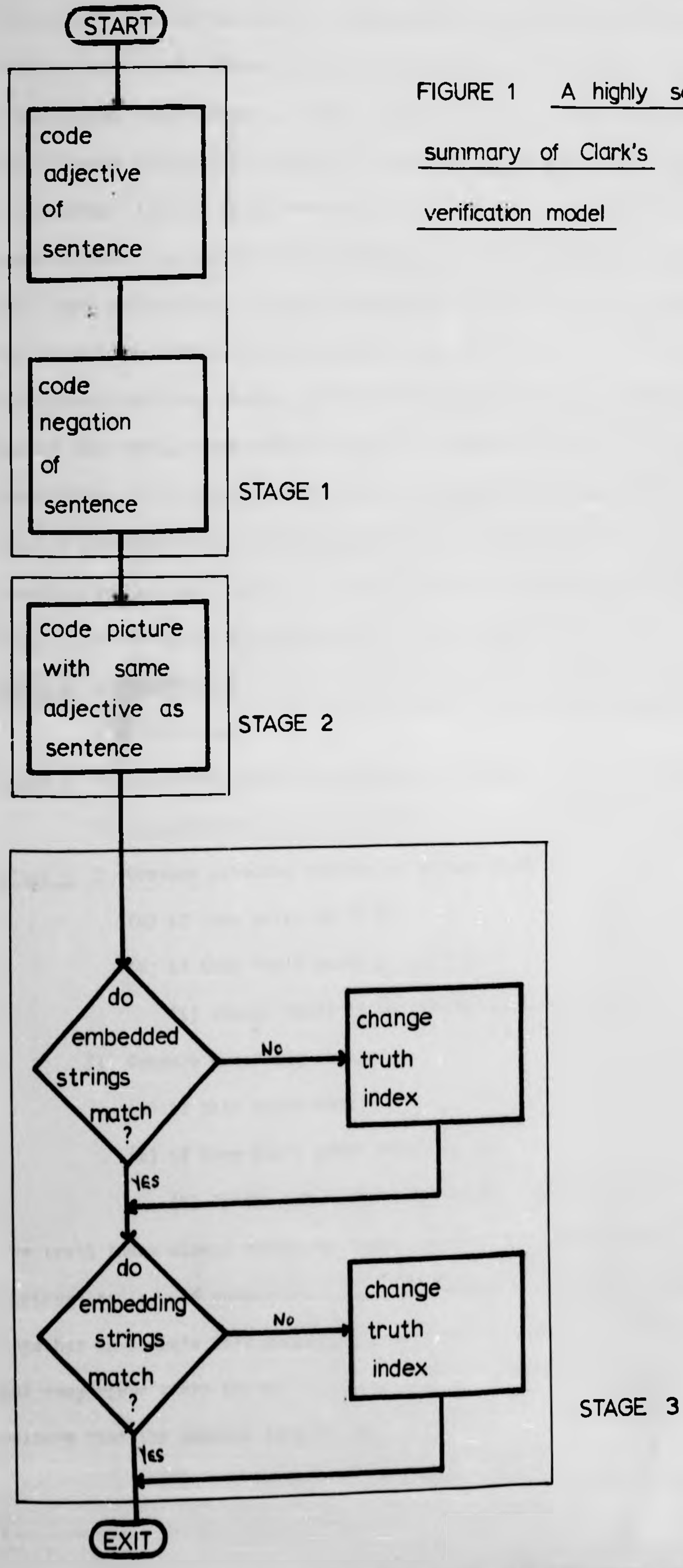


FIGURE 1 A highly schematic summary of Clark's verification model

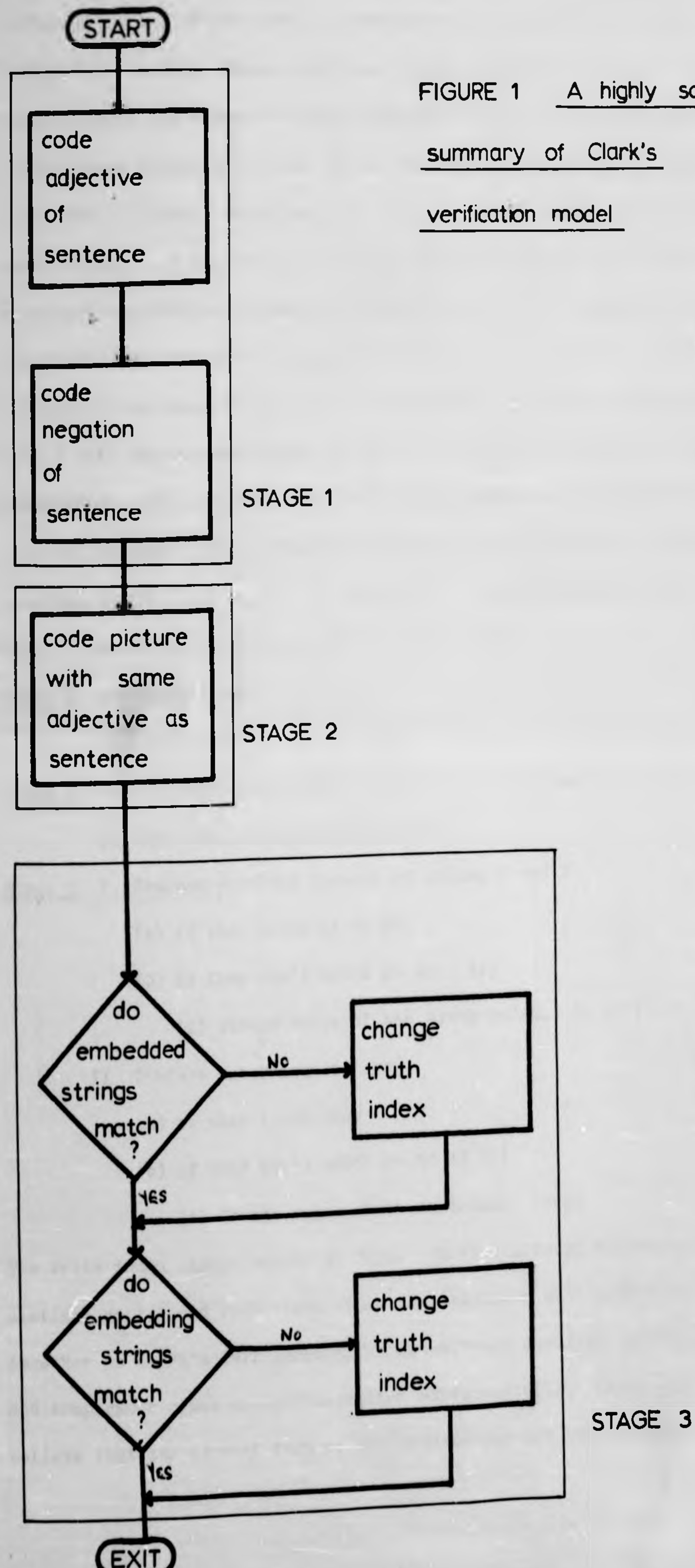


FIGURE 1 A highly schematic summary of Clark's verification model

performed in one of two ways : in accordance with the "true" or the "conversion" model. These models are quite different and lead to different predictions. The conversion model depends on the kind of artificial equivalences referred to above and is accordingly less general than the true model. In many tasks subjects can opt to use one method or the other : some subjects may use one and some the other. This inevitably makes data very hard to interpret unless one asks subjects about their strategies, and treats the resulting two groups separately. Failure to do this explains the rather confusing results in the literature according to Clark (1974). Here I will concentrate solely on the true model because of its greater generality. This model is explicitly designed to cope with negatives as well as premises with contrastive adjectives (such as lexically marked/unmarked pairs - see below). It has the following components (see also Fig. 1 which presents a flow-chart of the model):

<u>Stage 1</u> code adjective

code negation

<u>Stage 2</u> code picture with same adjective as the sentence (negative picture codings are not allowed).

<u>Stage 3</u> I Compare embedded strings of Stages 1 and 2

(a) if they match go to II

(b) if they don't match go to I (i)

(i) change value of the truth index. Go to II.

II Compare embedding strings

(a) if they match stop

(b) if they don't match go to II (i)

(i) Change value of truth index. Stop.

The truth index always starts at TRUE. On this account the conceptually distinct coding and comparison times for negatives will always be merged together in Clark's 1972 paradigm where sentence encoding, picture encoding and comparison times are not separated experimentally. Clark appears to believe that the general form of the process carries over to several other tasks. But he emphasises that special task demands or instructions (as in some experiments by Young and Chase referred to by Clark) may allow or encourage people to use a variety of conversion strategies. These will generally be easier with explicit negatives ("not present") than with implicit negatives ("absent").

Although Clark would seem to view the process of comparing sentences against pictures as essentially a mechanical one, he at times gives the impression that he believes the process can be altered depending on the situation (something which he presumably would want to emphasise in view of his later work e.g. Clark (1973)). For example he acknowledges (p.1331ff) the common use of the negative as a means of denying a prior assumption. Evidence for the greater ease of processing of a negative when it is being used to deny a prior assumption comes from Wason (1965), Greene (1970) and Johnson-Laird and Tridgell (1972) in a paper significantly entitled "When negation is easier than affirmation". But this seems to run counter to the overall pseudo-mechanical nature of the processing stages Clark describes. What is more he later acknowledges the possibility that pictures may be coded negatively if there is a prior expectation of a different state of affairs. This follows work by Olson and Filby (1972) which shows that the picture coding appears to be manipulable in a manner which seems to give the picture coding a voice (i.e. either active or passive). One might well ask what kind of processing of the sentence is taking place if it does <u>not</u> set up an expectation that it will be true : surely we normally expect things said to us to be true? It seems reasonable to suppose that in everyday life we code our perceptions of the world in accord with our prior expectations so that much of the burden of the process of comparison in the Clark model is carried by the state-of-the-world encoding stage. That is to say we do not normally code pictures or states of the world in a relatively context-free form - one unaffected by prior expectations. Even the simplest scene is too complex for this. Instead we selectively "interrogate" the picture on the basis of prior expectations (this is the position of Donaldson, 1974) only coding the relevant aspects. A corollary of this is that the picture is definitive : pictures are never falsified by sentences. But the Clark model fails to recognise this point.

Furthermore the process of finding a sentence false is more complex than noting simply that there is a mismatch - it also involves noting whereabouts the mismatch is (hence the inadequacy of always representing "not" as a blanket sentence negation, rather than as a denial that a particular aspect of the sentence is correct). Sentences are typically divided into a questionable part and an unquestionable part : the first being variously called the "focus", the "assertion" or the "new" part, the second usually called the "presupposition" or by some authors the "given" part. The latter does not normally fall within the scope of a negation though we have linguistic devices available which allow us to avoid any commitment to whereabouts an assertion falls down, and so allows the possibility of any aspect of the sentence being incorrect. For example "John Doe did not kill Richard Roe" allows the possibility that someone killed Richard Roe - but not John Doe (who may or may not exist); that John Doe did something to Richard Roe - but didn't kill him; that John Doe killed someone - but not Richard Roe (who may or may not exist); or a number of other more complex possibilities. These are subtleties which Clark is beginning to tackle in his more recent work, but which he has not yet attempted to cope with in any detail so far as his model of the verification process goes. It seems unlikely that the model can survive in its present form when the phenomena of dialogue and interlocutors' expectations are considered : the complexity of levels of what is taken for granted is too great (Rommetveit, 1967, 1975).

One finding from Clark's early work which is incorporated in the verification model is a phenomenon called the "lexical marking effect". The concept of marking is an exceedingly simple but nevertheless extremely powerful one. Its first application was to phonology where it continues to be a very useful tool both from the linguistic (Chomsky and Halle, 1968) and the psychological (Menyuk, 1971) viewpoint. The concept has been extended to semantics by Bierwisch (1970) and others, and has been shown to have psychological correlates by Herb Clark (1974) and to be of use in understanding semantic development by Eve Clark (1973). Jakobson has suggested that we might also think of the phenomenon as operating in the syntactic domain - for example viewing the active/passive contrast as one of markedness. If this last suggestion can be justified on linguistic grounds - something which is by no means clear (see Greenberg, 1966) - then there is a wealth of evidence to show faster processing of the unmarked (active) form. Since the concept of marking is at the centre of much of what follows, both in this review and in the subsequent experiments, it deserves substantial treatment here.

The linguistic work on marking has been thoroughly reviewed by Greenberg in his monograph on the subject (Greenberg, 1966) to which what follows is heavily indebted. Jakobson distinguished three criteria of marking in lexis ; (1) some words stand for the generic category as well as one member of the subordinate e.g. "man" ; (2) the unmarked form tends to be simpler e.g. "author" vs. "authoress" ; (3) distinctions present in the unmarked member are often absent in the marked e.g. in English pronouns the third person plural does not show the gender distinctions present in the third person singular. Hjelmslev in addition to these three criteria noted five others (some of them not purely lexical) ; (4) neutralisation : in some environments the distinction between marked and unmarked is suppressed and only the unmarked occurs ; (5) marked forms tend to be less morphologically irregular (presumably because of their lower frequency) ; (6) defectivation : this is very similar to syncretisation ((3) above) and refers to the lack of certain categories in marked forms, e.g. the future in the French subjunctive ; (7) dominance : where a heterogeneous collection is referred to the unmarked form is used e.g. the Spanish "los padres" ("parents") and "el hijo y la hija son buenos" (masculine - unmarked - form of the adjective) ; (8) frequency : the marked tends to be much less frequent. Many of these criteria are partially overlapping so that Clark's

(1969) two chief criteria for markedness in dimensional adjectives - namely the use of the unmarked to name the scale and to ask unbiased questions - can be seen to be the result of a number of the processes noted by Hjelmslev. One point of major importance in Greenberg's work is his thesis that the major criterion for markedness in grammar and lexis should be frequency of use. For the case of lexis in particular he notes that the masculine is much more commonly the unmarked form but we still have such obviously marked forms as "male nurse" and "male model". What is involved here giving rise to these forms is the nature of the real world and the way it is interpreted within different cultures, not any innate linguistic or psychological universals. This kind of view contrasts with some suggestions of H. Clark (1973) on the fundamental psychological nature of markedness in some dimensional adjectives as well as with Eve Clark's (1973) ideas on the nature of semantic development.

Support for the association of marking with basic psychological phenomena comes from several points in the psychological literature. Seymour (1969) in a verification task showed an asymmetry in people's ability to judge displays of an object above or below a reference point which he interpreted in terms of a general scanning response. Just and Carpenter (1975) in a rather similar experiment to Seymour's but with the verbal element eliminated (at least not explicitly involved) found much the same result although they attributed it to a general property of semantic coding rather than to a scanning strategy. Their results are rather complex however and the effect appears to vary quite considerably depending on exactly how the information is presented. In contrast to this result Chase and Clark (1971) failed to get Seymour's results when they eliminated the explicit verbal component of the task. Clark at that time believed the marking effect to be essentially linguistic.

Whether marking is a general property of representations or not there would still appear to be plenty of evidence to show that the marking effect is extremely reliable with linguistic materials. In addition to the work mentioned above (especially Clark, 1969) there is quite a lot of other material showing psychological correlates of linguistic marking. For example Hamilton and Deese (1971) have shown that subjects can be made to reliably sort marked from unmarked adjectives, though they suggest that all that may be involved is the evaluation dimension of Osgood's semantic differential (Osgood, Suci and Tannenbaum, 1957). Greenberg (1966) quotes data showing that associations to singular nouns are nearly always singular nouns, and though associations to plurals are usually plurals, they are six times more likely to be singular than associations to singulars are likely to be plural. Further, associations to positive adjectives are almost always positive but associations to comparatives are quite likely not to be comparatives - in fact they have a 0.29 probability of being positives (though they are almost never the more highly marked superlative). Harris (1973) did an experiment in which subjects had to guess the answer to questions involving either the marked or unmarked member of pairs of dimensional adjectives. This was a rather bizarre experiment in that people were simply asked questions like "How (much (little money was in the man's wallet?" and had to make a guess without any other information. However Harris showed that the variance of the guesses was much larger for most unmarked adjectives than for their marked partners (all guesses were in terms of numbers, of course). Harris interprets this result as showing that the unmarked term is being interpreted as asking an unbiased question - that it is simply the superordinate scale name. There was a hint that the exact nature of the materials might bias the results, though, and this was not controlled for in any formal way.

Clark and Clark (1968) studied memory for complex sentences describing time relations by means of either "before" or "after" and either marked or unmarked syntax (subordinate clause first and subordinate clause second respectively). (They also used "but first" and "and then" but these will be ignored here as they were not used in the Smith and McMahon study reported below). Their results showed (1) that people tended to order the clauses temporally so that the clause which came first referred to the event which came first ; (2) that there was a bias towards having the subordinate clause second ; (3) that accuracy for sense was not related to transformational complexity. They suggest that the results are best seen in terms of two marking factors : an unmarked order of mention being with the first event in the first clause, and unmarked syntax being with the main clause first. A tendency to remember the marked forms as unmarked but not the reverse would explain the results. There appears to be no consideration of <u>lexical</u> marking in this paper. If we assume that "before" is unmarked and "after" marked (an assumption which is partially justified by data on order of acquisition and frequency differences - though is by no means cut and dried) then we would get a rather different set of predictions. Consider the four sentences describing a state of affairs where X precedes Y (where X and Y stand for clauses describing events), schematised as 1 - 4 :-

1. Before Y, X	3. After X, Y
2. X before Y	4. Y after X

Now the Clarks' two principles lead to the following predictions (where "$\rightarrow$" stands for "will tend to be remembered as")

(A) Subordinate Clause Second is Unmarked:

1 $\rightarrow$ 2 (or 4)

3 $\rightarrow$ 4 (or 2)

(B) Order of mention the same as order of occurrence is unmarked:

1 $\rightarrow$ 2 (or 3)

4 $\rightarrow$ 3 (or 2)

But lexical marking as laid out above leads us to:-

(C) 3 $\rightarrow$ 1 or 2

4 $\rightarrow$ 1 or 2

A and B together appear to lead to different predictions than any of A plus C, B plus C, or A plus B plus C.

A series of experiments by Smith and M[c]Mahon (1970) casts some doubt on the Clarks' results. They used a question answering technique rather than a memorial or verification method. In their experiments on transitive sentences describing a single event in which two objects were ordered they found three effects consistent across six studies : (1) passive sentences were harder than actives (this, of course, can be interpreted as a syntactic marking effect since actives are more frequent, simpler and more neutral than passives in emphasis); (2) it takes longer to answer if the answer is the patient than if it is the actor; (3) it took subjects longer to respond when the answer was not the leader. Their results with sentences describing two events and of the same form as the Clarks' showed a superiority for main clause answers, for answering "What happened first?" as opposed to "What happened second?", for answering about sentences in which the subordinate clause is first ('contra' the Clarks), and for an order of mention not the same as the order of events in time (again 'contra' the Clarks). These experiments presented the question prior to the sentence. In a subsequent set of experiments they presented the question after the sentence, subjects being allowed to inspect the sentence for as long as they liked (though this was timed). Inspection times and question answering times were therefore measured independently. The inspection times were longer for sentences with "after" (as a marking account might predict : but not in line with the Clark's two principles which indicate that the "before" sentences are both the hardest and the easiest, with the "after" sentences intermediate); they were also longer for sentences in which the first event was in the first clause (again 'contra' the Clarks). Questions again showed more difficulty with "What happened second?" but now show more errors with the sentences beginning with the subordinate clause (a result consistent with the Clarks' results but not with Smith and M[c]Mahon's question-first experiments). Again it took longer to answer with the subordinate clause. A repetition of these results with a disruptive task between the sentence and question produced substantially the same results except that there was no longer any effect of the order of main and subordinate clause in the sentence. This seems to indicate a decay of surface structure information, but not deep structure information. (More rapid decay of surface information is a common enough result (see below)). Smith and M^{c}Mahon also replicated the Clark's original experiment getting a degree of agreement with their results significant beyond the p=0.01 level. Their null hypothesis here was that the replication would fail - which appears rather odd until one remembers that their other experiments suggested that the Clarks' result could not be replicated. Their conclusion is that memorial experiments are unreliable as indicators of linguistic structure (a conclusion shared by Fillenbaum (1973)) though they are clearly baffled about the correct interpretation of their results. They summarise them in terms of five conclusions: (1) the logical subject is more available than the logical object (this result is analogous to that found by Huttenlocher, Eisenberg and Strauss (1968) in their question answering interpretation of their placement task. It is a claim made also by Clark (1974) p.1349. It will recurr in the question answering experiment of the present thesis); (2) passives take longer to process (this is a very common result : see the review of voice effects at the start of the relevant chapter); (3) what is asserted in the main clause is more available than what is asserted in the subordinate clause, though there are no order effects; (4) whatever is asserted to be first in time is more accessible than whatever is asserted to be second; (5) "before" is easier than "after".

Apart from the unequivocal support which the Smith and M^{c}Mahon results give to the notion of marking as applied to voice phenomena in syntax and to the analysis of "before" as simpler than "after", their results are rather difficult to interpret in terms of marking. The order of mention marking is not given any support except by the replication of Clark and Clark; and support for the notion of syntactic marking in terms of the main/subordinate clause order is at best equivocal. A major blow to the notion of lexical marking (and especially to H. Clark's (1970) analysis of associations in terms of feature dropping) comes from some data collected by Brewer and Lichtenstein (1974). They used antonym pairs which had been rated on the basis of the degree of bias induced by the one relative to the other in asking questions (a classic markedness criterion for adjectives and adverbs). They presented sentences with marked and unmarked words in both the affirmative and the negative (negation being, as noted above, another form of marking) and asked their subjects to recall them. They observed significantly more shifts from the marked to the unmarked than in the opposite direction. But 737 of the shifts were meaning preserving, which means two features must have changed (the marking on the adjective as well as negation marker). This is clearly contrary to a theory of feature dropping in memory (as Clark, 1970). When they looked at memory for lists without negatives they found only 8 shifts in 1600 items - and 4 of those were marked to unmarked and 4 the opposite. These results appear to strongly disconfirm the marking theory as applied to memory. Again though, as Brewer and Lichtenstein themselves emphasise, this does not necessarily extend to other tasks, for example verification.

Clark in his review paper summarises a whole body of evidence which he is able to explain by using the markedness notion as an integral part of the more general model. None of that data is dependent upon memorial tasks. It is to that data that we now turn. As with the negation work he assumes a canonical encoding but he again suggests the possibility of alternative codes, for example in stating that "people encode pictures in terms of the figure they have attended to" (1974, p.1344), though acknowledging that the unmarked word will be used if there is no preference. In fact in the 1974 paper Clark has started to use the terms "positive" and "negative" instead of "unmarked" and "marked". This seems to me a mistake as they do not have the same generality. The "in terms of" here seems to mean that the locative phrase (the quotation comes from the section on locatives) will normally contain the reference noun. This is extended in Clark's discussion of the Huttenlocher placement tasks (Huttenlocher, Eisenberg and Strauss, 1968; Huttenlocher and Strauss, 1968) to transitive sentences. The logical object is considered the analogue to the noun in the locative phrase. Passives are considered harder because they make the actor the reference point - something which is at odds with the canonical deep structure. The form of the canonical coding would make it easier to place the actor in a placement task, and is also consistent with the superiority found by Smith and McMahon for the question-answering task when the actor is the answer. It is significant that Clark is here depending upon the ordering assumption noted earlier, of his deep structures (an ordering assumption common also to the deep structures of TG). The principle being embodied here is what Clark calls the "Primacy of Functional Relations" which "asserts simply that functional relations, like those of subject, verb, and direct object, are stored, immediately after comprehension, in a more readily available form than that of other kinds of information, like that of theme" (Clark, 1969, p.388). (the terms "subject", "verb" and "direct object" here refer to the base elements in a TG, not to the surface roles which they are used to refer to by the present author). This claim is one which will be challenged repeatedly in the experiments to be reported in the coming chapters.

It's worth noting that much of the evidence on memory for sentences (not cited by Clark) lends support to his notion of an early encoding of the sentence in a canonical deep form with surface details being rapidly forgotten. Sachs' (1967) work on recognition memory for sentences presented in connected discourse certainly seems to provide some support for Clark's position. She presented sentences for recognition which were either the same as the target or else changed in one of a number of ways - either syntactic or semantic - and measured correct rejections. Test sentences were presented immediately after the target, or else 80 or 160 syllables later. On immediate presentation correct rejections were as high for syntactic or semantic change but after 80 or 160 syllables percentage correct remained high for semantic change while it dropped away for syntactic change. This would be consistent with a model in which memory was said to be in terms of relatively abstract "propositional" chunks (perhaps something akin to a Chomsky (1957) type deep structure or a set of case roles). Nevertheless there is no doubt that a small residual memory for form remains - a result found also by Anderson (1974). Anderson used a hybrid memory and verification task : he presented subjects with passages of connected discourse not dissimilar to those of Sachs and later presented them with a probe sentence which they were to indicate was either true or false of the passage. His results are very complex but they show (amongst other things) effects of delay, the voice of the probe sentence, truth, a delay x probe voice interaction (with actives relatively better after a delay), input voice x probe voice, delay x input voice x probe voice (bigger effect in the immediate condition) and truth x input voice x probe voice (bigger effect with true sentences). The main effects here are all in the standard directions (passive, falsity and delay all impair performance), the input voice x probe voice effect is analogous to the Olson and Filby (1972) result with pictures and sentences. The results as a whole favour a model in which decoding of the surface form occurs at comprehension and most of the material kept in long term memory is of a deep nature.

There is a good deal of evidence to show that more is involved here than a simple depth/surface dichotomy. Indeed the presence of residual surface information after quite long delays in both the Sachs and Anderson experiments indicate something quite complex is occurring. Wright's (1969) experiment shows that even after quite long delays a mismatch in the voice of a sentence and the subsequent question about it, has effects. Begg and Paivio (1969) show that "imageability" of the sentence has a negative effect on verbatim recognition memory. They showed a greater ability to recognise substitution of a word by a synonym with "abstract" sentences than with "concrete" material sentences. The evidence appears insufficient to support their claim for verbatim storage of abstract sentences but some other kind of storage, in which mental images play a critical role, for sentences describing "concrete" material. An experiment by Kennedy (1973) shows that people are better able to reject as not seen associates of the surface subject of the sentence than of the surface object. In this situation subjects are presented with a sentence for memorization and then a series of 40 words. They have to indicate whether the word occurred in the sentence or not. Although Kennedy's results are only for passives he seems to believe that, taken in conjunction with an earlier experiment on actives, they point to the ability to store verbatim material (since the <u>surface</u> subjects are involved). However this result might well be the result of a strategy demanded by this experiment (perhaps a left to right matching scan through an image of the sentence?), and people may well not ordinarily store a verbatim record - or even be able to store one except for the special case of an image of a visually presented sentence.

Two experiments which seem, on the face of it, to support the deep structure trace model are those by Coleman (1965) and James, Thompson and Baldwin (1973). Coleman found a tendency to recode passives as actives, but not the reverse. This is obviously like a markedness effect, but it suggests at the same time a reversion to a more abstract form. That may not sound like a real contrast but in fact it can be. It depends upon one's interpretation of markedness shifting. Clark (1974 also 1969) makes two proposals : in one the sign of a feature is changed, in the other a whole feature is dropped. Consider "tall/short" : "tall" would be represented on the feature theory as [+ height]

[+ polar]

and "short" as [+ height]

[- polar]

but "tall" as in "How tall is John?" as [+ height] without specification as to polarity. Now if "short" is recalled as "tall" more often than the reverse, are we to say the sign of the polar feature has been altered to the less marked form or that the feature has been dropped? The former is a simple markedness effect, the latter more radical : a reversion to a more abstract form. The point is that Coleman's result may be a reversion to a more basic form or it may be a simple markedness effect : the active can signal emphasis on the actor (Johnson-Laird, 1968) or it can be neutral. For Clark's theory the actives really ought to be a manifestation of the neutral form. But the presence of passives suggests that at least some of the actives may also be non-neutral, which tends to go against the conversion to an abstract base model. In any case, as Coleman points out, all that may be involved is a response bias because actives are more common. The James et al.(1973) paper is in the same vein as the Begg and Paivio (1968) one. They attempted to give prominence to either the actor or the patient at encoding by either making it more "imageable" or by presenting it alone prior to the presentation of the sentence. Looking at recall they found more tendency for passives to be recalled as actives than the reverse but they also found a tendency to begin the sentence with the salient term. This suggests a response bias towards actives but also the possibility of coding thematic information which is certainly not in itself "surfacey" but which has some well-defined surface consequences! This result runs parallel to one of Anderson in the study discussed above. He found that passives were verified much more slowly if the grammatical subject of the sentence (i.e. for his set of sentences the first noun) was <u>not</u> the "topic" of the prose surrounding the sentence (he presented the sentence in connected discourse). This was not true for actives. Once again this result supports the notion of the active as unmarked (not subject to the same degree of environmental conditioning) but it does not appear to support the model of canonical form storage in long term memory. Some kind of thematic or topic information seems to be available and its influence is obviously much more important with the passive.

That reduction to a canonical form is not required either for verification or question answering is illustrated in experiments by Olson and Filby (1972) and the Wright (1969) study referred to above. Wright found that surface information appeared to be present after 5 sec. delay in that sentence/question voice matching led to more rapid responses than sentence/question voice mismatching. This demonstrates quite clearly that there is no need to, as it were, detransform a surface structure into the base in order to answer questions. Olson and Filby (1972) came to similar conclusions using a variety of tasks which either foregrounded the actor or the recipient of the action. If one assumes that foregrounding of the actor leads to coding of the picture in some way analogous to the active voice of sentences, and foregrounding of the recipient of the action leads to a coding analogous to the passive, then they found the same kind of result as Wright. If the codings match responses are faster than if they do not. In fairness to Clark he is ready to admit the possibility (in the 1974 paper) of non-canonical encoding if some object is especially foregrounded in the subject's attention, but that position is clearly antipathetic to the principle of the Primacy of Functional Relations described above (Clark, 1969). In particular it plainly contradicts the principle that deep structure functions (or, simply, cases) are primary in the coding process. Instead it emphasises the importance of thematic information (explicitly denied in Clark's (1969) statement of the principle). The importance of topic information is something which will come through repeatedly in the experiments to be reported in succeeding chapters.

Much of Clark's most convincing data on marking comes from work done by himself and others on comparatives. The primary concern is not with comparatives as a form which are themselves marked (as Greenberg, 1966, notes) but rather with the markedness of the positive counterpart of the comparative adjective (or adverb). Of some importance here is Clark's use of a two sentence base structure for comparatives. This is certainly not the only possible analysis (see Campbell and Wales, 1969). However it does fit neatly with Clark's "simple propositional" view of the base structures involved in the verification process. In Clark's notation the base of "John is better than Fred" would be something like "John is good more than Fred is good" or

((John is good +) (Fred is good))

Similarly "Concorde is slower than TSR2 was" would be

((Concorde is slow +) (TSR2 was slow)).

The latter example seems to the present author to show the analysis to be obviously incorrect. What is more the notation as Clark presents it appears to attribute the non-neutral interpretation of the unmarked term to at least one (possibly both) of the "propositions" underlying the surface structure.

However if one accepts his linguistic analysis then the data Clark collected are readily explained by his model. The fact that the model is readily extendable to so called "negative equatives" ("John is not as good as Fred") makes it particularly attractive. The model predicts greater encoding time for marked adjectives and for negatives, with these being additive. This means that "not as bad as" will (on the analogue of the "true" model of negation) be coded with two features more than "better than" even though they are truth conditionally equivalent. One needs also the principle of congruence which states that some sentences are easier than others at the level of functional relations. This means, for the case of comparatives, that questions are easier if the underlying string of the question and sentence is the same : that is, in a word, that the adjective matters more than the "more than" relation (see Clark, 1969).

Clark extends the model to cope with the three term series problem or "linear syllogism" : problems with two premises which altogether mention three objects and two relations, thus : John is taller than Mike.
Fred is shorter than Mike.
Who is shortest)
tallest)?

In problems where both relations are the same then what matters is the marking of the relation : marked premises taking longer. This clearly follows from the lexical marking principle. When the two are different what is said to matter is congruence between the question and the premise which contains the answer. This is so because the output of the process of understanding the two premises is three basic propositional forms, with amalgamation taking place between the forms of the underlying strings containing the noun mentioned in both premises.

Thus in the above problem the first premise analyser produces

((John is tall +) (Mike is tall))

and the second premise analyser

((Fred is short +) (Mike is short)).

The two Mike-premises become "(Mike is middle)" so that the end result is three simple structures. Either the marked or unmarked question is readily answered from this. This kind of premise set is harder than those in which the same term is used in both premises, according to Clark, because of a tendency to lose the second half of the first premise, so necessitating backtracking or "some other time consuming strategy" in order to produce a three proposition structure. Apart from this rather ad hoc assumption the model is elegant and simple and good at accounting for the results. Clark shows quite convincingly (Clark, 1972) that an alternative explanation by Huttenlocher (1969) based on her earlier studies of placement tasks with children (Huttenlocher, Eisenberg and Strauss, 1968) although it works well in these tasks cannot work as a general model for the three term series problem. This is because it fails to work for the negative equative problems. The attempt by Huttenlocher and Higgins (1971) to salvage the theory seems unnecessarily ad hoc. Clark however fails to consider the possibility that different processes may be at work in comparatives and negative equatives and Huttenlocher's suggestion may be one strategy for comparatives. Clark's model has the obvious advantage of greater generality, though.

The latter parts of the Clark (1974) review paper are concerned with voice effects. He suggests a notation for passives very much in the same vein as the notation for the various other types of sentence he considers.

Thus "John hit Fred" would be :

(John did (John hit Fred))

and "Fred was hit by John" would be :

((John hit Fred) happened to Fred).

This notation is interesting in several respects. Primarily because it appears to answer some of the questions about the canonical notation view Clark appeared to hold in his earlier work (and even in the discussion of negation in the 1974 paper!). Topic information is represented here by the embedding strings. There are a number of problems with the notation as it stands. Firstly it would appear to derive truncated passives from full passives by use of a variable in the embedded string. This is objectionable in the first instance because language is undoubtedly used in a way which utilises a given/new structure and it would seem appropriate to have questionable (variable) elements at the highest, rather than the deepest, level in any embedding - that is to say enclosed by the outermost rather than the innermost brackets in Clark's notation. The greater frequency of truncated passives relative to full passives (Svartvik, 1966), and their lower probability of being transformed into actives in memory tasks (Slobin, 1968) strongly indicates a model in which they are treated as a marked/unmarked pair with the truncated passive as the unmarked member. (This leaves open the question of the relative markedness of this contrast taken as a unit, compared to the active). A second objection concerns the form of coding of the active : it appears to leave the active as a marked form since it, like the passive, has both an embedding and an embedded string. If we assume that feature dropping (in this case dropping of embedding strings) occurs randomly then voice information is lost for the active as much as for the passive. More frequent recall of actives is then simply a response bias. This is difficult to interpret because of the ambivalence created by the double role of the active : this is the problem of the two interpretations of feature dropping again. A third objection is to the ordering of passive embedding strings after the embedded strings and actives before : this is unjustified and seemingly arbitrary. Fourthly despite the apparent presence of topic information in this model it cannot be counted as giving a perspicuous account of the topic effects in recall noted several pages back. Clark's model assumes embedded strings are matched before embedding strings, but the topic effect could not then be observed because it would need embedding strings to be matched before embedded strings. Finally it is worth noting that the presence of two strings in both active and passive and the identity of the string which will be used first in any comparison operation (namely the embedded string) seems to make any comparison as easy for the one as for the other. If the sentence is compared with a topicalised picture then the problems are equivalent for both voices (as Olson and Filby, 1972, show). If it is compared with a neutrally coded picture then presumably the picture coding will consist only of the embedded string (since, 'ex hypothesi', there is no topic information) so that both voices will be equally easy. It is plain that Clark would not wish to make quite this prediction but it is hard to see how he can avoid it on the basis of this notation. (It is not a prediction which will readily stand up to the data - e.g. that of Gough, 1966).

It's worth noting that the notation gives a simple solution to the problem of surface matching in the Wright question answering task and related question answering problems. It also correctly predicts for Wright's experiment that the interaction of sentence voice and question voice will not occur for verb questions (e.g. "What happened to Fred?" "What was done by John?") because of the nature of the matching process.

Given the overall emphasis of the Clark (1974) paper - and even more so of those which preceded it - it comes as something of a surprise to find Clark saying near the end of that paper (p.1408) : "One could conclude that actives and passives each have their own important place in the language, and when the proper conditions prevail, actives are easier than passives, or passives are easier than actives. It is just that actives are probably appropriate in a wider range or a more common set of contexts". This statement plainly ignores the fact that, if we are to regard actives and (full) passives as unmarked and (highly) marked respectively then the term "actives" here covers a distinction of major theoretical interest, namely that between the "markless" superordinate and the unmarked subordinate. In fact it may well be that the active is itself marked with respect to some other form (Halliday (1967) and others have suggested that the least marked case is what he calls the "middle" form - forms like "Susan washed/marched/seemed happy"). One does not have to be committed to that view to appreciate that here are genuine questions, questions that are being ignored in the search for more basic regularities, but at the same time prevent us from seeing those regularities. The fundamental operations may well be as Clark suggests they are, but it is impossible at the present time to either assent or dissent until we know more about the parameters of the processes governing what counts as, for example, a situation conducive to the passive, or one conducive to the active. Halliday (1970) points out that a marking contrast is one where the unmarked form is chosen unless there is a "good reason" to choose the marked. What constitutes a good reason? The view taken in the present work, and for which support will be given in the experiments as well as in the presentation of a particular model of language later in this Introduction, is that, for a small set of syntactic options in English "good reasons" are reasons of prominence, cohesion with prior discourse, and information structure.

That people tend to organise information into cohesive semantic structures can hardly be doubted at the introspective level : we are all aware of having a view of a coherent world, of a coherent life history for oneself and so on. There has been much debate in the psychological literature in recent years as to how this coherence is represented in our minds. The commonest view seems to be that our memories are essentially a set of "propositions" (a term used very loosely amongst psychologists) connected by labelled relations (Anderson and Bower, 1973; Kintsch, 1972). Objections to the view that knowing is always "knowing that", to use a slightly different distinction, (Ryle, 1949) have not always been fully understood. There are undoubtedly exceptions : Winograd's (1972) language understanding programme was widely acclaimed and one of its features is the assimilation of aspects of "knowing that" to "knowing how" by the use of procedural representations of data. However the view of remembering as, Ebbinghaus apart, a process of adding propositions and linking them together and of recall as simply locating the right node in the network and reading out the contents, remains a very powerful one in much psychological theorising. It is a view which is easily associated with the kind of view Clark apparently once held about the coding of linguistic information - a view which has been amply documented and criticised above.

The work of Bartlett over forty years ago provided plenty of evidence that storage and retrieval were dynamic, integrative and reconstructive. But the "paired associate model" of memory as a series of atoms with or without links between them, has always tended to dominate over the Bartlettian view in psychology, and it is it which gives credence to the "propositions" and (labelled) links approach. If propositions are stored then one would imagine this should be seen most easily if one looks at recall and recognition of simple sentences. If one can show a failure to memorise material of this sort in discrete chunks then there does not seem to be any possibility of upholding a "propositional" view. That was the strategy of Bransford and Franks and their co-workers (Bransford and Franks, 1973; Barclay, 1973). Their results are too well known to be detailed here. It is sufficient to note that their Bartlettian assumptions proved well founded : even using very simple sentences people appear to readily integrate them. The results are consistent with a model in which people try to integrate material into a single semantic structure (or as small a set as possible). They seem to lose most of the surface information and are unable to tell with great accuracy what sentences were presented to them, so long as these are consistent with the assumed model. Of course in the strictest sense the results are consistent with a model in which memory consists of a set of propositions some of which are linked and some not (a set of propositions with a set of relations defined over them constitute a structure). But the fact that the supposed atoms of this kind of structure (input sentences) are not recognised as well as more complex structures (sentences representing the information from several input sentences) argues against the simple proposition plus relation approach. The exact behaviour observed varies in a number of ways depending upon the instructions given to the subjects, the nature of the task (recognition and recall have been studied), the "abstractness" of the material (cf. Begg and Paivio, 1968), and the possibilities for inference in the material (Bransford and Franks, 1973). It is clear in all their many studies that integration regularly takes place and that other information may be added to what is explicitly presented if the person's world knowledge permits these additional inferences. Many of the results provide additional support for the studies on verbatim recall of verbal material cited earlier (e.g. Begg and Paivio, 1968; Anderson, 1974; Slobin, 1968).

However the Bransford studies can be faulted in one fundamental way : they do not point towards what it is in the nature of the material which encourages or discourages the integration process. The only structural parameter they seem to have investigated is voice and here they essentially only confirmed the conclusions on differences between full and truncated passives which Slobin had arrived at earlier. Bransford and Franks (1973, p.244-245) suggest that more complex syntactic structures may well not be more complex in a paragraph context and that they may serve to give the material more cohesion, but no systematic work has yet been presented on that subject. The basic problem which work of this sort faces is the lack of any model to classify linguistic phenomena on the basis of their contribution to cohesion. Frederiksen (1975 a) working very much in the same tradition as Bransford did an experiment in which he contrasted the "constructivist" view of himself and Bransford with what he calls the "interpretivist" views of Anderson and Bower (1973). The Bransford interpretation states that integration occurs at input, the view of Anderson and Bower that it is largely a retrieval phenomenon, a method of filling in gaps due to forgetting. So, he argues, repeated presentation ought to reduce the amount of inferred information present at recall on the Anderson and Bower view while his view would predict no effect. The results show no difference between a memory only and a memory plus problem solving group, with a big difference between these and a problem solving plus incidental memory group, in terms of a number of measures of what is recalled. Although the third group produced much less verbatim recall their memory for the "concepts and relations" of the original material was better than the other two groups. Subject to a "caveat" about the scoring procedure and the fact that the rejection of the Anderson and Bower model rests partly on a failure to reject the null hypothesis "vis-a-vis" groups one and two, the results seem to support Bransford rather than Anderson and Bower.

The primary difficulty with Frederiksen's results is that the scoring procedure is not as explicitly formal as one would like. This is partly remedied in a very long theoretical paper (Frederiksen, 1975 b) in which Frederiksen presents a detailed model of a logical and semantic network, much of it based on case relations, together with the outlines of a scoring procedure for checking on the accuracy of recall of the semantic information in a presented text. Although he claims that the system is essentially not tied to language, except indirectly (since both purport to represent the world), the model has a very strongly linguistic flavour. Given that, it is a pity that many of the distinctions made are not justified on linguistic grounds. The model is also more complex than would be necessary if the constraints on binary branching and one choice at a time were removed. The linguistic model to be presented shortly (systemic grammar) does not have these constraints.

Models like Frederiksen's (a related model has been suggested by Schank (1972)) clearly are of some interest and they promise to be a fruitful stimulus in the future in terms of the basic elements which they suggest. At present they lack justification either on logical or empirical grounds : their primary justification is that they are implementable on machines, and so the approach is rather pragmatic and eclectic. In consequence the light which they throw on the actual process of integrating a text into a coherent semantic structure is oblique. They are designed to show how it might be done, not how it is done.

A slightly more empirical approach comes from linguists and philosophers who have studied so called "text grammars" (van Dijk, 1973). These are grammars for texts which have as a subcomponent the kinds of grammars which we are familiar with from the work of Chomsky, Fillmore and Lakoff. Since they start from some of the observed phenomena such as noun phrase definiteness, anaphoric relationships, tensing, use of sentential adverbs and so on, they are of more direct interest. However at the moment they are very much in their infancy. The literature on the subject, though it attempts to use some of the apparatus available in advanced logics (such as modal logic), tends very much to ape the technical apparatus of sentence grammars, mainly of the Chomskyan variety. Furthermore since they are in the Chomskyan tradition the grammar is seen as generative in the traditional sense. That is it is a method of distinguishing "grammatical" from "ungrammatical" structures - essentially a decision procedure for well formed formulae. Such a notion seems much too impoverished to support a theory of how the psychological process of sentence or discourse generation occurs (Watt, 1970; Derwing, 1973).

A deep rooted shortcoming in all the discussions of cohesion from the point of view of TG is the failure to consider the meaning which might be attributable to syntactic choices. It is clear to everyone that the choice of one form rather than another is meaningful in the very weak sense of transmitting "information" (in the technical sense of that word). But very little attempt has been made within TG to explain what the function is of these various possibilities of conveying what is apparently the same content. This attitude can be traced back to Chomsky's concluding statements in Syntactic Structures (p.108) : "The notion of "structural meaning" as opposed to "lexical meaning", however, appears to be quite suspect, and it is questionable that the grammatical devices available in language are used consistently enough so that meaning can be assigned to them directly". Although he goes on to suggest that such correlations as there are "quite naturally" between semantics and syntax could be the subject of a "more general theory of language", later work in TG has hardly touched upon this whole area. The only phenomenon of cohesion to be treated in any depth at all seems to have been pronominalisation - and this was examined not from the point of view of cohesion so much as from the viewpoint of content : the interpretation of anaphoric pronoun reference in complex sentences can be ambiguous and hence involve difficulties of interpretation of the content of the sentence.

One grammatical theory which is both prepared to attribute meaning to grammatical choices and which has a highly developed theoretical apparatus for dealing with cohesion is "systemic grammar". Since a good deal of the interpretation of experimental results is given in terms of this theory a fairly long summary follows.

Systemic grammar (SG), as the name suggests, is a theory of English grammar whose goals embrace description of the choice structure of English syntax and the way this is organised into a system. It stresses the importance of accounting for the choice of one option rather than another, asking questions like "What kinds of purpose does this choice serve which that choice would not have served?". Indeed early forms of the theory were almost exclusively concerned with the choice structure without paying much attention to how the systemic options could be realised. The theory was capable of classifying sentences on the basis of the options chosen and in that respect was enlightening but there was no real attempt to produce an apparatus of realisational machinery. Although attempts were made in the very early stages to formalise the theory (e.g. Halliday, 1961) genuinely formal structures did not appear until ten years later with the publication of Hudson's English Complex Sentences and Winograd's Ph.D. thesis on computer processing of natural English which utilised a parser based on a fragment of SG (Winograd, 1973). In fact most of the early publications are by only one person (Halliday, 1961, 1963, 1966, 1967-68), although many of the notions used are derived from earlier members of the London school such as Firth and Malinowski (see Langendoen, 1968) as well as the functional approach of the Prague school (Mathesius, 1975; Danes, 1964). Halliday, despite the breadth of his interests, (e.g. 1973, 1974, 1975) has written very little on the mechanics of the realisation process. Although the main interest of SG for the purposes of the present work lies in the choice structure it is as well to get the role of this element of the grammar in perspective. For that reason I will first give an overview of the theory as presented by Hudson (q.v. for a more detailed account). Though Hudson's book contains much which Halliday has not written on he in fact claims that "in most respects this version of systemic theory represents the current views of Michael Halliday" (1971, p.vii).

The theory is based on a division of language into four (reasonably) distinct domains : semantics, phonology, grammar and lexis - though some writers, including Halliday, tend to merge the last two together. The grammar constitutes a formal object which can be used to assign structural descriptions at both deep and surface levels to strings of the language. Although the theory is primarily directed at English much of it is intended to be applicable to other languages also. The grammar uses neither phrase structure rules nor transformations, and it allows surface and deep units to be in the same order, generally speaking. Syntagmatic dependencies between immediate constituents of an item are left implicit because they can be deduced from the paradigmatic characteristics of the constitute which contains them. For example the relationship between "have" and "-en" in the English perfect which is made explicit in TG by having a phrase structure rule introduce them together and then a transformational rule position them correctly, is made explicit in SG by having them manifest the single paradigmatic feature (perfect). This ability to cope with discontinuous items in a simple fashion is a major advantage of having the same order in deep and surface structure. Moreover it avoids any possibility of having to have relationships between both deep structure and semantics and surface structure and semantics - something which has been suggested in recent years as a modification to the standard theory (see Jackendoff, 1972).

A characteristic feature of SG is the shallowness of its tree-diagrams. This is because it utilises very heavily the concept of a group (noun group, adjective group, verb group etc.) which may contain many ordered elements. For example the structure of a noun group (according to Winograd, 1972) is as follows:-

Determiner Ordinal Numeral Adjective* Classifier* Noun Qualifier*

(where * denotes the fact that more than one of these may be present - otherwise only one can occur). Winograd gives the following example:

Det.	Ord.	Num.	Adj.	Adj.	Clas.	Clas.	Noun
the	first	three	old	red	city	fire	hydrants

Qual. (Preposition group) Qual. (clause)

without covers You can find.

This example illustrates a simple but quite powerful feature of SG : namely "rank shift". This term refers to the possibility of having units of a higher rank (in this case clause) shifted down to operate within units of a lower rank (in this case noun group). This facility greatly simplifies the derivation process.

One very important part of the formal apparatus is the description of syntagmatic relations in terms of a large number of functions. Functions in SG can be combined into bundles so that a single constituent can carry several functions. Furthermore functions can enter into quite complex relations : one term in a structure building rule (see below) is always a function, others may be functions or may be a characteristic of the constitute (e.g. that it is an interrogative clause). Hudson comments (p.39) : "The functions that are set up are often rather unusual, if for no other reason because they have little or no direct connection with meaning; and there are far more of them, both in the grammar itself and in the description of any item, than is usual in other kinds of grammar. Functions have a very important part to play in the grammar, as mediators of environmental conditioning of all kinds, often bringing together under one category a number of different and rather complex environmental influences." They can thus be seen to be quite different from the feature system : in their arbitrary nature, in their lack of a direct connection with meaning, and in the complexity of the environmental conditioning which they handle.

Paradigmatic relations are relations in the system network. The selection of some features is dependent upon the selection of others, in other cases there may be a dependence on the selection of more than one other feature, while still other choices are relatively independent. Thus if a constituent is a verb it may be finite or non-finite, and if the former then past or present. Simultaneous with this set of choices is the dichotomy between grammatical and lexical verbs and if the latter copular, transitive or intransitive. This partially illustrates the kind of choice possible but the system is much more complex than this. Entry conditions for a set of choices can be simple as in the case of the verb feature just mentioned or they can be the intersection of two features (thus given "third person" and singular" there is the choice of "masculine", "feminine" or "neuter" in the English pronoun system), or the disjunction of two features (thus either "indicative" or "dependent" gives rise to the choice of "declarative" or "interrogative", though the system is such that the "indicative"/"imperative" choice can only be made given "independent" - these are choices at the clause level of course). A fact of major importance is that this kind of feature choice occurs at all levels in the language : clause, group and word. (Note also that the system which is entered can be as complex as the entry conditions : thus given "verb" one has the choice of either finite or non-finite <u>and</u> either grammatical or lexical. Recursion (i.e. a feature serving as its own entry condition) is also possible).[2,3]

One probably gets a feeling for the system most easily by considering the process of derivation. Although the grammar does not of itself imply a temporal order we shall assume one both because of the greater ease of expression this allows and also because of the importance of having a temporal component in any psychological interpretation of the model. (It cannot be emphasised too strongly, though, that the kind of temporal order suggested here is in no way a part of SG - See Hudson, p.81). The initial choice as to structure is in the system network. Here the clause system is entered and a set of features selected within the constraints of the system network e.g. [clause, independent, indicative, interrogative, non-polar, Wh-, subject

2. This account is extremely sketchy due to space limitations. Hudson (1971), from which the present account is almost entirely derived, gives a much more detailed description together with more concrete examples.

3. Hereafter, following Hudson, paradigmatic features are written in square brackets and small type and syntagmatic functions are written without brackets and in large type.

focus, transitive, passive, actor unspecified].[4] This set of features is then passed on to the feature realisation rules. These rewrite the features as a set of functions, thus:

clause	independent	indicative	interrogative	non-polar
↓	↓	↓	↓	↓
+ PROCESS	∅	+ SUBJECT	+ MOOD FOCUS	∅
		+ FINITE	(in the environment	
			[independent])	

Wh-	subject focus	transitive	passive	actor unspecified
↓	↓	↓	↓	↓
+ QUESTION	QUESTION	+ GOAL	GOAL =	?
= MOOD FOCUS	= SUBJECT	+ ACTOR	SUBJECT	
(in the	(given [Wh-])	(unless [actor	+ PASSIVE	
environment		unspecified])		
[independent])		+ TRANSITIVE		

This set of functions then passes through a set of unordered structure building rules which serves to order the functions and conflate those which are to be conflated, as well as add some additional functions. The output of these would in the present case be:

QUESTION	→	TRANSITIVE
MOOD FOCUS		PROCESS
SUBJECT		-EN
GOAL		PASSIVE
FINITE		

(Where "→" means "is to the left of"). This set of functions now passes into a set of function realisation rules which interpret the constituents indicated by the functions in terms of features. This set of features now constrains what may be selected for the lower units (in this case the group is next lowest). There will still of course be some measure of choice left

4. Features of one unit - here a clause - are represented within one set of brackets.

at the lower unit in the usual case and so it will add further information which will again serve to constrain choices at the next lowest level and so on to other levels of the linguistic description (e.g. phonology). It is of interest to note that the postulation of structure building rules entails a claim that there are levels of language which are purely structural : i.e. not determined by the meaningful options in the system network. But the theory, by making the systemic network the heart of the model and by having structure building rules as a separate level, emphasises the role of the choice structure in the language as well as the relatively superficial nature of the final structure. I take it that these are desirable consequences both from the linguistic and the psychological viewpoints. The main area of interest is clearly the system network : indeed the clause system network is the only aspect of the theory which is much discussed in what follows.

Halliday in a variety of discussions (1967, 1970, 1973) repeatedly divides up the clause into three components of meaning : the "ideational", the "interpersonal" and the "textual". These are labels which cover groups of choices in the system network : they are all primary in the sense that they are sets of choices which are entered simultaneously (to use the time metaphor again), but the first two are more basic in the sense that they are the real meaning options : the textual choices are subsidiary in that they concern choices over the means of communication rather than the content. The division between interpersonal and ideational functions of language is very deep rooted according to Halliday. He has suggested (Halliday, 1975) that there may be a basic dichotomy between the "mathetic" and the "pragmatic" functions of language : the former covering uses of language to classify and make sense of the world, the latter being its use to get on in the world - to obtain things, interact with people and so on. This view has interesting relationships with observations by Nelson (1975) of two styles of mother-infant interaction at an early stage in language development : what she calls a "referential" (or naming) style and an "expressive" (or interactional) style. Bruner (1973, 1975) has also suggested this may be an important distinction though he uses a framework of locutionary and illocutionary force (Austin, 1962) - albeit in a rather loose manner. The ideational function is the function of language as a means of representing and conveying states of the world. It includes the whole set of options in the systemic network which are concerned with the choice of the roles to be represented in the clause. This covers such options as whether the clause will be transitive or intransitive and if the first a material or mental process clause or a relational or verbal (quoting, reporting) clause. These choices decide what the basic "participant roles" will be in the clause : for Halliday these participant roles are case roles (see Halliday, 1967; Anderson, 1971). The particular set of case roles involved is decided by the set of features chosen for the clause : only the process itself and the affected role (Halliday, like Anderson, proposes an ergative structure) are selected automatically (see Halliday, 1973, p.40).

Parenthetically it's worth noting the distinctions Halliday makes with regard to voice options : this seems the right place to do that as voice is considered an option within the transitivity subcomponent of the ideational option (with reservations which will be expressed in discussing the textual component). Halliday recognises three voice options : middle, active and passive, middle being the unmarked option in the first instance but, given non-middle active being the unmarked choice between active and passive. In Halliday (1970) he gives the choices in the following table:

clause voice		roles	verb voice	example
middle		actor	active	the gazebo has collapsed.
non-middle	'active'	actor, goal	active	the Council are selling the gazebo
	'active'	actor (goal)	active	the Council won't sell.
	'passive'	goal	active	the gazebo won't sell.
	'passive'	goal actor	passive	the gazebo has been sold by the Council

clause voice		roles	verb voice	example
non-middle (cont'd)	'passive'	goal (actor)	passive	the gazebo has been sold.

Notice here that the verb and the clause voice are not necessarily coincident : the expressive resources of English and the underlying choice structure are not related in a straightforward way (this is where feature realisation and structure building rules come in). Even if one were to omit the active with goal suppressed and the passive clause, active verb type, as being options restricted to a small set of verbs, one still has a more complex structure than the standard marked/unmarked account allows. In actual fact the situation is made even more complex by the fact that the transitivity system in English is in a transitional phase between a transitive and an ergative type. The latter, which includes a compulsory "affected" role and an optional "causer" is more readily generalisable to non-material process clauses (see Halliday 1970, p.158 who notes the transition from "methinks" and "it likes me" to "I think" and "I like" : the actor/goal terminology is clearly inappropriate for the modern form). Having noted these problems we will, in what follows, mostly assume the following choice structure for the sake of simplicity (although it is obviously inadequate):

- Voice
  - middle*
  - non-middle
    - active*
    - passive
      - agent unspecified*
      - agent specified

where the "[" brackets denote "either/or" and "*" denotes the unmarked option. The reasons for this configuration cannot all be given here : the middle is adopted as the least marked action clause because of its necessity in an ergative account (see Halliday, 1967a) and because of its similarity to attributives (the "derivation" of adjectives from verbs is of course a subsidiary motive here - see Lyons, 1968); the choice of the active as unmarked "vis-a-vis" the passive I take it needs no justification; the justification for the agent unspecified form being the unmarked passive is its much greater frequency (see Svartvik, 1966), its greater simplicity (the lack of deletion operations in SG makes the "truncated" passive simpler rather than more complex as in TG), as well as such secondary considerations as its lesser likelihood of being transformed into the active (Slobin, 1968).

All this is part of the ideational component of the system network : if we are looking for a "logical" component of language then it is here we must look. The set of cases (plus the verb) approximates quite closely to current ideas about propositions (and to Clark's notion of a deep structure in its dependence upon functional relations which are equivalent to Clark's for the purposes for which he uses them). There is more content in a sentence than the set of case roles though. Of considerable importance are the options of mood and modality which Halliday brings together in the interpersonal component. Mood is obviously interpersonal : it relates in fairly direct ways to the speech act (Searle, 1969) which is being performed. Modality (the use of adverbs like "possibly", "probably", "perhaps" as well as modal verb forms like "will", "would", "can", "ought to" etc.) is less obviously interpersonal. It clearly involves a modification of the propositional content and as such might be thought of as logical also (see Hintikka, 1969). However it is clearly necessary to treat it separately from case information - something all linguists would agree on (see Fillmore, 1968; Wilson, 1972), and, as Halliday (1970) has argued, there are good reasons to treat mood and modality together.[5]

5. The problem here is not so much to show the similarities between mood and modality as to avoid having to include mental process expressions which one wants to keep firmly in the ideational component. All of them seem to involve a process of commitment which has obvious interpersonal elements.

However we treat the "interpersonal" component it is quite clear that the third major "macro-function" (Halliday, 1973), the textual function, is a separate set of options. It is concerned, in Halliday's terms, with the organisation of the clause as a message. In early work (Halliday 1967, pt.2; 1967 unpublished) Halliday distinguished three dimensions of message structure (see also Rommetveit, 1968 : known-unknown, given-new, and theme-rheme. The first of these appears to have become subsumed in later work under the last two. In any case it is a dimension primarily to do with identification clauses (e.g. "The leader is John") and since this is a type we shall not be concerned with here I will set it aside. With regard to the other two dimensions of textual structure it should be emphasised that these are choices as to the organisation of the clause which are independent, to a very large extent, of the choices made in the ideational or interpersonal functions. There are certainly some restrictions due to these other options : for example the voice dimension in transitive clauses gives scope for the organisation of material not present in relational clauses, similarly the nature of questions notably restricts the options realisable in terms of message structure. Nevertheless these restrictions are present mainly in the structure building rules and the options available in the system network remain essentially the same for all clause types. This is an important point as we shall want to apply given-new and theme-rheme restrictions to both transitive and relational clauses in the experiments which follow.

Theme-rheme is a distinction which derives from the work of the Prague school on functional sentence perspective (Mathesius, 1975; Danes, 1964; Vachek, 1966). It involves a recognition that first position in English is a special role in the clause. Halliday expresses this by saying that what is involved in first position is the speaker's point of departure for the sentence. We are certainly aware that first position is special although our intuitions here are confused by the fact that in the least marked case theme and given coincide. But we are dealing here with two quite distinct roles : "The difference may be summed up in the observation that, in dialogue, 'given' means 'what you were talking about' while 'theme' means 'what I am talking about'; and, as is well known, the two do not necessarily coincide". (1967, unpublished, p.9). We can think of theme options as being marked and unmarked but with the unmarked case being dependent on mood. "The unmarked theme is the subject in a declarative clause, the Wh- element in a Wh- interrogative and the finite verbal element in a polar interrogative. Any clause in which the element so designated does not occur initially is said to have marked theme". (1967, unpublished, p.10). Halliday points out that it is much more unusual to have a marked theme in the interrogative where, naturally enough, the speaker's main interest is in the questioned element (be it the Wh- item or the modal in a polar interrogative) than in the indicative "where the subject is merely a way of getting off the ground" (ibid, p.10). There has been some suggestion in recent years that it may be necessary to introduce the possibility of complex themes : perhaps a modal theme, a discourse theme and a clause theme (Martin Davies, personal communication). This point will be touched upon in dealing with questions in the introduction to that chapter, but for the moment we will merely talk of theme 'simpliciter'. The rheme is merely that part of the sentence other than the theme. (Note that theme-rheme here does not mean the same thing as it does in the works of the Prague school).

Given-new refers to the information structure in the clause. Given is what the speaker takes to be information which the hearer already knows or which he believes is readily recoverable from the context. New is what the speaker wishes to convey. The distinction is expounded or realised in intonation : the new element being the carrier of heaviest stress. In the unmarked case units not carrying stress are technically unspecified as to given and new structure but we may take it that they are usually given. In an unmarked indicative clause the new element will be post-verbal (or verbal and post verbal - the "predicate" on some TG accounts - though with two quite distinct roles : "The difference may be summed up in the observation that, in dialogue, 'given' means 'what you were talking about' while 'theme' means 'what I am talking about'; and, as is well known, the two do not necessarily coincide". (1967, unpublished, p.9). We can think of theme options as being marked and unmarked but with the unmarked case being dependent on mood. "The unmarked theme is the subject in a declarative clause, the Wh- element in a Wh- interrogative and the finite verbal element in a polar interrogative. Any clause in which the element so designated does not occur initially is said to have marked theme". (1967, unpublished, p.10). Halliday points out that it is much more unusual to have a marked theme in the interrogative where, naturally enough, the speaker's main interest is in the questioned element (be it the Wh- item or the modal in a polar interrogative) than in the indicative "where the subject is merely a way of getting off the ground" (ibid, p.10). There has been some suggestion in recent years that it may be necessary to introduce the possibility of complex themes : perhaps a modal theme, a discourse theme and a clause theme (Martin Davies, personal communication). This point will be touched upon in dealing with questions in the introduction to that chapter, but for the moment we will merely talk of theme 'simpliciter'. The rheme is merely that part of the sentence other than the theme. (Note that theme-rheme here does not mean the same thing as it does in the works of the Prague school).

Given-new refers to the information structure in the clause. Given is what the speaker takes to be information which the hearer already knows or which he believes is readily recoverable from the context. New is what the speaker wishes to convey. The distinction is expounded or realised in intonation : the new element being the carrier of heaviest stress. In the unmarked case units not carrying stress are technically unspecified as to given and new structure but we may take it that they are usually given. In an unmarked indicative clause the new element will be post-verbal (or verbal and post verbal - the "predicate" on some TG accounts - though "predicate" is a complex notion; see below and Sandmann, 1952). However, the lack of specification as to given and new of the other items in an unmarked clause allows the possibility that the whole of the clause is new. In the marked case there may be more than one information unit (one information unit = one tone group)[6] in the clause. If the information structure is marked : either by the focus falling on an element other than the final lexical item or by the presence of more than one stressed unit, then all other information in the sentence is to be taken as given.[7] Particular applications of the given-new and theme rheme distinctions will be given in discussion of the results of several of the experiments.

The richness of the multiple feature-multiple function view can be seen very clearly in Halliday's discussion of the notion of the "subject of a sentence". Traditionally linguists have felt obliged to distinguish several "subjects" in the sentence principally because of the recognition that the grammatical subject is not always the same as the "logical subject" of the action (the distinction is of most value in transitive clauses, particularly in describing the difference between passives and their corresponding actives). Some scholars also felt the need to distinguish the topic of the sentence from both these two because in some sentences what is the primary focus of attention is neither actor nor the grammatical subject. Halliday extends this multi-subject approach to encompass four distinct notions, each of which corresponds to a function or role (functions, remember are the output of the feature realisation rules - we

6. This is not quite right : the presence of "silent stress" makes the situation more complicated. See Halliday (1967 Pt.2) and Halliday (1967 b).

7. For additional information 'vide' Halliday references and Grimes (1975) Chs. 19 and 21 which provide an excellent summary of Halliday's discussions on textual structure together with a proposal that we introduce a role of "Highlighted" element. This is not needed for English, however.

are not here talking of such macro-functions as the interpersonal and textual components). Taking first the logical subject : TG may here be thought to have an advantage over SG in that the logical subject is definable very straightforwardly as the subject of the base structure (this, of course, is one of Clark's functional relations). This allows us to have a unitary characterisation of the whole notion : one which covers attributive sentences as well as transitives. It is not possible to produce such a simple characterisation in SG since the logical subject may be either in the agentive or the affected case, if the former is not present. However there would not seem to be any point in distinguishing logical from grammatical subject were it not for passives since otherwise they are always the same. Consequently this does not seem much of a disadvantage - especially in view of the fact that the term "logical subject" seems a misnomer anyway since there is no 'a priori' reason for distinguishing one term rather than another as argu ment rather than part of the predicate (with one place predicates). Strictly speaking the logical subject is whatever we choose to call the logical subject[8] and the traditional use of the term merely serves to pick out what is, as it were, the most active object referred to in the sentence. But this role has no independent logical status. Indeed one might reasonably expect it to be a function of the kind of process and not independently definable - which is precisely the state of affairs in SG.

On the notion "grammatical subject" Halliday comments (1970, p.160) "The notion 'grammatical subject' by itself is strange, since it implies a structural function whose only purpose is to define a structural function". But, he says, just as the logical subject is part of the transitivity system, so the grammatical subject is part of the mood system. The grammatical subject is that item in agreement with the verb. The two

8. This view of the matter is derived from Geach's view of the nature of singular terms and predication and is expounded in his Logic Matters (1973)

together are the primary locus of mood (and modality) options in English. These two form a unit which serves to expound mood and modality choices by such phenomena as subject omission (in the imperative) and subject-verb inversion in questions. This is all part of the interpersonal function of defining the communication rôle adopted by the speaker (Halliday, 1970, p.160).

On Halliday's analysis the psychological subject is itself said to be a complex notion, both parts of which are aspects of the textual macrofunction. The first of these is theme - what the speaker is primarily focussing on - "the peg on which the message is hung" (1970, p.161). The second of these is "given" : what the speaker takes the hearer to know already (or be able to infer) and hence the starting point of the message from the speaker's view of the hearer's point of view. Halliday stresses the fact that these two notions are quite distinct and, though they typically are expounded (realised) by the same surface item, the notion "psychological subject" (like that of "topic") is really a complex one.

The point of most relevance for the earlier discussion of psycholinguistic research, as well as for the research to be presented, is that all of these functions are subject to marking principles. In the least marked case they are all realised by one surface item, but this is subject to the good reason principle. If there is reason to separate them they will be separated. This does not necessarily mean that the process of interpretation is made more difficult by the presence of marked configurations, at least so long as the reason for choosing any particular set of options is apparent without too much effort. 9 Subsidiary to this

9. Compare this with Clark's (1976) more recent work on the generation of semantic structures and the process which he calls "bridging". This is the process of using world knowledge to infer cohesive relationships between sentences: e.g. our knowledge of the world tells us the link in "John got the picnic supplies out of the car. The beer was warm". If we can see the link we take less time to comprehend the sentence than if we cannot.

important point, is the fact that the independence of the four functions means any combination can occur (Halliday, 1967, Part 2, p.217-218 gives a full set of possibilities for the declarative). This is however subject to the qualification that some of the combinations are rather odd in that the situations in which they can arise are rather limited. (They often look odder written down than said because of the comparatively restricted expressive means of the written medium). From the point of view of the contribution to a theory of cohesion of this grammatical theory one can say the following. The functional account of some of the message structure options suggests that it should be possible to manipulate aspects of the surrounding linguistic environment in such a way as to render this or that selection of options easier or harder to understand. This follows from the good reason restrictions intrinsic to the structure of systemic choices because of their organisation on the basis of markedness principles. Although the major purpose of this thesis is the reexamination of the psycholinguistic data summarised earlier in the light of the new evidence to be presented, the value of systemic grammar as a psycholinguistic model will be assessed at a number of points with reference to particular experiments. Few of the experiments really constitute tests of SG - the number of parameters which would have to be controlled is too large to be manageable. Most of the use of SG here will be in suggesting interpretations for the results obtained. On the other hand there will be experiments in which the model appears to make clear predictions : on the whole the results provide support for SG as a psychological model of parameters which people consider in interpreting sentences.

A subsidiary interest running through several of the experiments is the role of definiteness marking in sentence interpretation. Intuitively one thinks of the use of "the" as picking out the topic of a sentence and "a" an element of lesser interest. That these intuitions are reliable has been shown by Grieve (1974). In a previous paper (Grieve and Wales, 1973) the hypothesis that emphasis is a simple function of voice and word order (the agent being emphasised in the active and the affected in the passive) was shown to be insufficient without consideration of definiteness marking. If the SG theory of the importance of initial position and the interaction between theme and old information is correct then one would expect interactions of these with definiteness marking which serves to indicate both importance and old information. This aspect of the work is extended in the fourth chapter to an examination of the relative efficacy of definitely marked noun phrases, indefinitely marked noun phrases, pronouns and names in facilitating comprehension of sentences in which a coreferential interpretation of a noun phrase in a previous sentence and the relevant noun phrase in the target sentence is intended. Although the exact nature of any interaction between syntax and the nature of the referring expression is impossible to predict due to the novelty of SG as a psycholinguistic theory, it seems likely to be quite complex. Other theories (specifically Clark's) would perhaps allow one to predict speeded comprehension with pronouns or definitely marked noun phrases but would require this to be a set increment and not to relate to syntax.

Chapter 2 : The Influence of Context on the Comprehension and Verification of Sentences Describing simple spatial relations between objects.

Introduction

This chapter, like much of what follows, is about two things : cohesion, and in particular the nature of topicalisation devices in the clause; and the psychological parameters involved in the process of understanding sentences. These two are in fact very intimately related : the main reason that I here separate them is not that I believe it is possible to do so in any strict fashion, but merely because, as noted in the introduction, there are many attempts to examine the one without taking account of the other. I shall here deal first with the notion of the topic of the sentence and then turn to details of the main relevant experiments on sentence comprehension (i.e. those on comprehension of locatives).

Why study topicalisation?

It's worth starting the discussion of topicalisation by giving some reasons for wanting to study it. Most accounts start from the simple intuition that one can divide a sentence into that part which consists of a reference to the object one is primarily interested in, and that part which consists of what one wants to say about that object.[1] This is the basis for the classical Prague school distinction between "theme" and "rheme". As a characterisation of the nature of the topic/comment distinction it is plainly inadequate since both the notions of "what one is interested in" and "aboutness" which are involved here,

1 "object" is here used in the way in which Strawson uses "individual" (Strawson, 1957), namely to refer to spatiotemporal particulars of all kinds, including, of course, people.

are far too vague to be of much use. What is needed is some formal criterion or set of criteria for distinguishing the two. We can see the need for formal criteria in linguistics if we look at work on "focus" and "presupposition" (Fillmore and Langendoen, 1971) and at the work on forming text grammars referred to previously (van Dijk, 1973). There seem to be linguistic choices which depend upon the difference between what is the focus of interest and what is not. If the goal of linguistics is an account of all the principled choices in a grammar then these phenomena must be considered. An additional bonus of any formal theory of the nature of topicalisation would be its interest to philosophers. There has been much debate in recent years over the nature of presuppositions and truth value gaps. One of the most celebrated contributions to this continuing debate came from Strawson (1964) who suggested we could go some of the way towards distinguishing cases where reference failure leads to falsity from those where it leads to a truth value gap by considering what the troublesome sentence is about. If the aberrant referring expression remains in the formulation of what the sentence is about then, says Strawson, we have a case which is relatively favourable to the truth value gap theory. If it does not remain then we have a case relatively more favourable to the falsity theory. Although Strawson does not attempt to give a formal criterion of "aboutness" - we are left to rely on our own intuitions - it is clear that a formal characterisation would tidy up his argument considerably. It could well lead to a more transparent formulation of the problem.

Finally the major reasons for studying the phenomenon and looking for formal criteria are psychological. As noted in the Introduction a good many experiments have shown that thematic or topic information is used by subjects in a variety of tasks. Since the importance of this kind of information has been questioned explicitly by Clark (1969) any additional evidence should be of some interest. The literature on psychological aspects of topicalisation will be considered later after we have examined some of the candidates for the delimitation of the notion of the topic of the sentence.

What picks out the topic?

One of the more obvious candidates for the criterion distinguishing the topic is that of previous mention. According to it a noun in the current sentence is the topic if it has been previously mentioned and any other nouns have not. Extensions of this theory are possible to deal with cases where there are more than two nouns in the sentence of interest, two or more of which have been previously mentioned, but this rapidly becomes extremely complicated. Furthermore it is easy to show that the whole argument (as an argument about formal criteria of topic/comment) is broken backed. It is both too strong and too weak. This is so because what matters is not identity of substrings; in many cases the very fact that two substrings are identical means that they cannot refer to the same object. Thus if I say a sentence with "an x" in and soon after a different sentence with "an x" in again, then you can be reasonably sure that the second x is not the same object as the first one. e.g. "John saw an ass. Fred saw an ass, too". What is important is not identity of substrings but rather referential or possibly denotative[2] identity. When we see this we can see straightaway that a new string can refer to an object already referred to. As I write this "the man who came the day before yesterday" refers to the same person as "the man who came on the 19th". These examples use complex strings which are partly the same, but they need not be: this is most clear with the anaphoric use of pronouns. Only insofar as we have developed a formal theory to cope with phenomena of this kind can we say that the previous mention criterion is formal. In fact the problems of reference involved here are so great that this is not yet in sight. Furthermore even if we had an adequate theory of previous mention we would not have all we need of a theory of topicalisation for the simple reason

2 For this distinction see Donnellan (1972).

that while previous mention may be a sufficient condition for picking out the topic (even this is debatable) it is undoubtedly not a necessary condition.[3] However cases like this will not be dealt with in what follows and we will assume that previous mention (in a loose sense) is enough to set up a topic. This assumption is essential if the rest of this discussion is to get off the ground - even though, in the long term, it is to be hoped we can do away with it.

An apparently straightforward method of distinguishing the topic from the comment is the use of definiteness marking or pronominalisation. Our intuitions indicate that the topic is much more likely to be marked definite and any other noun which might form part of the comment indefinite. This relates to our intuitions regarding previous mention in that nouns which refer to objects which have been previously mentioned in a discourse are marked with "the" while those which refer to objects not previously mentioned are marked with "a" (Grieve, 1974). There is empirical evidence (Grieve and Wales, 1973) that when presented with a sentence, subjects tend to see the object marked with "the" (in a transitive sentence with asymmetrical definiteness marking) as the topic. This is measured by presenting subjects with a series of sentences which purport to be answers to questions and asking them to construct plausible questions for which these would be possible answers. The topic is judged to be whatever element of the sentence is mentioned in the question (although if more than one is mentioned the response is not classified). The results are as one might expect:

(1) if one is "the" and one "a", the noun with "the" is mentioned more.

3 It can make the world of difference, though. For example Bransford's "washing clothes" paragraph quoted by Clark (197[illegible]), is about washing clothes, though these are never mentioned. I am a little reluctant to say that the topic of the paragraph is "washing clothes" unless one knows this - even though the paragraph remains the same.

(2) if both are "a" the event is mentioned more.

(3) this holds regardless of sentence voice.

This criterion fails to cope with the case where both nominals are definitely marked, where one would imagine it would predict an even split. But this is not what is observed and, what is more, the unevenness of the split varies significantly between active and passive. So although definiteness marking is of some use, there is clearly more involved than that.

One factor obviously involved in the Grieve and Wales study is syntax, in particular voice. There is a certain amount of evidence that certain syntactic constructions highlight the division of the sentence into topic and comment (Hornby 1971, 1972, 1973, 1974). Hornby (1972) has demonstrated that there is considerable reliability in people's judgements as to what a sentence is about. In his paradigm they have to choose between two pictures as representations of the intended reference behind an utterance. One of these pictures has a different actor from the sentence and one has a different patient. The choice of actor or patient does not seem to constitute a reliable effect (no significant difference was found) but there are very reliable effects with all the seven sentence types which he studied - the effect varying with the grammatical construction. (The seven types were active, stressed active, passive, pseudo-cleft actor, pseudo-cleft patient, cleft actor, cleft patient). Unfortunately Hornby fails to give an overall characterisation of his data, such that we would have a general method for picking the topic. It seems that, on his account, we can go no further than a listing of syntactic constructions together with, for each type, where we can find the topic in the surface structure associated with it. Furthermore this list could not be formulated on a theoretical basis : we would have to discover where the topic is for each type of construction. This fails to capture the generality of the notion in any way. In addition there are aspects of Hornby's data which favour an alternative explanation. These will be explored in due course.

A fourth putative criterion for picking out the topic of a sentence is the notion of the subject of the sentence. As was explained is some depth in the Introduction the view adopted here is based on that of Halliday and so recognises that there are in fact four separate criteria involved here. These are : the logical subject or actor, the grammatical subject (in English the noun in agreement with the verb), the theme (the item in first position in the clause) and the given (that part of the sentence which is not the focus of an information contour). I will take these in turn. Hornby's (1972) data appear to show that the logical subject or actor is unimportant. As we will shortly see there are several other studies which conflict with this result. One advantage of this criterion is that there are no demarcation problems with it : the transitivity system specifies clearly which is the subject in this sense. Specification of grammatical subject is equally straightforward and formal. So too, generally speaking is the specification of theme : it is the first nominal in the surface structure.[4] As indicated in the introduction specification of the given is more complex as this can be discontinuous in the surface structure and may be very large. However the specification is quite clear. As noted, all four notions of topic are expounded (realised) by the same item in surface structure if there is no reason for them to be separated (i.e. in the least marked case). This would perhaps explain why we think of topic/comment as a single distinction instead of a complex of distinctions. As formal distinctions they are quite clear but the question of their psychological validity is a separate issue - one we will turn to in a few moments.

Before doing that I want to briefly consider lexical marking as a possible weak topic indicator. It seems odd from one point of view that marked items should ever be used : if they are harder to process then why

4 There are objections to this but we can ignore them for the present. See chapter on questions and Halliday (1967, unpublished).

not simply change the word order. Instead of saying I "John is shorter than Fred" why not say II "Fred is taller than John". Of course one reason might be that one wants in addition to conveying the relative heights of Fred and John, to convey by the use of I that both are short. In that case I is clearly easier than III. III "Fred is taller than John, but neither of them are tall". But is this the only "good reason" that would dictate the use of the marked form? In view of the account given by Halliday of the notion "subject of a sentence" it seems rather unlikely that it is. If thematic position is important then it is possible that this overrides the preference for an unmarked form. One consequence of this is that the presence of a marked form is a weak indicator that there is a topic though it does not, of itself, indicate where the topic is. The other side of this is that if topicalisation is the kind of good reason that motivates the use of marked forms then one ought not to expect reaction time difference in favour of the unmarked form if there are possible topicalisation reasons for using the marked form. There is a very important assumption built into that statement. It is that the time taken to process a sentence is not independent of the possible function which the sentence is serving in a communicative act. In particular the coding of adjectives from marked/unmarked pairs is dependent upon possible topicalisation choices. If this is correct then simple additive stage models such as the one suggested by Clark (1974) would appear to be inadequate.

The criteria discussed so far although they do not constitute a very strict method of determining the topic of a sentence do go a long way towards meeting the demand for such a set of criteria. It is of some interest to see whether it is possible to find psychological correlates for each of them. The work by Grieve (1974) and Grieve and Wales (1973) has already been mentioned. Their results show quite clearly that definiteness marking has reasonably well-defined effects in terms of one measure of what people take to be the topic of a sentence. Grieve showed similar effects also in the production of sentences. Grieve and Wales were generally rather sceptical of the importance of the syntactic construction. They commented that "we do at times seem to depend on word order for indication of importance, but ..... this is restricted to certain types of construction : namely, full passives where both nominals are definitely marked" (p.181). This conclusion is of course to be qualified in the light of the fact that they were looking at the so called voice word order hypothesis (VWO) and definiteness marking - and not at any broader set of syntactic possibilities. The VWO states that sentence initial position is psychologically important, but that this is especially so in the passive. As such it attempts to express an intuition which is given more precise treatment by Halliday in his notion of theme - which on his account is only one of many foregrounding options available to a speaker. In fact there are several aspects of the Grieve and Wales data which suggest that word order may be more important than they state. Firstly if one considers those cases where marking does not distinguish between the two nouns (both marked "the" or both marked "a") then there is a huge difference between active and passive:-

		Mentioned in the Question			
		Noun 1	Event	Noun 2	a
Voice	Active	5	34	29	$\chi^2_2 = 48.69$
	Passive	35	38	1	$p < 0.001$

Secondly if one considers those sentences where the two are marked differently there is a significantly greater tendency to pick the first noun in the passive than in the active, viz:-

		Mentioned in the Question Noun marked with:-				
		"The"	Event	"A"	Article	Order
	Active	29	5	6	The	a
$\chi^2_5 = 12.13$		34	4	1	a	the
$p < 0.05$	Passive	36	4	0	the	a
		25	5	8	a	the

a. Here, as elsewhere in this thesis degrees of freedom for χ^2 and F are denoted by subscripts.

were generally rather sceptical of the importance of the syntactic construction. They commented that "we do at times seem to depend on word order for indication of importance, but ..... this is restricted to certain types of construction : namely, full passives where both nominals are definitely marked" (p.181). This conclusion is of course to be qualified in the light of the fact that they were looking at the so called voice word order hypothesis (VWO) and definiteness marking - and not at any broader set of syntactic possibilities. The VWO states that sentence initial position is psychologically important, but that this is especially so in the passive. As such it attempts to express an intuition which is given more precise treatment by Halliday in his notion of theme - which on his account is only one of many foregrounding options available to a speaker. In fact there are several aspects of the Grieve and Wales data which suggest that word order may be more important than they state. Firstly if one considers those cases where marking does not distinguish between the two nouns (both marked "the" or both marked "a") then there is a huge difference between active and passive:-

		Mentioned in the Question			
		Noun 1	Event	Noun 2	a
Voice	Active	5	34	29	$\chi^2_2 = 48.69$
	Passive	35	38	1	$p < 0.001$

Secondly if one considers those sentences where the two are marked differently there is a significantly greater tendency to pick the first noun in the passive than in the active, viz:-

		Mentioned in the Question Noun marked with:-				
		"The"	Event	"A"	Article Order	
$\chi^2_5 = 12.13$	Active	29	5	6	The	a
		34	4	1	a	the
$p < 0.05$	Passive	36	4	0	the	a
		25	5	8	a	the

a. Here, as elsewhere in this thesis degrees of freedom for χ^2 and F are denoted by subscripts.

Taking only the cases where a noun was picked gives us the following result:

	Mentioned in the Question		
	Noun 1	Noun 2	
Active	30	40	$\chi^2_1 = 6.18$
Passive	44	25	$p < 0.02$

This is a clear difference between the two voices. Finally if we conside[r] all the responses which fall clearly into one of the three classes event, noun 1, noun 2, there is a very marked difference between active and passive. Subjects tend to pick the second noun as the "topic" (by the Grieve and Wales criterion) more than the first noun in the active ($\chi^2_1 = 11.11$, $p < 0.001$). The reverse is true in the passive ($\chi^2_1 = 26.75$, $p < 0.001$); while the event is chosen about equally often in both voices.

	Mentioned in the Question		
	Noun 1	Event	Noun 2
Active	35	43	69
Passive	79	47	26

There can be no doubt that there is some sort of voice effect here in addition to the effects of definiteness marking. However the tendency to select the second noun in the active is contrary to the usual VWO hypothesis, though the results for the passive voice support it.

This result appears to contradict such earlier results as those of Johnson-Laird on both the interpretation of voice and the choice of voice in a communication task (Johnson-Laird 1968a, b). He showed evidence which appeared to support accounts which emphasise the importance of sentence initial position as a method of foregrounding. There does not seem to be any explanation of this discrepancy between the two methods of investigating voice. The simplest explanation for the Grieve and Wales results is that there is a bias towards producing questions which mention the patient. This parallels the Smith and McMahon (1970) and Wright (1969) results which show it is easier to answer a question with the actor as answer.

It is possible that people naturally focus on the actor : that it is the natural topic and the centre around which processing hinges.

A series of papers by Janellen Huttenlocher and her associates appear to support this view (Huttenlocher and Strauss, 1968; Huttenlocher, Eisenberg and Strauss, 1968; Huttenlocher and Wiener, 1971). These experiments showed that children found it easier to place objects in a display if the object to be placed is subject of the sentence. For instance given a sentence like "Make it so the green block is under the red block" it is easier to place the green block in a "ladder" with the red one already in place than it is to place the red one with the green one already in place. Now this is explicable in terms of theme (in Halliday's sense) : a natural focus of interest is on the block which has to be manipulated so that it is easier if this is foregrounded linguistically i.e. thematic. Huttenlocher interprets the results as due to a need to have some kind of natural congruence between sentence and situation. Huttenlocher believes that people performing these placement tasks think of the moveable block as like on actor and naturally associate this with the grammatical subject. She claims it is not a simple word order effect. Rather than use sentences in which the locative phrase precedes the copula to rule this out she decided to use transitives and vary voice. She found superior performance when the actor was the object to be placed. This seems to support her idea of natural relationships between sentence and situation. On the other hand she was forced to admit that some residual importance attached to the grammatical subject. Then again, she at no time separated the grammatical subject and the noun in first position so that it is not clear which of these is involved.

That the case role of the object to be placed should matter more than the surface structure role is odd given Johnson-Laird's results already referred to. However it is consistent with the Grieve and Wales data, as already noted. It is also consistent with the Smith and McMahon (1970) and Wright (1969) results.

An experiment by Hornby (1972), already referred to, runs counter to this interpretation. He presented subjects with pictures, neither of which were truly described by a sentence presented simultaneously and asked subjects to guess which picture was intended. An example might be : a picture of an igloo with an indian building it and a picture of a tepee with an eskimo building it together with the sentence "The igloo is being built by the eskimo". Hornby conjectured that different syntactic constructions would lead to different decisions as to which was the right picture. He found that the decision was affected in a very reliable fashion by the kind of syntactic construction used, but concluded that word order as such was unimportant. Now of Hornby's 7 constructions, four tend to lead to the picture in which the first noun's referent is represented being chosen significantly more often, and three to the picture in which the second noun's referent is depicted being chosen. So if there is no effect of word order per se there ought to be no change in the number choosing according to the grammar (as Hornby sees it) depending on whether the grammar predicts the first or the second noun will be thought of as topic. In fact this prediction is falsified : if the grammar (on Hornby's interpretation) predicts the first noun is topic then more people pick the picture depicting the first noun than pick the second noun if it predicts the second noun. Thus:-

	As per prediction		not as per prediction	
Grammar predicts noun 1	746	(67%)	374	(33%)

	As per prediction	not as per prediction
Grammar predicts noun 2	496 (59%)	344 (41%)

$$\chi^2_1 = 11.82, \quad p < 0.001$$

This result quite clearly indicates a residual theme effect. Of more immediate importance is the failure to find any effect of case role as such. If we consider only the six syntactic variations (and ignore the data for the stressed agent simple active) then the agent picture was selected 818 times and the object picture 862 times : a non significant ($\chi^2_1 = 1.15$, $p > 0.1$) tendency to select the non-actor. Thus Hornby's results provide no support for the actor-as-focus-of-processing hypothesis though they do provide support for the word order hypothesis. Hornby's own conclusion - that we cannot generalise beyond individual syntactic constructions - seems unnecessarily restricted.

Rationale and Predictions for Experiment 1

The experiment to be reported below sets out to look again at the process of understanding a sentence to see what happens to several of the parameters mentioned above when the major topicalisation factor, namely previous mention, is manipulated. The experiment is essentially a verification one but separate measures are taken of the time taken to understand the sentence and the time taken to verify it. This is achieved by presenting the sentence prior to the picture and allowing the subject to control the time for which the sentence is visible before onsetting the picture. Though this method must surely have its own peculiarities from the point of view of the kind of linguistic processing which occurs, it seems fair to say that it should allow us to make more well-founded inferences about any separate comprehension and verification parameters than the all-in-one method of Clark and Chase (1972). In that paradigm a single measure is taken of both comprehension and verification processes since "sentence" and "picture" were presented simultaneously. The use of full sentences rather than the skeleton sentences occasionally used in previous work (e.g. Clark and Chase's[5] "star above plus"), together with sketches of real objects (animals, people, vehicles etc.) rather than abstract symbols (such as typewriter symbols), was to encourage natural linguistic processing - insofar as there is such a mode and one can talk of it in an experimental situation of this sort. All sentences used were simple relational sentences with locative predicates and the following parameters were manipulated in the experiment:-

1. Whether only one sentence was presented or whether the target sentence was the last of a series of sentences all describing the picture.
2. Whether the two nouns mentioned in the sentence were accompanied by the definite or indefinite articles : each noun marking was manipulated independently.
3. Whether the relational term was marked or unmarked. Some workers will undoubtedly object to my terminology here. I shall continue to use it for several reasons:-
   (a) because there is no better terminology. Clark's later "positive" and "negative" seem too evaluative and only perspicuous in the case of prepositions if one accepts the linguistic-perceptual homomorphisms Clark suggests (see Introduction).
   (b) because Clark has demonstrated a clear RT difference between the American English lexical items corresponding to the British English items used in the present experiment. (see Clark, 1974, and Introduction). In this respect the prepositions to be used here behave like a marked/unmarked pair.
   (c) details of the results suggest the neutrality of one of the pair (viz : "in front of") "vis-a-vis" the other (viz : "behind"). This is a classical feature of a marked/unmarked pair.
   (d) the exact interpretation of marking effects is still unresolved. Clark (1974) gives two quite distinct possibilites. As noted in

5. Clark's work is again picked out for critical comment because it is the best example of a certain kind of approach to the problem.

the Introduction there are more possibilities than that.

4. Whether the lexical item referring to the object mentioned in the earlier sentences was the first or second noun in the sentence. (only one of the two objects was mentioned in the sentences prior to the target). If any form of word order hypothesis is true one would expect this factor to have some effect, though there is a good deal of room for disagreement over what kind of effect one might expect.

5. a final factor sought to investigate reaction times to a separation of theme and grammatical subject by the use of a marked syntactic option (e.g. "Behind the bus is a sports car"). Investigations of voice separate grammatical from "logical" subject but to my knowledge no previous work has looked at this possibility.

The experiment investigates the following hypotheses:-

1. that failure to use the correct article ("the" or "a") - correctness here being dictated by previous mention - will lead to longer reaction times. In fact for a variety of reasons the analysis by noun position rather than previous mention will be the one most discussed in what follows. However both analyses were performed. (By "analysis in terms of noun position" I mean defining the two definiteness variables in terms of the place of the article in the sentence - i.e. with the first or second noun - rather than in terms of whether the noun it is with was previously mentioned or not. I call these the Def_1/Def_2 and the Def / Def_o analyses respectively - the subscripts denoting the noun viz : noun 1, noun 2, "topic" noun and "other" noun respectively).

2. that the reaction time difference to the marked and unmarked lexical items will only occur in the one-sentence condition. Where there is the possibility of the marked item being chosen for topicalisation reasons (i.e. in the case of the sentence being embedded in a series of sentences) this effect should not occur.

3. A parallel prediction to (2) for the syntactic option. If we can see the choice of locative phrase in sentence initial position as the

selection of a marked option then this too should be subject to a good reason principle. It is hard (or impossible) to see what the reason might be in the one sentence case. There ought therefore to be a main effect of this factor in the one sentence ("no text") data. Some of the combinations of factors in the several sentence ("text") condition ought to make sense or constitute a reasonable selection. There ought therefore to be cases in the text condition where the marked syntax is understood at least as fast as the unmarked. This is not to say necessarily that we will not find a main effect of syntax, merely that there will be sentences with marked syntax which take no longer to understand than corresponding sentences with unmarked syntax.

4. that the position of the noun referring to the object already referred to in the preamble will have an effect on reaction time. This is the weakest possible formulation of the word order hypothesis - or so one might think. However it is not clear just why the previously mentioned object (the "topic") should be the main focus of attention in the target sentence. One might equally well expect (in the spirit of Huttenlocher) that people in listening to the target sentence are focussing upon the new object since the task calls for them to work out its relation to the topic ready for the verification. In Huttenlocher's terms they have to add the new object to the mental display which already contains the topic. In that case the new noun might be supposed to be the focus of attention. If the interpretation of the marked syntax as a thematic foregrounding device is correct then one might well expect interactions between it and the topic position. But again it is not clear quite how one would expect this interaction to work. If the locative foregrounding is analogous to the use of passives in transitive sentences (and this is by no means clear) then one should observe the following interaction. The marked syntax should be as easy as the unmarked if the previously mentioned object is first, but much harder if it is second. There should

be little effect of the topic position with unmarked syntax. This follows from Halliday's analysis of the passive as a means of maintaining the patient as theme (1967, Part 2). On his analysis passives in context would be as easy as actives if the patient is topic of the discourse preceding the target, but otherwise not. Actives are supposed to be either neutral or fairly weak in their distribution of emphasis, and so less affected by the position of the topic. Although the marked/unmarked options in relational sentences bear some resemblance to passive/active options, they are clearly different in a number of respects. For example one can switch theme by keeping syntax constant and changing the relational term. This kind of difference makes any analogical argument necessarily very weak. Nevertheless it does not seem unreasonable to expect some kind of interaction of topic position and syntactic type with possibly the relational term also playing a part.

One should not of course expect any effects of the topic position in the one-sentence case.

So far the discussion has centred on the effects one might expect in the comprehension data. However, as noted in the Introduction there is also some interest in the actual process of verification. Gough (1966) showed the voice effects occurred even after a delay. One has already discussed some of the literature on the diminution of "surfacey" information over time. The delay between "comprehension" and verification in the present experiment is quite small (about 5 seconds) so that one would expect similar effects for both the text and the no text conditions on verification times. On the other hand it seems possible that topic information will have some effect. On a Hallidayan account topic information is in the deep structure - context and surface structure provide pointers to it. SG therefore predicts different results for the text and no text verification times, due to the different topic structures of the target sentences (even though the surface structure may be the same).

The kind of account given by Clark and Chase (1972) stresses the role of canonical structures in verification. Although they might expect some (presumably very small) effect of surface structure on the verification process, they would certainly not expect this to differ in the two conditions.

A relatively straightforward prediction can be made from this approach and equally straightforward predictions from the "surfacey" and "topic" viewpoints:-

1. the canonical view predicts that if one considers the means for the 32 sentence types (defined by the variables specified above) then the verification times ought to be more highly correlated with one another than either is with the comprehension times. This allows for the possibility of topic information affecting comprehension times, but assumes more or less complete reduction to a canonical form for verification.

2. the surfacey approach predicts a high correlation between both sets of verification times and their corresponding comprehension times as well as a high correlation with one another. Indeed a truly surface approach would predict high correlation between all four sets of results.

3. A Hallidayan approach predicts positive correlations between both sets of comprehension times and their corresponding verification times but little or no correlation between the text and no text results.

The verification data are important in that they may distinguish between these three hypotheses, and hence lean us towards one or another view of the comprehension process. Primarily, though, one is interested in comprehension and it will be that data which is considered in most depth.

Method

1. Subjects

43 undergraduates fulfilling a course requirement for an introductory psychology course at Stirling University. Modal age approximately 18 years. 28 were female, 15 male.

2. Apparatus and Materials

Subjects sat in a quiet chamber designed to mask equipment noise. They looked through a plain glass window into a tachistoscope equipped with an electronic card changer. Two fields of the tachistoscope were used : one to maintain a background level of illumination and one to display the cards with stimuli on them. Each card had either a typed stimulus sentence or a photocopy, in black and white, of a Letraset picture stuck onto it with Sellotape. Sentences to be verified also had a red mark at the beginning of the sentence, ½" from the start of the sentence. Subjects sat with a small box in their hands with three buttons on it : one in black was the button to indicate when they had understood each sentence, the other two in red were to signal the truth or falsity of the target sentence. The buttons were arranged in the shape of an equilateral triangle with the change button at the apex and the two "truth" buttons at the base. The "true" button was always on the right.

Pressing any button extinguished the field with the stimulus in it and advanced the card changer by one card. It also stopped the clock. The type of response and reaction time was then punched by means of a data transfer unit onto paper tape. A centisecond timer was used. Once the card changer had removed the old card from the stimulus field it returned to its resting place and after a fixed delay of five seconds the stimulus was illuminated and the clock started.

The stimulus cards comprised 16 stimulus sets. Each stimulus set was composed of 4 topic setting or preamble cards, 32 target sentences made up as detailed below, and 2 pictures. The preamble sentences always described the object depicted in the centre of the picture. They were short and simple and designed to consist of easily imageable information. The pictures, as already noted, were made up from black and white photocopies of letraset rub-on pictures. The visible part of each card could be divided into three panels. The object described in the preamble was always in the middle panel with the object mentioned only in the target sentence in one of the other two panels. False sentences always had the items in the wrong order. Only the two objects mentioned in the target sentence were depicted. All objects were pictured in profile and facing towards the subject's left. All stimulus sets were used twice for each subject.

3. Design and Procedure

The design was a simple 2^6 multifactorial one with five factors within subjects and one factor between subjects. The between subjects factor was the presence or absence of the preamble of four sentences, hereafter called the Text/No Text factor. The within subjects factors were as follows:-

1. whether the second noun is marked with the definite or indefinite article ("the" or "a").

2. whether the first noun is marked with the definite or indefinite article.

The decision to label these factors on the basis of order in the sentence is based on two considerations:

(a) several authors maintain that word order is of considerable importance, in particular Halliday in his proposal of a function of theme.

(b) since the topic designation is effectively arbitrary for the no text condition we might expect order to be more important here than the topic. However analyses were also done in terms of the definiteness marking of the noun referring to the object mentioned in the preamble, and of the noun only mentioned in the target sentence. These are not presented in full below as they generally reveal little not shown by the analysis in terms of definiteness marking of the first and second nominals. They will be referred to occasionally though.

3. this factor concerns the position of the noun referring to the object mentioned in the preamble, in the target sentence (whether it is the first or second nominal). For the no text condition there is clearly no difference between the two nouns in each stimulus set since neither is mentioned previously. This factor is therefore assigned by correspondence with the condition in which there is a preamble. It should be noted that this factor is completely arbitrary if one considers only the sentences of the no text conditions. It is non-arbitrary in relation to the pictures of the no text condition because the nominal designated topic always refers to the object in the centre of the picture. It is therefore possible that there might be effects of this factor on the verification times of the no text condition but there ought not to be any effects on the comprehension times.

4. this factor concerns the relational term : whether it was "behind" or "in front of". Clark (1974) has shown that "in front of" is significantly faster understood than "in back of". However "in back of" is not a lexical item in British English and "in front of" contrasts with the simpler "behind". These terms are here used in the "bus queue" rather than the "depth" sense. That is to say that the truth of sentences involving them is not dependent upon the position of either speaker or hearer, and "behind" is not equivalent to "beyond". Subjects were asked at the start of the experiment what they meant by "behind" and "in front of". About 50% gave one type of meaning, about 50% the other. All subjects who did not see both meanings had them explained.

5. this factor is concerned with the use of normal or marked syntax in the target sentence. Despite Huttenlocher's claims to the contrary (Huttenlocher, Eisenberg and Strauss, 1968) there is no necessity in English locatives to have the subject earlier in the sentence than the locative phrase. There is a perfectly natural locative foregrounding option which makes the locative phrase thematic. Thus "John is behind Fred" can be expressed as "Behind Fred is John".

Subjects were told that they would sit in the quiet booth and look through the window into the tachistoscope. There a series of sentences would be displayed one at a time. When they had read and understood each sentence they were to press the black button. This would immediately wipe out the sentence which would be followed after a short delay by another sentence. The sentences would all describe one object which they were to try to imagine. The fifth sentence would describe the relation between that object and another object. It would have a red mark to its left. After the fifth sentence a picture would be displayed. They were to press either the "true" or the "false" button to indicate whether the fifth sentence was true or false of the picture. The other sentences would all be true but they were not to ignore them : trying to imagine the object would probably help them with the verification task. Instructions were suitably modified for the no text group.

Subjects were not given any practice trials. This was done in order to minimise the probability of any local strategy effects being picked up in the data. All subjects responded to only one sentence of each type. All subjects had the stimulus sets in the same order (though initial order was random) but order of presentation of sentence types was randomised separately for each subject with the exception that subjects in the two conditions received the same random order. This meant that subject one in the text condition had the same order as subject one in the no text condition and so on. It was not possible to carry out this procedure completely as some subjects' results were dropped due to their high error rate.

There was an interval of at least 10 seconds between trials as a blank card was inserted between each batch of trial cards (2 changes of the card = 10 seconds).

The experiment lasted 15 - 20 minutes for the no text group and 40 - 50 minutes for the text group.

Results

As noted in the introduction to this experiment two sets of results were analysed:

1. the time taken to indicate that the target (5th) sentence had been understood, hereafter called the comprehension time. It's worth noting that subjects may not have fully understood the sentence in this time but rather have relied on the 5 second gap to complete processing. The results make this suggestion a little implausible, as we shall see.

2. the time taken to respond "true" or "false" after the illumination of the field containing the picture. The fact that subjects only responded once to each sentence type means that the experiment contains a lot of variance (this is made especially severe because of the absence of practice trials). This is made even more acute for the verification times because of the fact that "true" and "false" responses take different lengths of time and subjects times for each sentence type are not averaged over "true" and "false" responses (since for any sentence type they only responded one or the other). This problem does not, of course, affect the comprehension times. It was felt that 64 trials (the number required to obtain one response per sentence type per truth value per subject) was too great a number and would encourage the development of special strategies. Since we are interested in natural linguistic processing rather than the process of verification "per se" in tasks of this nature, it is desirable to avoid these as far as possible.

The basic analyses for both sets of data are six factor analyses of variance of the reaction. As noted in discussing the design, analyses were also carried out with the definiteness variables redefined. These will not be presented in any detail, though they will be referred to in places.

A central problem with experiments with only one response per subject per cell is what to do about errors. There is no solution which is really satisfactory. In the present experiment 7 subjects' data were discarded because the error rate exceeded 12½%. This left 18 subjects per condition. 10 females and 8 males in the text condition and 12 females and 6 males in the no text condition. All their responses were treated as homogeneous with no correction for errors, since errors appeared to be more or less randomly distributed. Of these 18 subjects per condition there were 38 errors in the no text condition and 30 in the text condition. The distribution over sentence types is given in Tables 1 and 2.

Comprehension

Tables 1 + 2 give the mean reaction times to each of the 32 sentence types for both conditions. There is a tendency for the text times to be shorter than the no text times : in fact 26 of the 32 sentences are reacted to faster in the text condition. This figure is highly significant on a sign test ($p < 0.005$). There is no correlation between the times for each sentence type in the two conditions ($r = -0.05$, df 30, $p > 0.1$). This result strongly suggests that the processes taking place in the two conditions are quite different.

The six way analysis of variance for the comprehension data is presented in Table 3 with the main effects summarised in Table 6. The difference between the two conditions here fails to reach significance ($F_{1,34} = 3.12$, n.s.) despite a mean difference of 602 msec. (Text 2600 msec., No Text 3202 msec.). As we will see there are several interactions involving this factor, which explains why the main effect fails to reach significance here when it was highly significant on the sign test.

The definiteness factors also fail to reach significance, however one defines them. Defining in terms of position of the noun in the sentence the first noun factor an F value of less than one ($F_{1,34} = 0.24$, n.s.) though reaction times are slightly slower if the first noun is indefinitely marked (2872 msec. vs. 2931 msec.). The F value for the second noun marking factor is also non-significant ($F_{1,34} = 1.64$, n.s.) though reaction times are slightly shorter if this noun is indefinitely marked (2960 msec. vs. 2842 msec.).

Table 1 Mean Reaction Times and Error Totals : Text Data.[a]

Sentence Type	Errors	Reaction Times	
		Comprehension	Verification
In front of the topic is the other	1	2977	2964
In front of the topic is an other	0	2168	2120
In front of a topic is the other	0	2663	2677
In front of a topic is an other	1	2807	2192
In front of the other is the topic	1	3140	1911
In front of the other is a topic	2	2509	3004
In front of an other is the topic	2	3304	2450
In front of an other is a topic	1	3192	2563
Behind the topic is the other	0	2279	2441
Behind the topic is an other	0	1973	2093
Behind a topic is the other	0	2356	2281
Behind a topic is an other	0	2397	2442
Behind the other is the topic	0	2799	2309
Behind the other is a topic	0	3226	2493
Behind an other is the topic	2	2807	2297
Behind an other is a topic	0	2660	2402
The topic is in front of the other	1	2709	2558
The topic is in front of an other	2	2405	2385
A topic is in front of the other	3	2790	2708
A topic is in front of an other	2	2371	1987
The other is in front of the topic	1	2460	2449
The other is in front of a topic	1	2783	3530
An other is in front of the topic	0	2100	2152
An other is in front of a topic	1	2172	1845
The topic is behind the other	2	2440	2778
The topic is behind an other	0	2030	2288
A topic is behind the other	0	2614	2447

a. Reaction Times are milliseconds. Each mean is based on an N of 18. Errors are totals, not means.

Table 1 Mean Reaction Times and Error Totals : Text Data. (contd.)

Sentence Type	Errors	Reaction Times	
		Comprehension	Verification
A topic is behind an other	1	2780	2467
The other is behind the topic	1	2410	1940
The other is behind a topic	1	2662	2140
An other is behind the topic	3	2315	2497
An other is behind a topic	1	2412	2625

Table 2 Mean Reaction Times and Error Totals : No Text Data.[a]

Sentence Type	Errors	Reaction Times	
		Comprehension	Verification
In front of the topic is the other	2	3251	1601
In front of the topic is an other	1	2971	1667
In front of a topic is the other	0	3082	1680
In front of a topic is an other	1	2795	1556
In front of the other is the topic	1	2829	1502
In front of the other is a topic	0	2915	1760
In front of an other is the topic	1	3248	1910
In front of an other is a topic	2	2702	1832
Behind the topic is the other	3	3374	2119
Behind the topic is an other	1	3518	1752
Behind a topic is the other	0	3764	1760
Behind a topic is an other	0	3985	2070
Behind the other is the topic	1	3965	1934
Behind the other is a topic	3	2999	2179
Behind an other is the topic	1	3478	2044
Behind an other is a topic	0	3420	1780
The topic is in front of the other	2	2754	1590
The topic is in front of an other	0	3062	1796
A topic is in front of the other	2	3266	1469
A topic is in front of an other	1	3162	1495
The other is in front of the topic	1	3040	1908
The other is in front of a topic	1	3049	1514
An other is in front of the topic	1	2703	1563
An other is in front of a topic	2	2741	1648
The topic is behind the other	2	3817	1951

a. Reaction Times are Milliseconds. Each mean is based on an N of 18.

Errors are totals, not means.

Table 2 Mean Reaction Times and Error Totals : No Text Data.

Sentence Type	Errors	Reaction Times	
		Comprehension	Verification
The topic is behind an other	1	2945	1515
A topic is behind the other	2	2537	1512
A topic is behind an other	1	3603	1739
The other is behind the topic	1	2999	1720
The other is behind a topic	0	2943	1638
An other is behind the topic	2	3923	1778
An other is behind a topic	2	3094	1333

Defining the definiteness variables in terms of topic and new noun is scarcely any different (see Table 4). The topic noun F ratio is less than one ($F_{1,34} = 0.11$, n.s.) with virtually identical times for the two levels (2897 msec. when "the", 2906 msec. when "a"). The other (new) noun factor shows a similar lack of effect ($F_{1,34} = 1.1$, n.s.; when "the" 2936 msec., when "a" 2867 msec.).

There is no main effect of the topic noun position. However there is a highly significant interaction between the topic and text factors ($F_{1,34} = 12.91$, $p < 0.01$). This is due to a superiority with the topic noun first in the text condition (2516 msec. vs. 2684 msec.) but with it second in the no text condition. It is not clear why there should be such a difference between the two levels of this factor in the no text condition as it is a pseudo-factor as far as the comprehension data goes.[6] Interpretation of the interaction is made more complex by the presence of other interactions involving the topic position factor. There is an interaction between it and the syntax factor such that overall reaction times are faster to marked syntax if the topic is the first noun (2960 msec. vs. 3074 msec.), but to unmarked syntax if the topic is the second noun (2833 msec. vs. 2737 msec.). However, this result appears to be due entirely to an interaction of text, topic position and syntax. This three-way interaction is significant at the $p < 0.05$ level ($F_{1,34} = 6.92$). The no text data show what appears to be only a tendency to longer RTs with the topic first (for marked syntax the figures are 3405 msec. and 3194 msec., for unmarked 3149 msec. and 3061 msec. for topic first and topic second respectively). The text data on the other hand show a very large difference between topic first and topic second for the marked syntax (2515 msec. and 2955 msec. respectively) but a small difference in the opposite direction with unmarked syntax (2517 msec. vs.

6 It is not entirely a pseudo factor for the verification part of the task. As already noted the topic noun always refers to the object in the centre of the picture.

Table 3 : Comprehension Data : Analysis of Variance, Part One

Factor	$F_{1,34}$ Value	Factor	$F_{1,34}$ value
A	3.12	AEF	* 6.09
B	1.64	BCD	2.95
C	0.24	BCE	* 5.42
D	0.03	BCF	0.36
E	* 4.44	BDE	0.90
F	**10.60	BDF	0.76
AB	0.16	BEF	0.49
AC	0.22	CDE	0.08
AD	***12.91	CDF	0.65
AE	***14.04	CEF	0.97
AF	0.08	DEF	0.41
BC	1.39	ABCD	0.07
BD	0.01	ABCE	** 8.06
BE	0.50	ABCF	1.17
BF	1.41	ABDE	0.26
CD	3.35	ABDF	0.49
CE	0.25	ABEF	0.24
CF	1.23	ACDE	1.14
DE	0.03	ACDF	0.29
DF	* 4.47	ACEF	2.04
EF	0.85	ADEF	0.17
ABC	0.01	BCDE	2.19
ABD	2.27	BCDF	0.02
ABE	1.45	BCEF	0.75
ABF	0.42	BDEF	0.00
ACD	3.03	CDEF	* 6.58
ACE	0.03	ABCDE	0.00
ACF	0.83	ABCDF	0.92
ADE	0.62	ABCEF	0.79
ADF	* 6.92	ABDEF	0.89

Factor	$F_{1,34}$ value
ACDEF	0.57
BCDEF	3.60
ABCDEF	* 5.70

* $p < 0.05$

** $p < 0.01$

*** $p < 0.001$

A : Text/No Text

B : Second Nominal Marking

C : First Nominal Marking

D : Topic Position

E : Relational Term

F : Syntax

2414 msec.). Another way of expressing this is to say that there is no difference between marked and unmarked syntax so long as the sentence is in a text and the topic is the first noun. If the topic is the second noun there is a huge difference in favour of the unmarked syntactic form (actually over 500 msec. in the present data).

This interaction is partly responsible for the significant main effect of syntactic form. This is highly significant ($F_{1,34} = 10.60$, $p < 0.01$) and indicates faster RTs for the unmarked syntactic form by an average of over 200 msec. (2785 msec. vs. 3017 msec.). The interaction of the syntax and text factor fails to reach significance ($F_{1,34} = 0.08$, n.s.). As we have just seen marked syntax is not necessarily always more difficult in the text condition, though this is not true for the no text condition.

The lexical marking factor shows an overall significant effect in favour of the unmarked term ("in front of") though this is quite small (2816 msec. vs. 2986 msec., $F_{1,34} = 4.44$, $p < 0.05$). In fact there is a highly significant interaction between this factor and the text factor ($F_{1,34} = 14.04$, $p < 0.001$). "Behind" is reacted to faster in the text condition (2541 msec. vs. 2659 msec.) but slower in the no text condition (3432 msec. vs. 2973 msec.). The overall significant main effect cannot therefore be taken at face value. This is especially true in view of a significant three way interaction between the text, relational term and syntax factors ($F_{1,34} = 6.09$, $p < 0.05$). There appears to be no effect of syntax in the no text condition with "in front of", but an effect with "behind"; while with text there is only a small effect of syntax with "behind" (167 msec. slower when marked) but a large effect with "in front of" (375 msec. slower when marked).

There are a number of other interactions which involve the relational term. It interacts with both of the definiteness variables ($F_{1,34} = 5.42$, $p < 0.05$) and also with these and the text factor ($F_{1,34} = 8.06$, $p < 0.01$). The three way interaction is probably simplest expressed as follows :

"behind" leads to faster RTs if the two nouns are marked differently (2784 msec.) than if they are marked the same (3232 msec.). There is no such effect with "in front of" : all that is apparent is a main effect of the definiteness of the second noun (2894 msec. when it is definitely marked, 2737 msec. when indefinitely marked). This interaction needs to be seen in the light of the four way interaction with the text factor. The "behind" results are not dissimilar for text and no text data while the "in front of" results are. In both conditions "behind" is noticeably harder when both nouns are marked with "the", and easiest when the first is "the" and the second "a". This is true also for "in front of" in the text condition. In fact the simplest way to characterise these overall is to say that "Behind" (with or without text) and "in front of" (with text) are easier when the two nouns are marked differently. The figures for same and different marking are 2498 msec. and 2869 msec. ("behind" : text) 3269 msec. and 3594 msec. ("behind" : no text) and 2590 msec. and 2733 msec. ("in front of" : text). The "in front of" : no text figures tend in the opposite direction viz. 3037 msec. and 2907 msec. respectively. This perhaps explains why the "in front of" : text figures show a smaller effect than the "behind" ones (only 143 msec. as opposed to well over 300 msec. for both sets of "behind" data).

Redefining these variables in terms of definiteness marking of the topic and other noun is scarcely more revealing (The ANOVA is presented in Table 4 with a summary in Table 7). The no text results do not really change in any noticeable way. The text results show an effect which can be summarised as follows. There is no effect of the definiteness of the other noun if the topic noun is indefinite (figures for "behind" are 2628 msec. and 2687 msec. and for "in front of" 2686 msec. and 2635 msec. for "the other" and "an other" respectively). If the topic noun is definite there is a substantial effect in favour of "an other" (figures in the same order as above are 3052 msec., 2281 msec., and 2821 msec., 2494 msec.). (Any interpretation in terms of same marking/different marking is of course

Table 4 : Comprehension Data : Analysis of Variance. Part Two

Factor	$F_{1,34}$ value	Factor	$F_{1,34}$ value
A	3.12	AEF	* 6.09
B	1.10	BCD	2.95
C	0.11	BCE	* 5.43
D	0.03	BCF	0.34
E	* 4.43	BDE	0.44
F	**10.61	BDF	1.14
AB	0.58	BEF	2.87
AC	3.17	CDE	0.17
AD	***12.92	CDF	1.90
AE	***14.03	CEF	1.60
AF	0.08	DEF	1.52
BC	1.39	ABCD	0.07
BD	0.09	ABCE	**3.05
BE	1.35	ABCF	1.17
BF	0.31	ABDE	1.06
CD	* 5.38	ABDF	0.01
CE	0.11	ABEF	0.16
CF	1.16	ACDE	0.00
DE	0.03	ACDF	0.03
DF	* 4.48	ACEF	0.15
EF	0.85	ADEF	0.17
ABC	0.01	BCDE	2.19
ABD	0.00	BCDF	0.02
ABE	0.16	BCEF	0.75
ABF	1.58	BDEF	3.24
ACD	0.01	CDEF	0.20
ACE	1.51	ABCDE	0.00
ACF	0.05	ABCDF	0.92
ADE	0.62	ABCEF	0.79
ADF	* 6.92	ABDEF	0.03

Factor	$F_{1,34}$ value
ACDEF	3.18
BCDEF	3.60
ABCDEF	* 5.70

* $p < 0.05$

** $p < 0.01$

*** $p < 0.001$

A : Text/No Text

B : "Other" Nominal Marking

C : Topic Nominal Marking

D : Topic Position

E : Relational Term

F : System

left unchanged by this redefinition of both definiteness variables). What is especially interesting about these results is the fact that the two definiteness factors lead to effects on RT which are very clearly non-additive. If topic noun marking is "incorrect" i.e. indefinite then mean RT is 2659 msec. But if it is "correct" then (1) if the other noun marking is correct RTs are smaller (mean 2387 msec.) but (2) if the other noun marking is incorrect RTs are much longer (mean 2936 msec.). Hence one cannot talk of the time taken to process the articles without phrasing it in terms of quite complex conditionals.

In addition to these effects involving definiteness there is a significant 4 - way interaction between topic position, definiteness of the first noun, relational term and syntax ($F_{1,34} = 5.58$, $p < 0.05$). This however is both complex and rather uninformative, especially in view of the higher order interaction involving these factors. Any effect of topic position and definiteness marking will be uninterpretable unless the text factor is also involved.

These comments apply to a two way interaction found on the topic definiteness/other definiteness analysis. This is between the topic definiteness and topic position factors. It appears to show superior performance when the topic is first noun and definitely marked (2823 msec.) or second noun and indefinitely marked. The other two possibilities yield identical RTs (2970 msec.). Again though this interaction does not involve the text factor and is obviously of little interest in view of that. This is especially so in view of the fact that the importance of text when definiteness is involved is demonstrated amply above. Finally one must mention the overall significant six way interaction ($F_{1,34} = 5.70$, $p < 0.05$). This inevitably reduces the degree of certainty with which one can accept lower order interactions. However of itself it adds little information because of the impossibility of grasping it as a whole. It remains as a warning of the complexity involved in even such simple sentences as those used here. On the other hand one should perhaps not overrate this effect :

some of the sentences used in the present experiment were rather odd and may have led to peculiar processing because of this.

Verification.[7]

There are only four effects present in the verification data. Firstly the text times are longer than the no text times by 705 msec. (2435 msec. vs. 1730 msec.); this result is significant at the $p < 0.05$ level ($F_{1,34} = 5.42$). If one looks at sentence types all 32 types take longer in the text condition ($p < 0.001$: Sign test) though the correlation between times for the 32 types in the two conditions is nil (actually $r = 0.06$).

There is an interaction between the text, topic position and second noun definiteness factors which curiously was not significant in the comprehension data. (There $F_{1,34} = 2.27$, $p > 0.1$, here $F_{1,34} = 5.30$, $p < 0.05$). This result simply shows faster RTs when the topic is first and the second noun indefinite, and the topic second and the second noun definite. This is only present in the text data as one might expect.

There are two other effects in the verification data. The first of these is an interaction between the two definiteness factors and the relational term ($F_{1,34} = 10.98$, $p < 0.01$). This effect is similar to that found in the comprehension data but the interaction with text, which there appeared the dominant effect, here fails to reach significance. ($F_{1,34} = 2.89$, $p > 0.1$). The results seem to show superior performance with "behind" when the two nominals are differently marked (2044 msec. for different, 2159 msec. for same). The reverse is true for "in front of" (2149 msec. for different and 1972 msec. for same). This makes the overall data here similar to the no text condition in the comprehension data. To that extent the results favour a reduction-to-canonical form analysis.

However this result must be seen in the light of the other interaction.

7 Data are presented in Tables 2, 5 and 8.

Table 5 Verification Data : Analysis of Variance

Factor	$F_{1,34}$ value	Factor	$F_{1,34}$ value
A	* 5.42	AEF	0.96
B	2.87	BCD	2.62
C	1.55	BCE	** 10.98
D	0.05	BCF	0.72
E	0.09	BDE	3.51
F	3.44	BDF	0.18
AB	0.07	BEF	0.01
AC	0.76	CDE	0.65
AD	0.00	CDF	0.20
AE	2.31	CEF	2.48
AF	1.60	DEF	0.63
BC	0.01	ABCD	0.58
BD	2.25	ABCE	2.89
BE	0.84	ABCF	1.45
BF	0.00	ABDE	1.84
CD	0.01	ABDF	0.06
CE	2.18	ABEF	0.00
CF	2.43	ACDE	2.11
DE	0.95	ACDF	0.07
DF	0.42	ACEF	1.60
EF	0.36	ADEF	1.20
ABC	0.79	BCDE	0.22
ABD	* 5.30	BCDF	0.07
ABE	0.52	BCEF	0.02
ABF	0.13	BDEF	2.35
ACD	1.18	CDEF	2.55
ACE	2.46	ABCDE	** 10.51
ACF	0.51	ABCDF	0.01
ADE	0.50	ABCEF	1.68
ADF	0.18	ABDEF	0.19

Factor	$F_{1,34}$ value
ACDEF	2.57
BCDEF	2.43
ABCDEF	0.00

* $p<0.05$

** $p<0.01$

*** $p<0.001$

A : Text/No Text

B : Second Nominal Marking

C : First Nominal Marking

D : Topic Position

E : Relational Term

F : Syntax

Table 6 : Comprehension Data : Summary of Major Effects, First Analysis.[a]

TOPIC POSITION x TEXT $F_{1,34} = 12.91, p < 0.01$

	TEXT	NO TEXT
TOPIC FIRST	2516	3277
TOPIC SECOND	2684	3128

RELATIONAL TERM $F_{1,34} = 4.44, p < 0.05$

"BEHIND"	"IN FRONT OF"
2986	2816

RELATIONAL TERM x TEXT $F_{1,34} = 14.04, p < 0.001$

	TEXT	NO TEXT
"BEHIND"	2541	3432
"IN FRONT OF"	2659	2973

SYNTAX $F_{1,34} = 10.60, p < 0.01$

UNMARKED	MARKED
2785	3017

TOPIC POSITION x SYNTAX $F_{1,34} = 4.47, p < 0.05$

	UNMARKED	MARKED
TOPIC FIRST	2833	2960
TOPIC SECOND	2737	3074

TOPIC POSITION x SYNTAX x TEXT $F_{1,34} = 6.92, p < 0.05$

	TEXT		NO TEXT	
	UNMARKED	MARKED	UNMARKED	MARKED
TOPIC FIRST	2517	2515	3149	3405
TOPIC SECOND	2414	2955	3061	3194

a. All figures are milliseconds.

Table 6 : Comprehension Data : Summary of Major Effects, First Analysis (contd.)

RELATIONAL TERM x SYNTAX x TEXT $F_{1,34} = 6.09, p < 0.05$

	TEXT		NO TEXT	
	UNMARKED	MARKED	UNMARKED	MARKED
"BEHIND"	2458	2625	3239	3625
"IN FRONT OF"	2474	2845	2972	2974

FIRST NOUN MARKING x SECOND NOUN MARKING x RELATIONAL TERM

$F_{1,34} = 5.42, p < 0.05$

FIRST NOMINAL :	"THE"		"A"	
SECOND NOMINAL:	"THE"	"A"	"THE"	"A"
"BEHIND"	3358	2787	2980	3106
"IN FRONT OF"	2894	2732	2894	2742

FIRST NOUN MARKING x SECOND NOUN MARKING x RELATIONAL TERM x TEXT

$F_{1,34} = 8.06, p < 0.01$

TEXT

FIRST NOMINAL :	"THE"		"A"	
SECOND NOMINAL:	"THE"	"A"	"THE"	"A"
"BEHIND"	3052	2473	2523	2687
"IN FRONT OF"	2821	2466	2714	2635

NO TEXT

FIRST NOMINAL :	"THE"		"A"	
SECOND NOMINAL:	"THE"	"A"	"THE"	"A"
"BEHIND"	3664	3101	3438	3525
"IN FRONT OF"	2968	2909	3075	2850

FIRST NOUN MARKING x TOPIC POSITION x RELATIONAL TERM x SYNTAX

$F_{1,34} = 6.58,\ p < 0.05$

UNMARKED

	TOPIC FIRST		TOPIC SECOND	
	"BEHIND"	"IN FRONT OF"	"BEHIND"	"IN FRONT OF"
FIRST NOMINAL "THE"	2808	2732	2753	2833
FIRST NOMINAL "A"	2396	2833	2936	2429

MARKED

	TOPIC FIRST		TOPIC SECOND	
	"BEHIND"	"IN FRONT OF"	"BEHIND"	"IN FRONT OF"
FIRST NOMINAL "THE"	2911	2842	3247	2848
FIRST NOMINAL "A"	3250	2837	3091	3111

FIRST NOUN MARKING x SECOND NOUN MARKING x TOPIC POSITION x RELATIONAL TERM x SYNTAX x TEXT $F_{1,34} = 5.70,\ p < 0.05$

TEXT

UNMARKED

"BEHIND"

FIRST NOMINAL :	"THE"		"A"	
SECOND NOMINAL :	"THE"	"A"	"THE"	"A"
TOPIC FIRST	2440	2030	2614	2780
TOPIC SECOND	2410	2662	2315	2412

"IN FRONT OF"

FIRST NOMINAL :	"THE"		"A"	
SECOND NOMINAL :	"THE"	"A"	"THE"	"A"
TOPIC FIRST	2709	2405	2790	2371
TOPIC SECOND	2460	2783	2100	2172

(Table continued on next page).

COMPREHENSION DATA (CONTINUED)

MARKED

"BEHIND"

FIRST NOMINAL :	"THE"		"A"	
SECOND NOMINAL :	"THE"	"A"	"THE"	"A"
TOPIC FIRST	2279	1973	2356	2897
TOPIC SECOND	2799	3226	2807	2660

"IN FRONT OF"

FIRST NOMINAL :	"THE"		"A"	
SECOND NOMINAL :	"THE"	"A"	"THE"	"A"
TOPIC FIRST	2977	2168	2663	2807
TOPIC SECOND	3140	2509	3304	3192

NO TEXT

UNMARKED

"BEHIND"

FIRST NOMINAL :	"THE"		"A"	
SECOND NOMINAL :	"THE"	"A"	"THE"	"A"
TOPIC FIRST	3817	2945	2587	3603
TOPIC SECOND	2999	2943	3923	3094

"IN FRONT OF"

FIRST NOMINAL :	"THE"		"A"	
SECOND NOMINAL:	"THE"	"A"	"THE"	"A"
TOPIC FIRST	2754	3062	3266	3162
TOPIC SECOND	3040	3049	2703	2741

MARKED

"BEHIND"

FIRST NOMINAL :	"THE"		"A"	
SECOND NOMINAL :	"THE"	"A"	"THE"	"A"
TOPIC FIRST	3874	3518	3764	3985
TOPIC SECOND	3965	2909	3478	3420

(Table continued on next page)

COMPREHENSION DATA (CONTINUED)

MARKED

"IN FRONT OF"

FIRST NOMINAL:	"THE"		"A"	
SECOND NOMINAL:	"THE"	"A"	"THE"	"A"
TOPIC FIRST	3251	2971	3082	2795
TOPIC SECOND	2829	2915	3248	2702

These data are depicted graphically in Figure 1

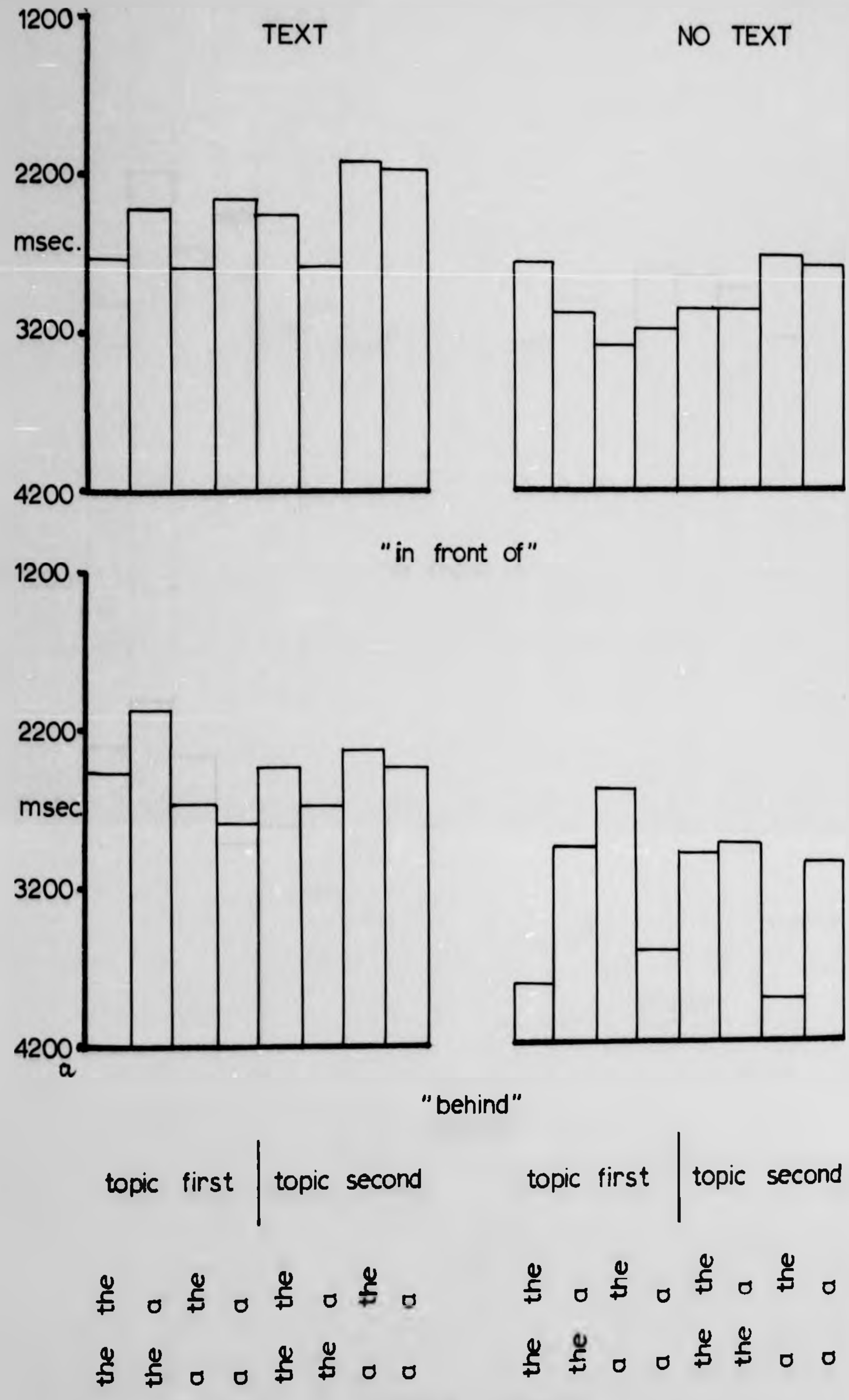


Figure 1 Part 1 — Reaction times to each sentence type, unmarked syntax

a. scales are inverted to facilitate comparison with figs. 1-6 in Chapter 3

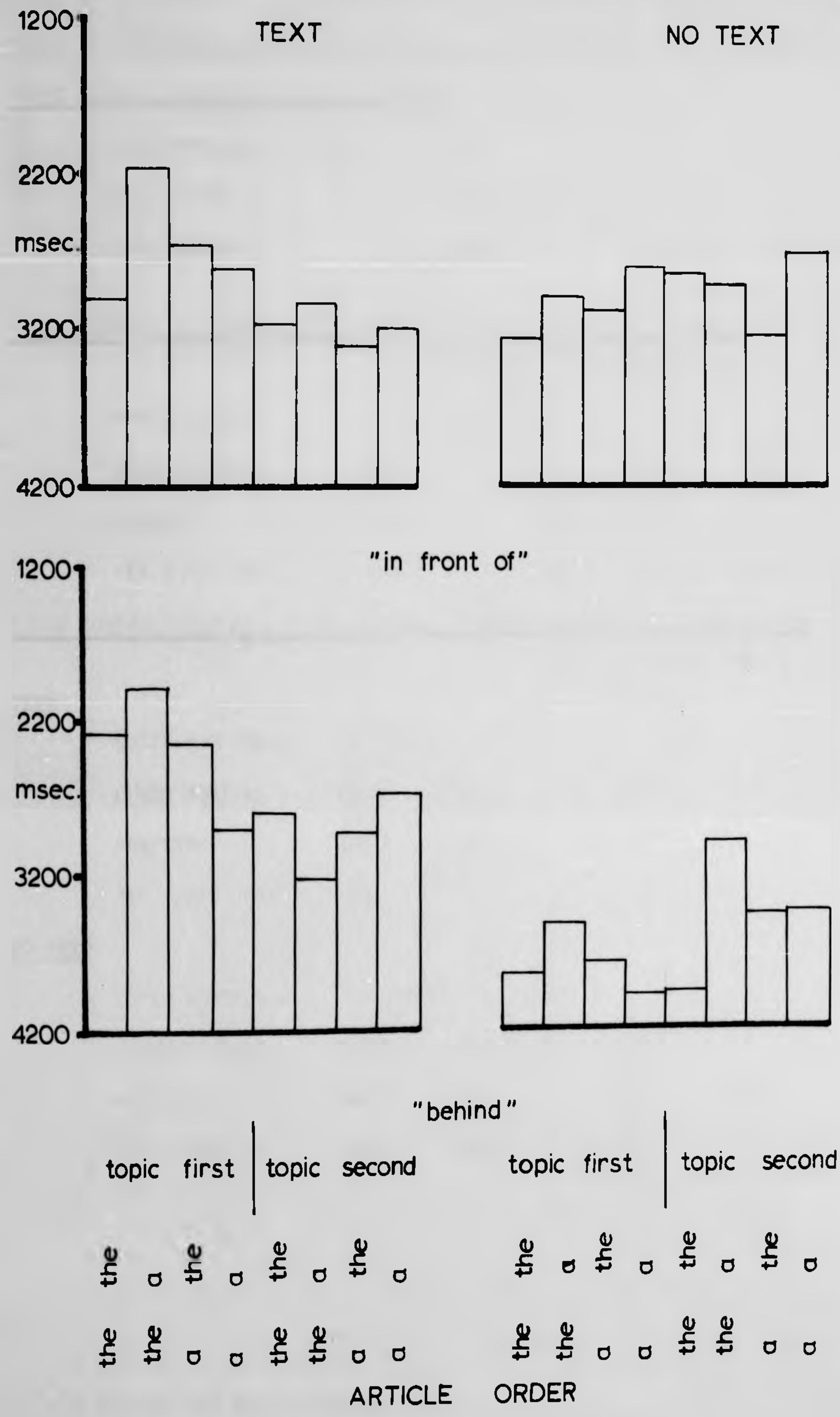


Figure 1 Part 2 — Reaction times to each sentence type, marked syntax

Table 7 : Comprehension Data. Summary of Major Effects, Second Analysis.[a]

TOPIC NOMINAL MARKING x TOPIC POSITION $F_{1,34} = 5.38$, $p < 0.05$

TOPIC NOMINAL :	"THE"	"A"
TOPIC FIRST	2823	2970
TOPIC SECOND	2970	2842

TOPIC NOMINAL MARKING x OTHER NOMINAL MARKING x RELATIONAL TERM

$F_{1,34} = 5.42$, $p < 0.05$

TOPIC NOMINAL :	"THE"		"A"	
OTHER NOMINAL :	"THE"	"A"	"THE"	"A"
"BEHIND"	3358	2973	2850	3106
"IN FRONT OF"	2894	3245	2882	2742

TOPIC NOMINAL MARKING x OTHER NOMINAL MARKING x RELATIONAL TERM x TEXT

$F_{1,34} = 8.06$, $p < 0.01$

TEXT

TOPIC NOMINAL :	"THE"		"A"	
OTHER NOMINAL :	"THE"	"A"	"THE"	"A"
"BEHIND"	3052	2281	2628	2637
"IN FRONT OF"	2821	2494	2686	2635

NO TEXT

TOPIC NOMINAL :	"THE"		"A"	
OTHER NOMINAL :	"THE"	"A"	"THE"	"A"
"BEHIND"	3664	3466	3073	3525
"IN FRONT OF"	2968	2996	3078	2850

a. All figures are milliseconds.

COMPREHENSION DATA (CONTINUED)

TOPIC NOMINAL MARKING x OTHER NOMINAL MARKING x TOPIC POSITION x RELATIONAL TERM x SYNTAX x TEXT $F_{1,34} = 5.70, p < 0.05$

TEXT

UNMARKED

"BEHIND"

TOPIC NOMINAL :	"THE"		"A"	
OTHER NOMINAL :	"THE"	"A"	"THE"	"A"
TOPIC FIRST	2440	2030	2614	2780
TOPIC SECOND	2410	2315	2662	2412

"IN FRONT OF"

TOPIC NOMINAL :	"THE"		"A"	
OTHER NOMINAL :	"THE"	"A"	"THE"	"A"
TOPIC FIRST	2709	2405	2790	2371
TOPIC SECOND	2460	2100	2783	2172

MARKED

"BEHIND"

TOPIC NOMINAL :	"THE"		"A"	
OTHER NOMINAL :	"THE"	"A"	"THE"	"A"
TOPIC FIRST	2279	1973	2356	2397
TOPIC SECOND	2799	2307	3226	2660

"IN FRONT OF"

TOPIC NOMINAL :	"THE"		"A"	
OTHER NOMINAL :	"THE"	"A"	"THE"	"A"
TOPIC FIRST	2977	2168	2663	2807
TOPIC SECOND	3140	3304	2509	3192

COMPREHENSION DATA (CONTINUED)

NO TEXT

UNMARKED

"BEHIND"

TOPIC NOMINAL :	"THE"		"A"	
OTHER NOMINAL :	"THE"	"A"	"THE"	"A"
TOPIC FIRST	3817	2945	2587	3603
TOPIC SECOND	2999	3923	2943	3094

"IN FRONT OF"

TOPIC NOMINAL :	"THE"		"A"	
OTHER NOMINAL :	"THE"	"A"	"THE"	"A"
TOPIC FIRST	2754	3062	3266	3162
TOPIC SECOND	3040	2703	3049	2741

MARKED

"BEHIND"

TOPIC NOMINAL :	"THE"		"A"	
OTHER NOMINAL :	"THE"	"A"	"THE"	"A"
TOPIC FIRST	3874	3518	3764	3935
TOPIC SECOND	3965	3478	2999	3420

"IN FRONT OF"

TOPIC NOMINAL :	"THE"		"A"	
OTHER NOMINAL :	"THE"	"A"	"THE"	"A"
TOPIC FIRST	3251	2971	3082	2795
TOPIC SECOND	2829	3248	2915	2702

Table 8 : Verification Data : Summary of Major Effects

TEXT $F_{1,34} = 5.42,\ p < 0.05$

TEXT	NO TEXT
2435	1730

SECOND NOUN MARKING x TOPIC POSITION x TEXT

$F_{1,34} = 5.30,\ p < 0.05$

	TOPIC FIRST		TOPIC SECOND	
SECOND NOMINAL :	"THE"	"A"	"THE"	"A"
TEXT	2607	2247	2534	2354
NO TEXT	1710	1699	1769	1742

FIRST NOUN MARKING x SECOND NOUN MARKING x RELATIONAL TERM

$F_{1,34} = 10.98,\ p < 0.01$

FIRST NOMINAL :	"THE"		"A"	
SECOND NOMINAL :	"THE"	"A"	"THE"	"A"
"BEHIND"	2211	2012	2076	2107
"IN FRONT OF"	2060	2222	2076	1835

FIRST NOUN MARKING x SECOND NOUN MARKING x TOPIC POSITION x RELATIONAL TERM x TEXT $F_{1,3[illegible]} = 10.51$, $p < 0.0$[illegible]

TEXT

"BEHIND"

FIRST NOMINAL :	"THE"		"A"	
SECOND NOMINAL :	"THE"	"A"	"THE"	"A"
TOPIC FIRST	2609	2190	2364	2454
TOPIC SECOND	2374	2316	2397	2513

"IN FRONT OF"

FIRST NOMINAL :	"THE"		"A"	
SECOND NOMINAL :	"THE"	"A"	"THE"	"A"
TOPIC FIRST	2761	2252	2692	2089
TOPIC SECOND	2180	3267	2301	2204

NO TEXT

"BEHIND"

FIRST NOMINAL :	"THE"		"A"	
SECOND NOMINAL:	"THE"	"A"	"THE"	"A"
TOPIC FIRST	2035	1633	1636	1854
TOPIC SECOND	1827	1908	1911	1556

"IN FRONT OF"

FIRST NOMINAL :	"THE"		"A"	
SECOND NOMINAL :	"THE"	"A"	"THE"	"A"
TOPIC FIRST	1595	1731	1574	1525
TOPIC SECOND	1705	1637	1736	1765

This is a five way interaction involving all factors except syntax. ($F_{1,34} = 10.51$, $p < 0.01$). There is no trace of it whatsoever in the comprehension data ($F_{1,34} = 0.00$). Again it is rather too complex to trace as a whole and is confused by effects in the no text data involving the topic : effects which it is hard to see as anything other than chance occurrences given the nature of the topic factor for this condition. There is little point in describing the result here and the reader is referred to Table 8. The result does serve to make the point that there is still some influence of text even after such a long period. In fact if one includes the time between sentence and picture presentation the average time from onset of the sentence to the true/false response is roughly 10 seconds (10035 msec. for the text condition and 9932 msec. for the no text condition). Even excluding the response time to the picture, times are quite long (7600 msec. in the text condition and [illegible] msec. in the no text condition).

	TEXT COMPREHENSION	TEXT VERIFICATION	NO TEXT COMPREHENSION	NO TEXT VERIFICATION
TEXT COMPREHENSION	-			
TEXT VERIFICATION	+ 0.32†	-		
NO TEXT COMPREHENSION	- 0.05	+ 0.19	-	
NO TEXT VERIFICATION	+ 0.24	+ 0.06	+ 0.53***	-

$\dagger$ $p < 0.1$ df 30

*** $p < 0.001$ df 30

A much simpler set of results appear if we consider the correlation matrix for the mean reaction times for each of the 32 sentences. This is represented above. The only result reaching significance is the correlation between the comprehension and verification times for the no text data ($p < 0.001$). The corresponding result for the text data is significant only at the $p < 0.1$ level. The correlations between the sets of comprehension data is virtually zero. The same applies for the verification data.

Discussion

Comprehension Data

The most striking result of the experiment is undoubtedly the failure to find in the comprehension data a single main effect of any of the factors which is not dependent on some higher order interaction. Only in the verification data is there a clear difference attributable to a main effect, namely the very considerable difference between the text and no text conditions. We will return to this later. In the meantime I want to turn to the main focus of interest of the experiment namely the comprehension times. It was to these that the hypotheses outlined in the introduction to this experiment refer. Since the effects involving the definiteness variables are complex I will treat them last.

Lexical Marking

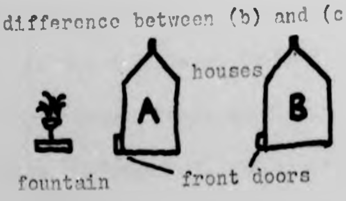
The results concerning the two relational terms constitute a clear replication of Clark's unpublished study (referred to in Clark, 1974) on reaction times to sentences involving these two lexical items. That is, where only one sentence is involved "in front of" appears to be consistently easier to process than "behind". Linguistic reasons for distinguishing these as a marked/unmarked pair are, at best, extremely thin on the ground. There is no sense in which one of them could be said to have superordinate status and indeed there does not even appear to be a superordinate for these two. Even the weakest candidate for marking is absent viz: the frequency difference between the unmarked and marked. This is due not (necessarily) to the fact that their frequencies are similar, but to the fact that we do not know what their frequencies are. This is for a variety of reasons : (1) "in front of" though one (or two) lexical item(s) is three words and frequency studies are in terms of words. (2) "front" can be used in a variety of ways other than in a relational one (3) even "in front of" has at least three distinct meanings : (a) between the observer and the object mention in the locative phrase, (b) independently of the observer, to refer to the area adjacent to an object and towards which that object is facing (c) a third meaning is closely related to the second but in this meaning it is a transitive asymmetrical relation.[8] This third meaning may or may not have a different reference from the second depending on the situation. In the present experiment it did not. These problems together make it impossible to derive an estimate of frequency from the literature on that subject (e.g. Thorndike and Lorge, 1944).

Given all this one might be tempted to discard the marking terminology as irrelevant for this pair. However the RT difference observed means that, from the point of view of a standard psychological measure, they behave like a marked/unmarked pair. Moreover the only alternative terminology which has been put forward (namely "positive" and "negative") has several disadvantages. Firstly it links the words with a theory of the perceptual basis of marking which has been rejected in the introduction. Secondly it is overly suggestive of the evaluative dimension of the semantic differential. Thirdly it is confusing since in phonology the unmarked term is the term with a less complex structure (e.g. unvoiced as opposed to voiced) and hence more accurately labelled negative than positive, whereas Clark's proposed usage for semantics is the opposite of this. Finally it tends to pre-empt any discussion of the relationship between negation and marking.

There is another reason for accepting that this pair constitutes a marked/unmarked pair while at the same time rejecting Clark's original

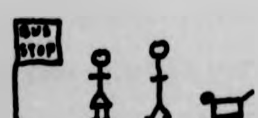
8 to see the difference between (b) and (c), consider the following.

(I)



The fountain is in front of house A, but not house B.

(II)



The girl is in front of both the man and the dog.

interpretation of the marking effect. Clark suggested two alternative interpretations in both of which the unmarked member of the pair has a less complex representation. I suggested in the Introduction that the superordinate may be less complex but that the two subordinates do not differ. They are distinguished only by the fact that the unmarked term will be chosen unless there is a good reason to choose the marked. One of the major predictions outlined in the Introduction to the present experiment is that the importance of topicalisation choices will lead to an interaction between lexical marking and the presence of text. This prediction is amply substantiated by the data. Only when single sentences are presented is the standard marking effect observed. The high order interactions only serve to add detail to this result, they do not force one to qualify it in any important way.

Syntactic Marking and Topic Position

The results with the marked/unmarked syntactic option are less straightforward. Here we observe a main effect of syntax present in both text and no text data. This is not qualified by any simple interaction with the text factor. The prediction outlined in the introduction states that there ought to be some combinations of factors which show no superiority of the unmarked syntactic form in the text case. This is consistent with an overall main effect, though a simple interaction with the text factor would obviously be stronger support. The interaction between text, topic position and syntax is sufficiently strong to give some support to this rather weak hypothesis. This interaction shows no superiority in the text condition of the unmarked over the marked so long as the topic is the first noun (only 2 msec. difference in the means). Having the topic second leads to slower RTs with the marked syntax and faster RTs with the unmarked. This accounts for the significant main effect, together with the fact that there is a constant superiority of the unmarked form in the no text condition.

This three way interaction between text, topic and syntax factors appears to be chiefly responsible for two other effects. The significant interaction between text and topic factors shows a superiority with the topic first in the text condition. As we have just seen this is due to the inferiority with topic second and marked syntax in this condition. The faster RTs with topic second in the no text condition are inexplicable. As noted in the Results section the topic factor is a pseudo-factor for the comprehension latencies in the no text condition. For both these reasons this result is best set aside. The same thing applies to the topic x syntax interaction. This seems to be clearly due to the topic x text x syntax interaction. It shows a pattern of results similar to those for the text condition : superior performance with marked syntax when the topic is first, but superior for unmarked when the topic is second.

The final effect not involving the definiteness factors is that between text, relational term and syntax. This is perhaps best expressed as follows : with text there is no difference between the relational terms when unmarked syntax is presented but "behind" is easier with marked syntax; without text there is no difference between the two types of syntax if "in front of" is chosen but the marked syntax is much harder with "behind". This is rather difficult to interpret. Possibly a better way of expressing the result is to say that "in front of" plus marked syntax is a noticeably harder type of sentence than the other three sorts in the text case. In the no text case it is "behind" plus marked syntax which stands out as taking rather longer. On either description I can see no easy explanation for the result.

Summary of Results not involving definiteness factors and suggested explanations.

It may be as well to summarise the discussion so far. Essentially I have said that the three results requiring explanation are:

(a) the text x relational term interaction

(b) the text x topic x syntax interaction

(c) the text x relational term x syntax interaction.

The other results are derivable from these.

Result (a) supports hypothesis 2 (that the marked term does not necessarily take longer to understand).

Result (b) supports both hypothesis 3 and hypothesis 4 (that the marked syntax does not necessarily take longer to understand and that the position of the topic in the sentence matters). I have not so far explained this result.

Result (c) was not predicted and I cannot explain it. Since it includes the factors of Result (a) it must affect interpretation of that Result. It does not affect the rejection of Clark's versions of the marking theory but it will be relevant to the construction of an alternative such as that already suggested.

The following explanation is suggested for Result (b). First position in the sentence is important because it represents a foregrounding of the speaker's point of departure for the sentence (in Halliday's terms). In the present situation in the text condition there are two natural points of departure : the object already talked about because it is already given prior to the target sentence, and the new object because it has to be "solved for" (in Huttenlocher's terms) - i.e. one has to find out where it fits into the picture. It was suggested in the introduction that the marked syntax may serve like the passive to keep the theme = old noun without the necessity for marked theme : this is a way of starting with what is given. This explanation is consistent with the fact that there was no difference found between syntactic types for the text condition when the topic was first. The speaker has another option : to take the addition of a second object as his point of departure. In that case, on analogy with the passive, the marked syntax should be harder. And this is what we find. However this explanation fails to account for an important distinction between these relational sentences and transitive sentences. That is the possibility of changing the relational term, without change of syntax, in order to change the order of nouns - and so the 'point of departure for the sentence' - while preserving truth. Why use the syntactic option when the lexical option is available? The answer to this may be simply that the lexical choice is a choice at a different level. It is simply made at a point "later" than the choices about theme. Only in the case where thematic options seem unlikely to have mattered in the construction of the sentence (i.e. in the one sentence case) does any consideration of the lexical option as a choice in itself arise.

At first sight one is tempted to adopt what looks like a much simpler explanation to account for the topic x syntax interaction. We have already encountered Clark's suggestion that the locative phrase is the natural reference point for the sentence and Huttenlocher's related suggestion that it is easier to add an object to a display when it is the grammatical subject of a sentence describing what the array will be like when it is added. Tasks like the present one can be considered as "mental analogues" of Huttenlocher's placement tasks : in the text case one is listening to (reading) the sentence with a view to "placing" the new item in the mental array already containing the topic. Now the topic position x syntax interaction can be expressed by saying that sentences are easier when the previously mentioned noun is in the locative phrase or, equivalently, that they are easier when the new item is grammatical subject of the sentence. This view has two major objections to it. Firstly it is not capable of extension to transitive sentences other than by vague analogy.[9] Secondly it cannot account for the difference in the effect depending on syntax. On the present data (with the text condition) the effect is only 103 msec. with unmarked syntax but 440 msec. with marked syntax. This result seems inexplicable on the Clark/Huttenlocher account.

9 In fact Huttenlocher's account of relationals is by analogy with transitives which she considers primary. Clark seems to have drawn the analogy in the opposite direction.

Results involving definiteness marking

Turning now to the effects involving definiteness marking. As noted in the Results section these are exceedingly complex. Contrary to Hypothesis 1 of the Introduction there does not appear to be any simple effect of the definiteness marking of either the topic or the other noun. This contrasts with the convincing evidence of the importance of definiteness marking presented by Grieve and Wales (1973). However there are a number of complex interactions in which definiteness marking plays a part. As already noted the definiteness variables can be defined in two ways : in terms of definiteness of the topic noun and the other noun, and in terms of definiteness of the first noun and the second noun.

There is a significant interaction of first noun, second noun and syntax, but this appears to be largely the result of a four way interaction involving these three and text. As pointed out in Results "behind" in both conditions and "in front of" in the text condition only, all show faster RTs when the two nominals are differently marked. "In front of" in the no text condition leads to faster RTs when the two nominals are marked the same (be it both with "the" or both with "a"). The "in front of" results for the text condition show a noticeably smaller effect than the "behind" results. Both sets of "behind" results show a superiority of over 300 msec. when the nominals are marked differently, but this is reduced to an average of only 143 msec. for "in front of". It is this which leads to the significant three way interaction (without the text factor).

We can redefine the variables here in terms of marking of the topic and other nouns. This, as one might expect given the pseudo-nature of this variable for the no text condition, does not really affect the picture for the no text data. It does give a slightly different picture for the text data. It appears that there is no effect of marking of the other noun if the topic is marked indefinite (mean 2659 msec., range 2628 msec. to 2687 msec.) whereas if the topic noun is definitely marked there are much faster RTs if the other noun is indefinite (2936 msec. when definite, 2387 msec. when indefinite). As noted in Results these are clearly non additive :

both nominals marked correctly[10] is fastest but the intermediate times are for an average of both incorrect and only one incorrect (i.e. topic marked wrongly with or without the other nominal marked wrongly); the longest times are for only one incorrect (i.e. topic correct and the other noun not). These data bring out the importance of the relationship between definiteness marking and previous mention, though it is hard to see quite what the process is which gives rise to such results.

The three other results involving the definiteness factors are less useful than these. The four way interaction between definiteness of the first noun, topic position, relational term and syntax is not very meaningful in that it includes the topic position factor but not the text factor. The same thing applies to the topic position x topic noun marking interaction found on the topic marking/other marking analysis. The final effect, namely the six way interaction, though it is potentially of considerable interest, is really much too complex to explain - or even to describe concisely. It remains as a warning that one is likely to be oversimplifying if one does not explain it, so that conclusions will have to be suitably tentative.

To return for a moment to the four way interaction involving text, relational term and both definiteness factors. We have seen that this can be expressed in two ways depending on how the definiteness variables are defined. Either way the results for the no text condition remain essentially the same : "Behind" appears to give rise to faster RTs when the nominals are marked differently than when they are the same; the reverse seems to be true of "in front of". Although the "in front of" results are more like the "behind" results when text is presented there still seems to be some residual tendency towards better performance than "behind" for the cases where both nominals are marked the same.

10 "correctly" and "wrongly" are here used on the basis of the assumption that the topic should be marked with "the" and the new noun with "a".

This result is interesting in that it relates to the hypothesised importance of topic decisions in choosing the relational term. I have suggested that topicalisation choices constitute a good reason for choosing the marked term. In the absence of any desire to topicalise one noun rather than another the unmarked lexical item should be chosen. The clearest case where one would not wish to differentially emphasise the nouns seems to be (as a naive, intuitive first approximation) where one marks both nominals indefinitely. Marking both definitely is a more complex case but there are probably less reasons for differential emphasis here than where the two nominals are marked differently. It is not surprising that "in front of" comes out better with these two "same marking" cases than with the "different marking" cases. On the account I am giving that is simply another aspect of its lexical unmarkedness.

Before going on to the verification data, here is a summary of the five effects which seem to be of most importance.

1. Text x Relational Term. This was considered to be a falsification of Clark's view(s) of marking and support for the alternative view presented here.

2. Text x Topic position x syntax. This shows the importance of the position of the previously mentioned item. It constitutes partial support for the very weak hypothesis of Hornby viz that different types of syntax have different topic/comment structures. It also provides support for the psychological interpretation of systemic grammar given here. An alternative explanation based on the Clark/Huttenlochor work and stressing the importance of the locative phrase as a reference point and/or the extra ease of having the grammatical subject as the new item, receives partial support also. However it fails to account for the fact that the effect is bigger with marked syntax.

3. Text x Relational Term x Syntax. This result was not predicted and remains unexplained.

4. Text x first noun marking x second noun marking x relational term <u>OR</u>

Text x Topic noun marking x other noun marking x relational term. This result was not predicted but a partial explanation resting upon the nature of the lexical marking effect is suggested.

5. The six way interaction. This result was not predicted nor is it explicable, because of its complexity. It serves as a warning of the complexity of the phenomena and forces one to be fairly tentative about one's conclusions.

<u>Verification Data</u>

The verification data are of only secondary importance in the present experiment. Only one analysis was performed for that reason. For this the definiteness variables were defined in terms of the position of the nominals in the sentence.

Models of the verification process based upon canonical structures tend to minimise the importance of surface structures as an influence on verification. Even with the relatively long time lapses between initial presentation of the sentence and onset of the picture one might expect some influence of surface structure. That was what Gough found. But it is hard to see quite what influence the presence or absence of previous text should have on verification times if one believes the canonical structure model. Accordingly it is revealing that three of the four significant effects which come out on the RT analysis involve the text factor, including an overall main effect of that factor.

This fact runs strongly counter to <u>both</u> the canonical form view and the surface structure encoding view, as well as it would seem any hybrid such as the Garrod and Trabasso model. The presence of effects dependent on both previous presentation of text and surface structure militates against the canonical form view. But the influence of previous presence of text is not in any obvious way accountable for by a purely surface structure view. What seems to be needed is a model which (a) stresses the importance of topicalisation features and (b) relates these in a fairly simple way to surface structure. Such a model is embodied in SQ, though of course this is not the only possible model with those properties.

This conclusion is reinforced by the correlations performed between the 32 means for the different sentence types. As detailed in the introduction to the present experiment the following predictions seem to follow from the three global views.

1. Canonical form : some surface information may affect verification but in general the two sets of verification times ought to be more highly correlated than each is with its corresponding comprehension set.

2. Surface structure : this predicts high correlations between all four sets of data.

3. Topic structure (including SG) : a low correlation between the two sets of comprehension data. Positive correlations between each set of comprehension data and its corresponding verification set, perhaps higher in the single sentence case because of the negligible influence of topicalisation interpretations on the comprehension data.

In fact the data tend to support the third view : the two sets of comprehension data do not correlate ($r_{30} = -0.05$); neither do the two sets of verification data ($r_{30} = 0.06$). The only significant correlation is between the no text comprehension and verification times ($r_{30} = 0.53$, $p < 0.001$), with the corresponding figure for the text data not quite significant ($r_{30} = 0.32$, $p < 0.1$). Predictions 1 and 2 are clearly falsified by these data.

The high correlation between comprehension and verification data for the no text condition suggests that similar processes were at work in both cases. The rather lower correlation for the text data suggests the possibility that different (or partially different) processes may have been involved. However this set of data is not really sufficient to base any positive conclusions upon.

Chapter 3 : Two experiments on the production of sentences describing simple spatial relationships.

The present chapter reports two experiments on the production of simple relational sentences like those used in the experiment reported in the last chapter. The first of these experiments requires subjects to write down a sentence to describe a picture presented to them. They are constrained to write sentences of the sort used in the verification experiment by the use of sentence frames in which slots are labelled with the name of the "part of speech" which is to be inserted there ("noun", "article", "relational term"). This method has the disadvantage that the syntactic type is wholly determined by the experimenter so that this factor can only be examined in so far as it interacts with other factors. On the other hand the experimenter does have the option of inserting lexical items in to one or more of the "parts of speech" slots and so has the possibility of systematic manipulation of variables. In this written sentence study an attempt is made to manipulate the text factor by presenting subjects with a passage to read prior to the filling in of the sentence frame. The passage describes one of the objects in the picture. This method has the disadvantage that subjects are forced to switch from the passive role of reader to the active role of writer. This must inevitably make the passage of less relevance to the writing task than one would wish.

The second study is designed to complement the first in several respects. Firstly sentences are spoken instead of written. Secondly the only manipulations are rather indirect, and there is no constraint on the kind of sentence to be used. Thirdly the disadvantage of having subjects switch from the passive to the active role is avoided by having subjects generate the prior text themselves. The chief disadvantages of this method are that it severely limits the power of the experimenter to systematically manipulate the situation, and the yield of usable data is very low.

Experiment 2

Method

Subjects

60 subjects fulfilling a course requirement for an introductory psychology course at Stirling University.

Apparatus and Materials

72 coloured slides each depicting two objects seen in profile so that one could be said to be "behind" the other. Slides were mostly of model animals made by Britains Ltd. but some were of people, vehicles (real and model) and sundry other objects with fronts and backs. There were 2 slides for each pair of objects - one with one object in front and one with the other.

Slides were back projected onto a screen about 1 metre square about 1½ metres from the subject who sat at a desk with a remote control for the projector, a pen and the book of sentences frames. Each book contained 36 sentence frames. "Basic frames" were of two sorts, corresponding to the two syntactic types viz.

______ Relational Term ______ Article ______ Noun is ______ Article ______ Noun

or

______ Article ______ Noun is ______ Relational Term ______ Article ______ Noun

Most frames had one lexical item inserted already : a noun or an article or the relational term. Details of this are given below. Half the subjects had an additional deck of 36 cards on each of which was typed a short paragraph[a] describing one of the objects in the accompanying picture.

Design and Procedure

Subjects were evenly divided into two groups. One group received the deck of context cards, one group did not. All subjects received a book of sentence frames made up as follows.

1. four frames - two of each type - had no lexical items (other than "is") inserted. One of each type was presented with a picture in which the topic

[a] These sentences are listed in Appendix A.

was in front, and one with the topic behind. As in the previous experiment the topic was specified for the no context condition by correspondence with that for the text condition. As before, though, the topic was always in the middle of the picture with the other object either to the left or right of it. (Unlike the previous experiment orientation of the picture was random to left or right).

2. four frames - two of each type - had "the" inserted in the first article slot.

3. Similarly for "a" in the first article slot,

"the" in the second article slot,

"a" in the second article slot,

the topic noun in the first noun slot,

the topic noun in the second noun slot,

"in front of" inserted

and finally "behind" inserted.

Although each subject responded to two sentence frames of each type one of each of those was presented with a picture in which the topic was in front and one with the topic behind. In effect then there is only one response per subject per sentence/picture type.

The 36 picture/object-pairs occurred in the same order for all subjects but the 36 sentence types were randomly assigned to the pictures with a different randomisation for each subject. The only constraint to this was that subjects were matched across the text/no text factor for random order.

Subjects were brought in and sat down at the table and told that we were interested in how people described pictures given certain constraints. They were shown a sample picture and two blank sentence frames, and told that for most cases a word would already be inserted. They were to complete each sentence in as natural a manner as possible and were then to rate it for acceptability. For this they used a 5-point scale as follows:-

1	2	3	4	5
Perfectly Acceptable	Acceptable	Marginal	Unacceptable	Perfectly Unacceptable

Most subjects had no difficulty in representing this rating task to themselves, usually in terms of vague notions of what is good grammar. A few had difficulty and were encouraged to think along those lines.

Though the rating task is interesting in itself no presentation of the results derived from it is given here. Its main function was to maintain subjects' interest.

Subjects controlled rate of presentation themselves. They were told not to turn the page of the sentence booklet until the companion slide was visible. Those who received context cards were told to read these prior to advancing the slide projector to the slide to which the context card referred.

The experiment lasted about 15 - 20 minutes without context cards and about 40 minutes with context cards. Subjects' reactions to it varied from great interest to mild boredom.

Results

The results of this experiment were tabulated in terms of contingency tables using six criteria : definiteness of the first nominal, definiteness of the second nominal, position of the topic noun, relational term, syntax and presence of text. The results for blank frames, frames with an article inserted and frames with the relational term or topic inserted were tabulated separately. The reasons for this are:-

(1) the blank frames leave the subject relatively unrestricted so they ought to produce more "natural" results.

(2) the relational term and topic noun insertion are tabulated together because the presence of either of these fully determines the other given truth and a particular syntactic type.

(3) insertion of an article leaves the choice as to relational term or topic position relatively unconstrained, but presumably not so unconstrained as in the blank frames.

Because of the complex nature of the results it is rather hard to pick out effects directly from the contingency tables. Since there is only one response per subject per sentence type/ picture combination one could make out a case for assuming independence of the observations. Obviously there are counter-arguments to this, but it does not seem an unreasonable position to adopt at this stage. If one is prepared to accept it then the appropriate method of analysis would be in terms of some kind of multi-dimensional χ^2. Unfortunately there is no readily accepted method of calculating χ^2 for the m-way table ($m > 3$). Goodman (1969) presents a method of calculation for the three-way table, but uses a different method for the general case (Goodman, 1971). Calculation of χ^2 for the six way tables in the present experiment is extraordinarily difficult using this method. Given the arguable nature of the independence assumption it seems unnecessary to go to such lengths.

For this reason a much simplified method is used here. This involves a straightforward subtractive logic, which is best explained by means of an example. Imagine the following contingency table.

	Y_1		Y_2	
	X_1	X_2	X_1	X_2
Z_1	5	7	8	4
Z_2	10	3	6	9

We first calculate χ^2_x to assess for main effects of x; similarly for y and z. Next the χ^2s for the x y, yz and xz 2 x 2 tables are calculated. Finally the χ^2 for the xyz table is calculated. But the interaction χ^2 are inflated by the presence of main effects so these are then subtracted. So we have the following:-

$$
\begin{aligned}
\text{"real" } \chi^2_x &= \text{Calculated } \chi^2_x \\
\text{" } \chi^2_y &= \text{" } \chi^2_y \\
\text{" } \chi^2_z &= \text{" } \chi^2_z \\
\text{" } \chi^2_{xy} &= \text{" } \chi^2_{xy} - \chi^2_x - \chi^2_y \\
\text{" } \chi^2_{xz} &= \text{" } \chi^2_{xz} - \chi^2_x - \chi^2_z \\
\text{" } \chi^2_{yz} &= \text{" } \chi^2_{yz} - \chi^2_y - \chi^2_z \\
\text{" } \chi^2_{xyz} &= \text{" } \chi^2_{xyz} - \chi^2_x - \chi^2_y - \chi^2_z - \text{"real" } \chi^2_{xy} \\
&\quad - \text{"real" } \chi^2_{xz} - \text{"real" } \chi^2_{yz}
\end{aligned}
$$

This method has the disadvantages that (a) there is not a uniform probability of rejecting Ho for all interactions (b) lower order effects tend to be overestimated and (c) the nature of the subtractive logic means that there is a cyclical error as one goes from lower to higher order interactions (if one overestimates a main effect one underestimates first order interactions, overestimates second order interactions etc.).

These are fairly grave disadvantages, but they are offset by the fact that the method is used here only as a kind of shorthand - we are not, strictly speaking, using it as a method of testing hypotheses. Given the difficulties I will here restrict discussion to those effects which are estimated by the present method to have a probability of less than 0.001 of occurring by chance. Results are tabulated in Tables 1 - 7. Table 1 gives the χ^2 results for all three data sets. Table 2 gives the full categorisation of responses from data set 1 and Table 3 lists all the effects significant at $p < 0.05$ or better for that data set complete with means. Table 4 gives the second set of results completely categorised and Table 5 lists the effects for that set (relational term or topic inserted). Tables 6 and 7 give the corresponding information for the blank sentence frame data. The data from Tables 2, 4 and 6 are also presented graphically in Figures 1 - 6 (Two figures for each data set).

The three χ^2 tables show very similar results on the $p < 0.001$ criterion. In all three the first nominal is more likely to be marked by "the", and the second more likely to be marked by "a". In all three these effects are enhanced if the topic is first and reduced if it is second, though this two way interaction is clearly largely due to the three way interaction with text as topic position tends to be the overriding influence in the text condition. This is because there is a strong tendency to mark the topic noun "the" and the other noun "a". All three sets of data also show a tendency to have the two nominals marked differently for definiteness,[1] though this is larger when text is presented. All three also show a very strong tendency to have the second nominal indefinitely marked with marked syntax.

Some effects appear in only one set of data. Only the data from the cases where either the topic or the relational term is inserted show a two way interaction between text and marking of the first nominal - the no text data showing a very strong tendency to have this "the". This effect appears to be additional to the effects mentioned in the last paragraph. The blank frame data shows two effects not present in the other two data sets. Firstly there is a clear tendency to put the topic first. This effect appears to be unmodified by the type of syntax (unlike in the last experiment). Secondly there is an interaction between topic position, text and marking of the two nominals (though this only reaches the $p < 0.01$ level). The tendency to have the two nominals marked differently appears to be restricted chiefly to cases where the topic is

1. this effect falls below the α level for the article-inserted data.

first and the articles are ordered "the - a" in the text condition. The data for cases where an article is inserted show a tendency for the first article to be "the" if "behind" is the relational term, but this is not replicated in the other data.

There are a number of cases where effects occur in one set of data but fall below the α level for the others, though still with some effect. Thus the second set of data (relational term or topic inserted) show a greater likelihood of the first noun being "the" with marked syntax. This effect is much reduced in the other data, though both the second and third (blank frame) sets show a strong tendency for the first nominal to be definite and the second indefinite with marked syntax. This is further modified by the presence of a four way interaction involving both definiteness factors, text and syntax. This result shows that the tendency to have the two nominals marked differently for definiteness is restricted to cases of marked syntax with the order "the - a" or unmarked syntax with both orders in the text condition; in the no text condition it is restricted to marked syntax with the order "the - a". Finally there is an interaction between the two definiteness factors, topic position and syntactic type present in both the first (articles inserted) and second sets of data. This showed that the greater frequency of "the - a" in marked syntax is greatly reduced when the topic is second - though this is only in the text condition. Even then it still remains the most common combination with marked syntax. This effect presumably does not occur in the blank frame data because of the topic main effect, and it is greatly reduced in the article-inserted data because of the relative freedom which that allows.

Table 1 Experiment 2 : χ^2 Tables.[a]

Effect (All df = 1)	Article Inserted	Relational Term or Topic Inserted	Blank Frame
A	0.77	0.01	0.11
B	* 6.59	*** 17.30	*** 22.81
C	*** 41.21	***157.40	*** 70.70
D	2.01	0.19	*** 13.68
E	* 4.68	0.13	2.83
F	0.11	0.03	0.04
AB	0.37	1.21	1.31
AC	0.66	*** 25.92	0.01
AD	0.17	0.00	2.84
AE	0.55	0.01	2.39
AF	0.02	0.02	0.04
BC	* 4.98	*** 26.61	** 11.77
BD	*** 75.99	*** 67.56	*** 10.87
BE	1.17	1.36	0.04
BF	*** 39.5	***223.26	*** 40.84
CD	*** 70.21	*** 61.15	*** 22.82
CE	*** 12.82	0.19	2.83
CF	* 6.6	*** 17.30	3.81
DE	0.00	0.03	0.01
DF	0.03	0.00	0.37
EF	*** 16.99	0.01	* 6.19
ABC	*** 70.20	*** 22.01	** 7.60
ABD	*** 45.66	*** 39.56	*** 22.80
ABE	0.55	0.93	0.77
ABF	1.17	** 7.27	2.39
ACD	*** 60.38	*** 42.98	***12.69
ACE	0.90	0.05	0.55

a Factors are : A : Preamble/No Preamble
B : Second Nominal Marking : "the" or "a"
C : First Nominal Marking : "the" or "a"
D : Position of the Topic : First or Second
E : Relational Term : "Behind" or "In front of"
F : Syntax : marked or unmarked.

(Cont'd.)

Effect (All df = 1)	Article Inserted	Relational Term or Topic Inserted	Blank Frame
ACF	0.00	0.06	0.36
ADE	0.02	0.19	0.00
ADF	0.54	0.25	0.00
AEF	0.90	0.35	0.12
BCD	0.89	0.69	1.99
BCE	3.33	2.04	0.22
BCF	1.16	*** 112.68	*** 31.17
BDE	0.77	0.40	0.00
BDF	0.45	0.40	1.30
BEF	* 4.97	2.24	** 6.88
CDE	0.08	1.36	1.01
CDF	0.10	1.86	1.30
CEF	3.09	1.51	0.04
DEF	0.65	0.18	0.04
ABCD	1.33	0.26	** 10.01
ABCE	0.46	0.26	0.76
ABCF	0.56	*** 32.39	* 5.45
ABDE	0.30	0.00	0.37
ABDF	0.12	2.88	2.84
ABEF	0.65	0.68	1.31
ACDE	0.00	0.59	0.77
ACDF	2.23	0.82	0.13
ACEF	2.41	0.19	0.36
ADEF	0.12	0.00	0.04
BCDE	0.02	3.25	0.76
BCDF	* 4.40	*** 27.31	2.00
BCEF	0.07	0.05	0.55
BDEF	0.66	1.52	0.12
CDEF	0.05	0.20	0.04
ABCDE	0.00	0.92	0.38
ABCDF	0.44	2.42	2.44
ABCEF	-0.09 b	0.00	1.45
ABDEF	0.11	1.22	0.10
ACDEF	0.28	0.78	0.27
BCDEF	0.68	-1.62 b	0.23
ABCDEF	0.01	2.37	-1.22 b

* $p < 0.05$

** $p < 0.01$

*** $p < 0.001$

b. These estimates of χ^2 are obviously inaccurate, but probably not by very much.

Table 2 Experiment 2 : Results, Article Inserted data.[a]

Text

		"Behind"				"In front of"			
Second Nominal:		"The"		"A"		"The"		"A"	
First Nominal :		"The"	"A"	"The"	"A"	"The"	"A"	"The"	"A"
Unmarked Syntax	Topic First	15	2	35	5	14	3	32	11
	Topic Second	11	35	4	3	8	35	3	6
Marked Syntax	Topic First	11	0	48	11	7	0	35	4
	Topic Second	14	27	16	10	7	22	7	10

No Text

Unmarked Syntax	Topic First	17	9	10	10	22	10	12	19
	Topic Second	22	8	8	8	23	10	9	14
Marked Syntax	Topic First	19	1	37	16	11	5	14	14
	Topic Second	17	0	25	16	13	2	7	17

a. These data are illustrated in Figures 1 and 2.

Figure 1 Article inserted data, unmarked syntax

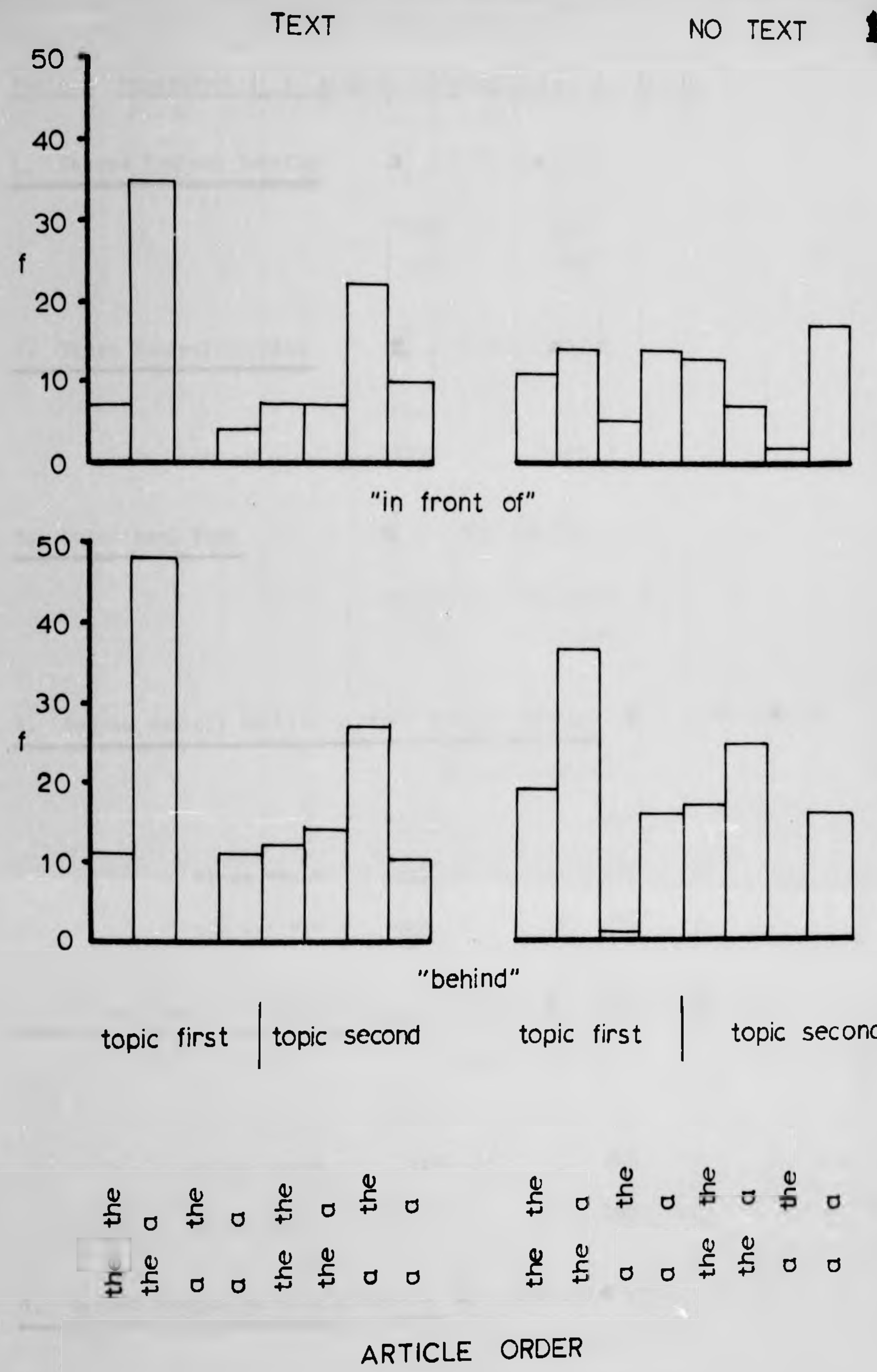


Figure 2 Article inserted data, marked syntax

Table 3 Experiment 2 : Article Inserted Data : Effects

1. Second Nominal Marking $\chi^2_1 = 6.59$, $p < 0.05$

"The"	"A"
400	476

2. First Nominal Marking $\chi^2_1 = 41.21$, $p < 0.001$

"The"	"A"
533	343

3. Relational Term $\chi^2_1 = 4.68$, $p < 0.05$

"Behind"	"In front of"
470	406

4. Second Nominal Marking x First Nominal Marking $\chi^2_1 = 4.98$, $p < 0.05$

	Second Nominal	
	"The"	"A"
First "The"	231	169
Nominal "A"	302	174

5. Second Nominal Marking x Topic Position $\chi^2_1 = 75.99$, $p < 0.001$

	Second Nominal	
	"The"	"A"
Topic First	146	313
Topic Second	254	163

6. Second Nominal Marking x Syntax $\chi^2_1 = 39.5$, $p < 0.001$

	Second Nominal	
	"The"	"A"
Unmarked Syntax	244	189
Marked Syntax	156	287

7. First Nominal Marking x Topic Position $\chi^2_1 = 70.21$, $p < 0.001$

	First Nominal "The"	First Nominal "A"
Topic First	339	120
Topic Second	194	223

8. First Nominal Marking x Relational Term $\chi^2_1 = 12.82$, $p < 0.001$

	First Nominal "The"	First Nominal "A"
"Behind"	309	161
"In front of"	224	182

9. First Nominal Marking x Syntax, $\chi^2_1 = 6.6$, $p < 0.05$

	First Nominal "The"	First Nominal "A"
Unmarked Syntax	245	188
Marked Syntax	288	155

10. Relational Term x Syntax $\chi^2_1 = 16.99$, $p < 0.001$

	"Behind"	"In front of"
Unmarked Syntax	202	231
Marked Syntax	268	175

11. Text x First Nominal Marking x Second Nominal Marking $\chi^2_1 = 70.20$, $p < 0.001$

Text	First Nominal "The"	First Nominal "A"	No Text	First Nominal "The"	First Nominal "A"
Second "The"	87	124		144	45
Nominal "A"	180	60		122	114

12. Text x Second Nominal Marking x Topic Position $\chi^2_1 = 45.66$, $p < 0.001$

Text	Second Nominal "The"	Second Nominal "A"	No Text	Second Nominal "The"	Second Nominal "A"
Topic First	52	181		94	132
Topic Second	159	59		95	104

13. Text x First Nominal Marking x Topic Position $\chi^2_1 = 60.38$, $p < 0.001$

Text	First Nominal "The"	First Nominal "A"	No Text	First Nominal "The"	First Nominal "A"
Topic First	197	36		142	84
Topic Second	70	148		124	75

14. Second Nominal Marking x Relational Term x Syntax $\chi^2_1 = 4.97$, $p < 0.05$

"Behind"	Second Nominal "The"	Second Nominal "A"	"In front of"	Second Nominal "The"	Second Nominal "A"
Unmarked Syntax	119	83		125	106
Marked Syntax	89	179		67	108

15. Second Nominal Marking x First Nominal Marking x Topic Position x Syntax $\chi^2_1 = 4.40$, $p < 0.05$

Unmarked Syntax	Topic First		Second Nominal "The"	Second Nominal "A"	Topic Second	Second Nominal "The"	Second Nominal "A"
	First	"The"	68	89		64	24
	Nominal	"A"	24	45		88	31
Marked Syntax							
	First	"The"	48	134		51	55
	Nominal	"A"	6	45		51	53

Table 4 Experiment 2 : Results : Topic or Relational Term Inserted Data.[a]

Text

		"Behind"				"In front of"			
	Second Nominal:	"The"		"A"		"The"		"A"	
	First Nominal :	"The"	"A"	"The"	"A"	"The"	"A"	"The"	"A"
Unmarked	(Topic First	22	2	26	6	19	1	26	10
Syntax	(Topic Second	15	41	0	3	4	33	2	5
Marked	(Topic First	2	1	50	5	1	0	50	7
Syntax	(Topic Second	11	16	18	9	8	16	17	14

No Text

Unmarked	(Topic First	34	7	10	8	40	2	4	13
Syntax	(Topic Second	33	11	3	8	29	5	13	9
Marked	(Topic First	3	1	52	4	1	0	42	9
Syntax	(Topic Second	7	0	40	9	11	3	37	5

a These data are illustrated in Figures 3 and 4.

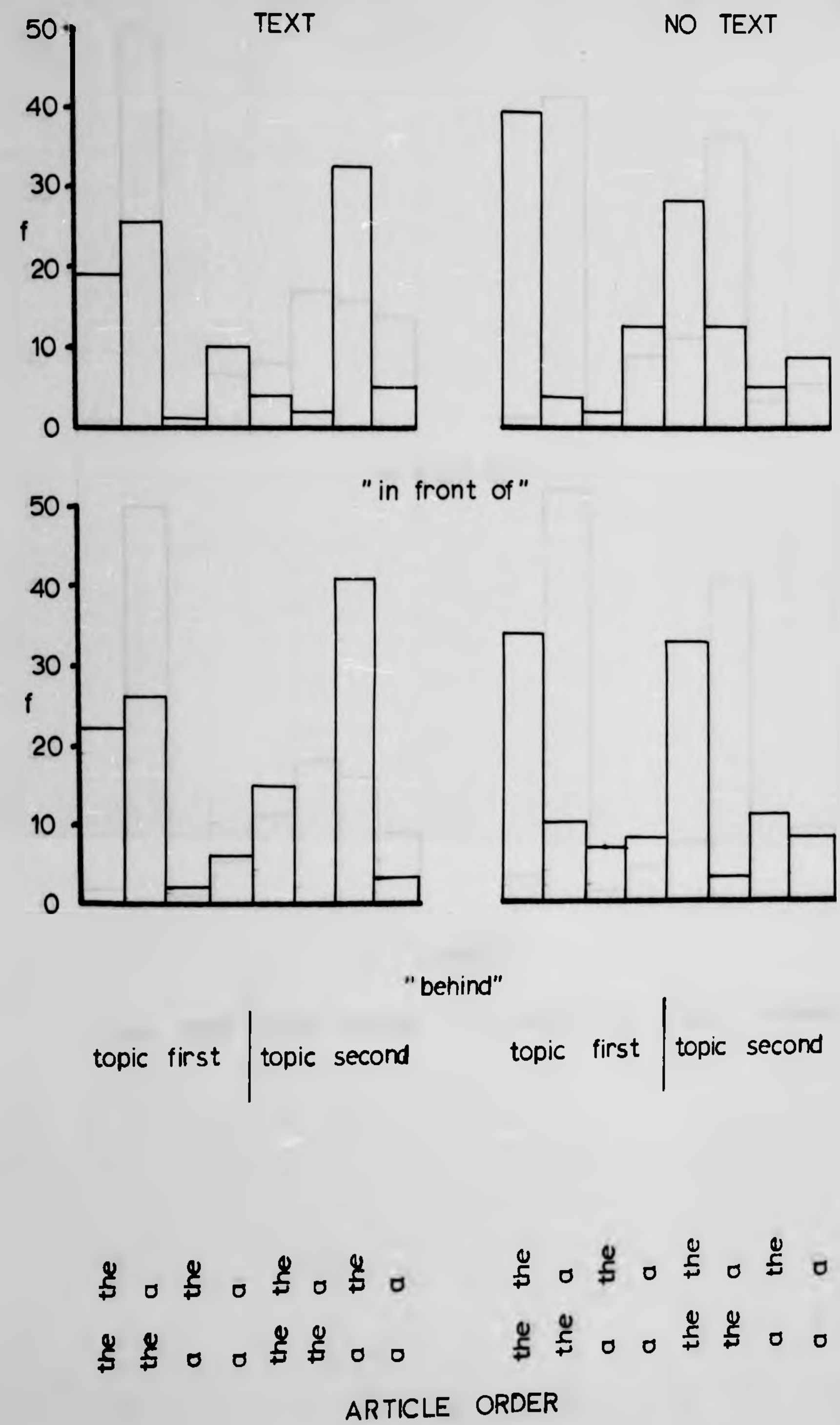


Figure 3 Relational term or topic inserted data, unmarked syntax

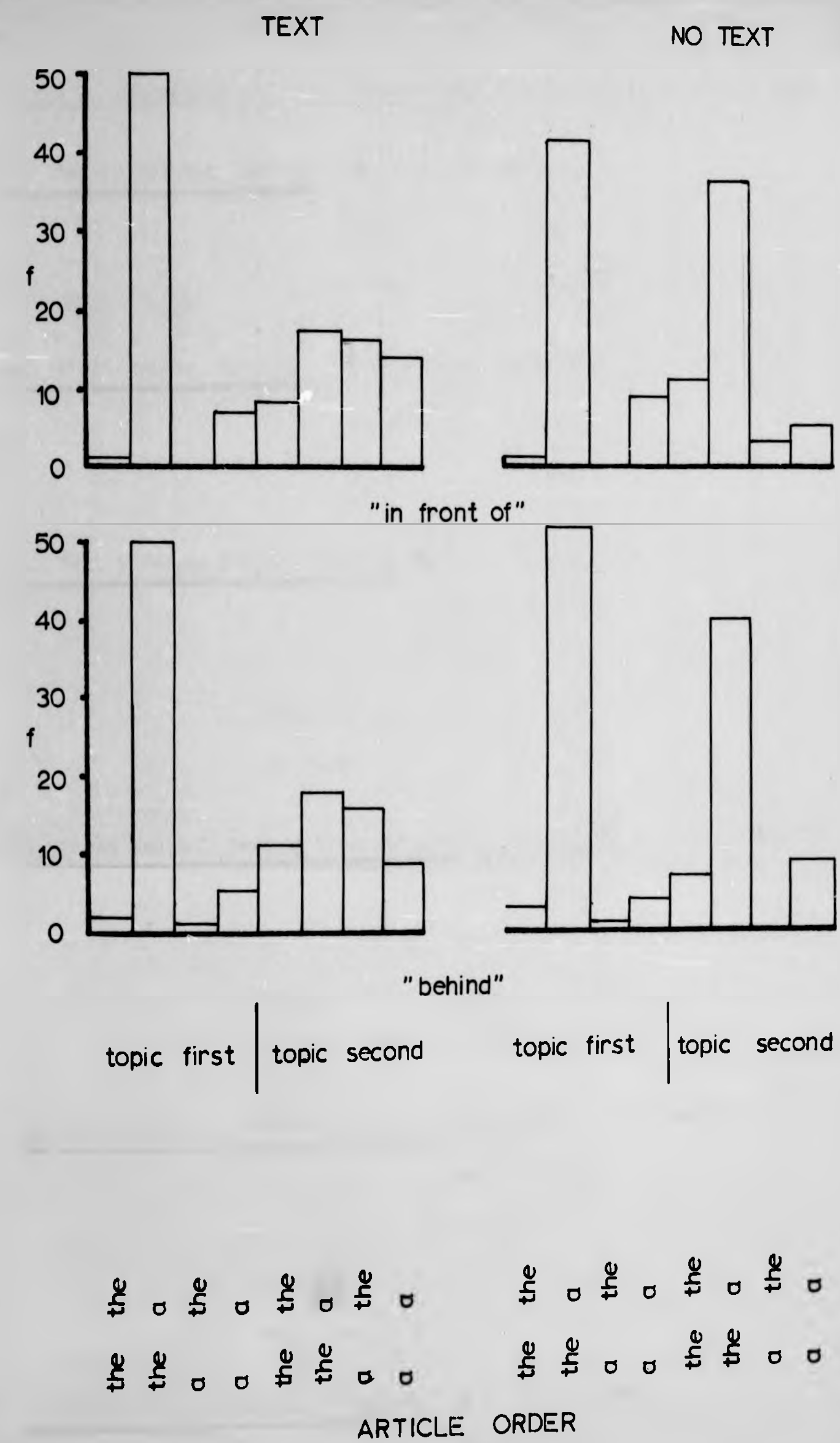


Figure 4 Relational term or topic inserted data, marked syntax

Table 5 Experiment 2 : Relational Term or Topic Inserted Data : Effects

1. Second Nominal Marking $\chi^2_1 = 17.30,\ p < 0.001$

"The"	"A"
389	514

2. First Nominal Marking $\chi^2_1 = 157.40,\ p < 0.001$

"The"	"A"
640	263

3. Text x Second Nominal Marking $\chi^2_1 = 25.92,\ p < 0.001$

	Second Nominal	
	"The"	"A"
Text	202	248
No Text	187	266

4. First Nominal Marking x Second Nominal Marking $\chi^2_1 = 26.61,\ p < 0.001$

		Second Nominal	
		"The"	"A"
First	"The"	250	390
Nominal	"A"	139	124

5. Second Nominal Marking x Topic Position $\chi^2_1 = 67.56,\ p < 0.001$

	Second Nominal	
	"The"	"A"
Topic First	136	322
Topic Second	253	192

6. Second Nominal Marking x Syntax $\chi^2_1 = 223.26,\ p < 0.001$

	Second Nominal	
	"The"	"A"
Unmarked Syntax	308	146
Marked Syntax	81	368

7. First Nominal Marking x Topic Position $\chi^2_1 = 61.15$, $p < 0.001$

	First Nominal "The"	First Nominal "A"
Topic First	382	76
Topic Second	258	187

8. First Nominal Marking x Syntax $\chi^2_1 = 17.30$, $p < 0.001$

	First Nominal "The"	First Nominal "A"
Unmarked Syntax	290	164
Marked Syntax	350	99

9. Text x First Nominal Marking x Second Nominal Marking $\chi^2_1 = 22.01$, $p < 0.001$

	Text: Second Nominal "The"	Text: Second Nominal "A"	No Text: Second Nominal "The"	No Text: Second Nominal "A"
First Nominal "The"	92	189	158	201
First Nominal "A"	110	59	29	65

10. Text x Second Nominal Marking x Topic Position $\chi^2_1 = 39.56$, $p < 0.001$

	Text: Second Nominal "The"	Text: Second Nominal "A"	No Text: Second Nominal "The"	No Text: Second Nominal "A"
Topic First	48	180	88	142
Topic Second	154	68	99	124

11. Text x Second Nominal Marking x Syntax $\chi^2_1 = 7.27$, $p < 0.01$

	Text: Second Nominal "The"	Text: Second Nominal "A"	No Text: Second Nominal "The"	No Text: Second Nominal "A"
Unmarked Syntax	147	78	161	68
Marked Syntax	55	170	26	198

12. Text x First Nominal Marking x Topic Position $\chi^2_1 = 42.98$, $p < 0.001$

Text	First Nominal "The"	First Nominal "A"	No Text First Nominal "The"	No Text First Nominal "A"
Topic First	196	32	186	44
Topic Second	85	137	173	50

13. First Nominal Marking x Second Nominal Marking x Syntax $\chi^2_1 = 112.68$, $p < 0.001$.

		Unmarked Syntax Second Nominal "The"	Unmarked Syntax Second Nominal "A"	Marked Syntax Second Nominal "The"	Marked Syntax Second Nominal "A"
First	"The"	206	84	44	306
Nominal	"A"	102	62	37	62

14. Text x First Nominal Marking x Second Nominal Marking x Syntax

$\chi^2_1 = 32.39$, $p < 0.001$

Text			Unmarked Syntax Second Nominal "The"	Unmarked Syntax Second Nominal "A"	Marked Syntax Second Nominal "The"	Marked Syntax Second Nominal "A"
	First	"The"	70	54	22	135
	Nominal	"A"	77	24	33	35
No Text	First	"The"	136	30	22	171
	Nominal	"A"	25	38	4	27

15. First Nominal Marking x Second Nominal Marking x Topic x Syntax

$\chi^2_1 = 27.31$, $p < 0.001$

			Unmarked Syntax Second Nominal "The"	Unmarked Syntax Second Nominal "A"	Marked Syntax Second Nominal "The"	Marked Syntax Second Nominal "A"
Topic First	First	"The"	115	66	7	194
	Nominal	"A"	12	37	2	25
Topic Second	First	"The"	91	18	37	112
	Nominal	"A"	90	25	35	37

Table 6 Experiment 2 : Results, Blank Sentence Frame Data.[a]

		"Behind"				"In front of"			
Second Nominal :		"The"		"A"		"The"		"A"	
First Nominal :		"The"	"A"	"The"	"A"	"The"	"A"	"The"	"A"
Text									
Unmarked Syntax	Topic First	7	0	13	0	3	0	11	4
	Topic Second	4	5	0	0	2	4	0	1
Marked Syntax	Topic First	1	0	20	2	0	0	13	0
	Topic Second	1	2	7	4	2	1	1	0
No Text									
Unmarked Syntax	Topic First	10	2	1	1	11	2	3	4
	Topic Second	6	0	1	3	4	2	4	4
Marked Syntax	Topic First	2	0	13	3	1	0	9	2
	Topic Second	2	1	11	1	0	0	10	0

a These data are illustrated in Figures 5 and 6.

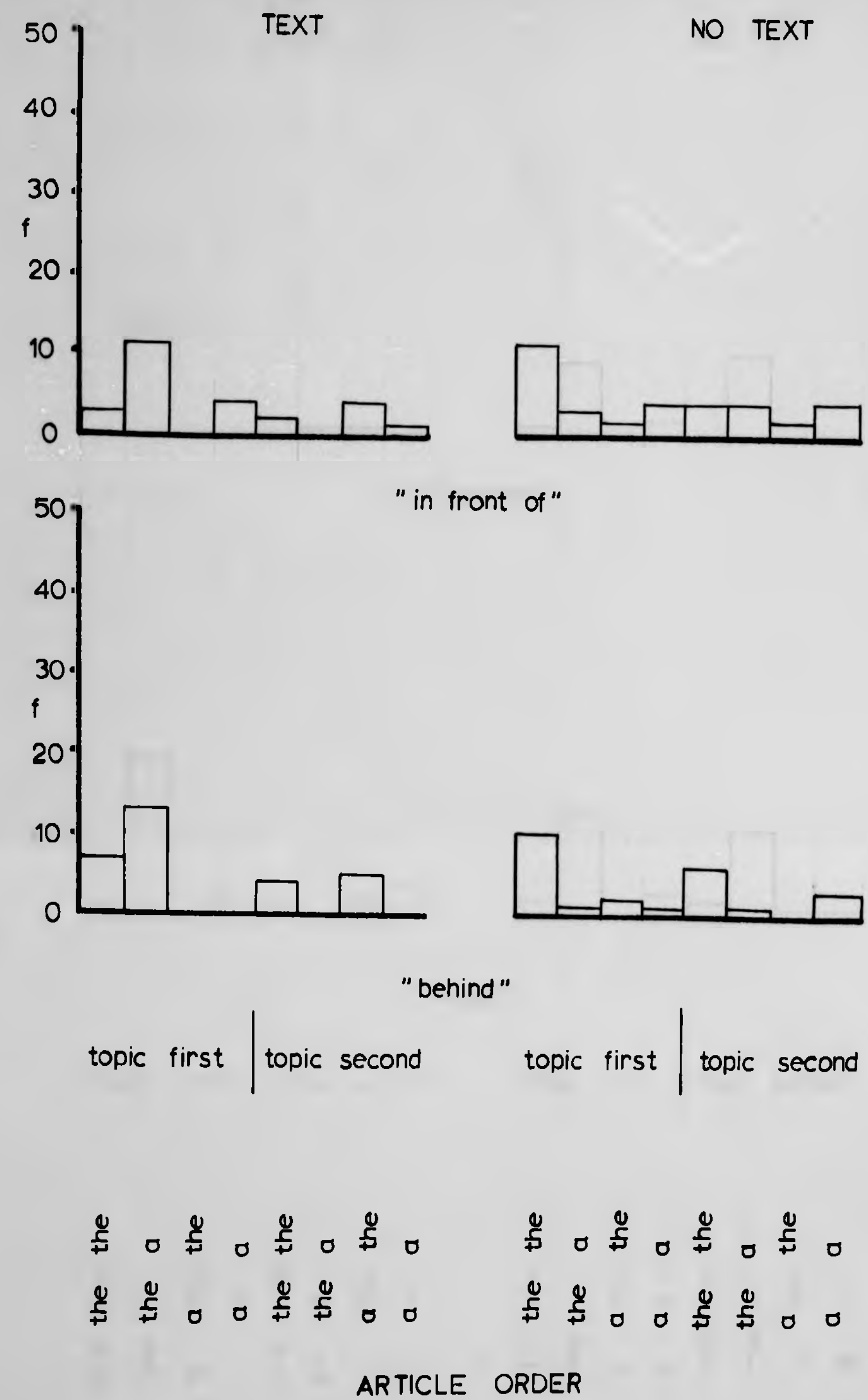


Figure 5 Blank frame data, unmarked syntax

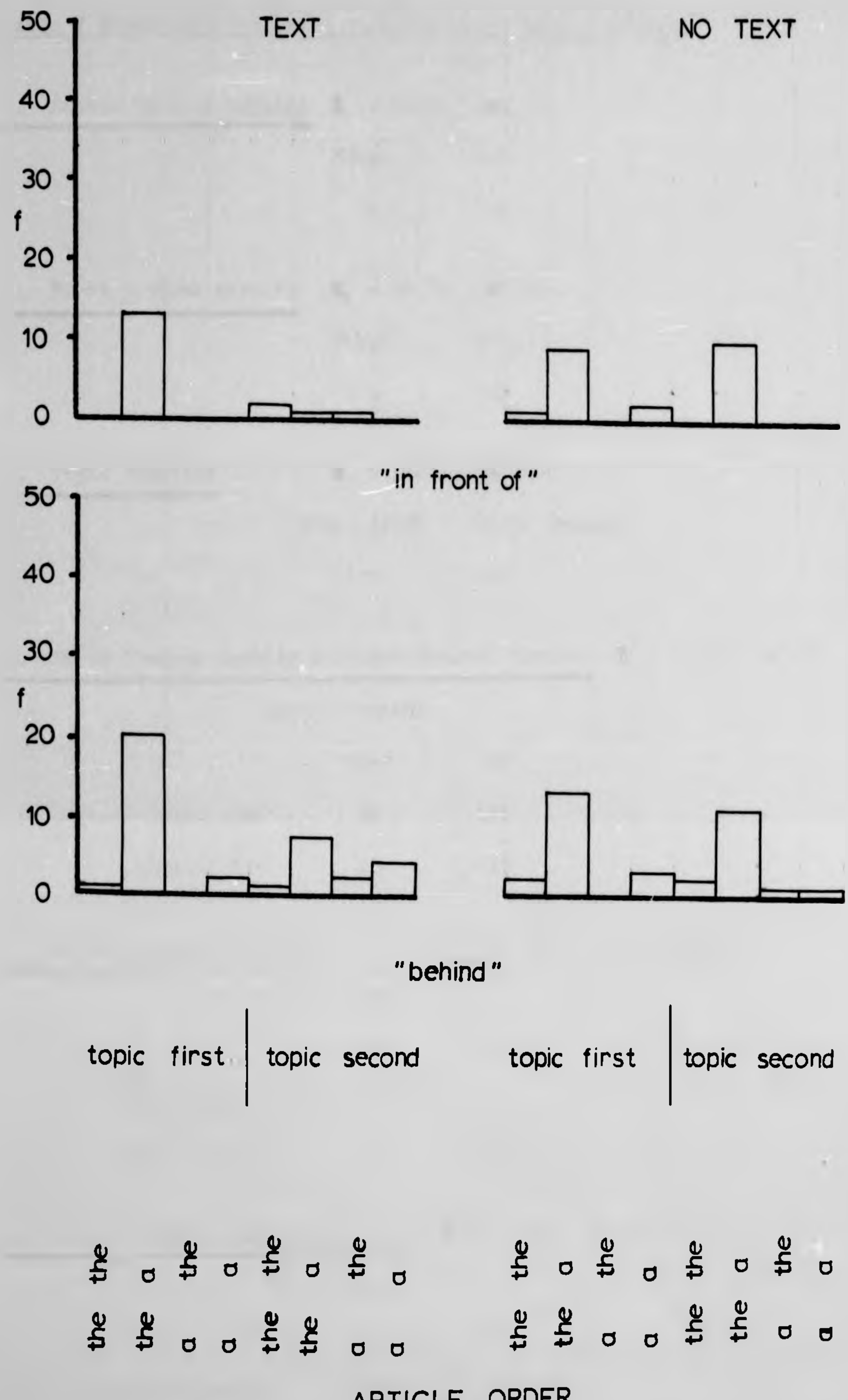


Figure 6 Blank frame data, marked syntax

Table 7 Experiment 2 : Blank Sentence Frame Data : Effects.

1. Second Nominal Marking $\chi^2_1 = 22.81$, $p < 0.001$

"The"	"A"
75	146

2. First Nominal Marking $\chi^2_1 = 70.70$, $p < 0.001$

"The"	"A"
173	48

3. Topic Position $\chi^2_1 = 13.68$, $p < 0.001$

Topic First	Topic Second.
138	83

4. First Nominal Marking x Second Nominal Marking $\chi^2_1 = 11.77$, $p < 0.01$

	Second Nominal	
	"The"	"A"
First "The"	56	117
Nominal "A"	19	29

5. Second Nominal Marking x Topic Position $\chi^2_1 = 10.37$, $p < 0.001$

	Second Nominal	
	"The"	"A"
Topic First	39	99
Topic Second	36	47

6. Second Nominal Marking x Syntax $\chi^2_1 = 40.64$, $p < 0.001$

	Second Nominal	
	"The"	"A"
Unmarked Syntax	62	50
Marked Syntax	13	96

7. First Nominal Marking x Topic Position $\chi^2_1 = 22.82,\ p < 0.001$

	First Nominal	
	"The"	"A"
Topic First	118	20
Topic Second	55	28

8. Relational Term x Syntax $\chi^2_1 = 6.19,\ p < 0.05$

	"Behind"	"In front of"
Unmarked Syntax	53	59
Marked Syntax	70	39

9. Text x Second Nominal Marking x First Nominal Marking $\chi^2_1 = 7.60,\ p < 0.01$

Text	Second Nominal		No Text	Second Nominal	
	"The"	"A"		"The"	"A"
First "The"	20	65		36	52
Nominal "A"	12	11		7	18

10. Text x Second Nominal Marking x Topic Position $\chi^2_1 = 22.30,\ p < 0.001$

Text	Second Nominal		No Text	Second Nominal	
	"The"	"A"		"The"	"A"
Topic First	11	63		28	36
Topic Second	21	13		15	34

11. Text x First Nominal Marking x Topic Position $\chi^2_1 = 12.69,\ p < 0.001$

Text	First Nominal		No Text	First Nominal	
	"The"	"A"		"The"	"A"
Topic First	68	6		50	14
Topic Second	17	17		38	11

12. Second Nominal Marking x First Nominal Marking x Syntax

$\chi^2_1 = 31.17,\ p < 0.001$

Unmarked Syntax	Second Nominal "The"	Second Nominal "A"	Marked Syntax	Second Nominal "The"	Second Nominal "A"
First "The"	47	33		9	84
Nominal "A"	15	17		4	12

13. Second Nominal Marking x Relational Term x Syntax $\chi^2_1 = 6.88,\ p < 0.001$

Unmarked Syntax	Second Nominal "The"	Second Nominal "A"	Marked Syntax	Second Nominal "The"	Second Nominal "A"
"Behind"	34	19		9	61
"In front of"	23	31		4	35

14. Text x First Nominal Marking x Second Nominal Marking x Topic Position

$\chi^2_1 = 10.01,\ p < 0.001$

Text	Topic First	Second Nominal "The"	Second Nominal "A"	Topic Second	Second Nominal "The"	Second Nominal "A"
	First "The"	11	57		9	8
	Nominal "A"	0	6		12	5
No Text						
	First "The"	24	26		12	26
	Nominal "A"	4	10		3	8

15. Text x First Nominal Marking x Second Nominal Marking x Syntax

$\chi^2_1 = 5.45,\ p < 0.05$

Text	Unmarked Syntax	Second Nominal "The"	Second Nominal "A"	Marked Syntax	Second Nominal "The"	Second Nominal "A"
	First "The"	16	24		4	41
	Nominal "A"	9	5		3	6
No Text						
	First "The"	31	9		5	43
	Nominal "A"	6	12		1	6

Discussion

Difficulties with the method of analysis must temper any conclusions one can draw from this study. As already noted the χ^2 method used is not the best such method and the independence assumption necessary to any χ^2 analysis is slightly dubious for the present data. However, to repeat, the method has here been used only really as a shorthand way of picking out dominant trends.

The most noticeable feature of the present results is undoubtedly the overwhelming importance of choices of definiteness marking. Almost all the major effects involve at least one of the definiteness factors. This is in sharp contrast to the previous experiment in which these seemed to be relatively peripheral. On the other hand it reinforces the Grieve (1974) conclusion on the importance of definiteness in production experiments. However it is clear from even a casual glance at the present data that more is involved in definiteness marking than simply a decision as to whether the referent of the nominal has been previously mentioned or not. All three data sets show noticeable main effects of both definiteness factors : the first nominal tending to be marked with "the" and the second marked with "a".

These trends are modified by several factors. Taking first the first nominal definiteness factor; there seem to be four influences on it:-

1. it is even more likely to be definite if the syntax is marked. This could be interpreted as a reflection of the functional nature of the syntactic choice : namely, like the passive, it is to keep the given information as theme. If this is its usual function then subjects could be performing some kind of frequency matching.

2. it is more likely to be definite if the second nominal is indefinitely marked. This may be simply due to interactions with both the topic and syntax factors - see below.

3. it is more likely to be definite if the topic is first. This relates in an obvious manner to our intuitions about definiteness marking and previous mention.

4. for some reason there appears to be a minor tendency towards an increased probability of having the first nominal definitely marked if the relational term is "behind". However this may well be due to the increased probability of having it definitely marked with marked syntax together with the greater likelihood of having "behind" with marked syntax. In any case this tendency is only of any size in the case where an article is inserted.

Second nominal definiteness is influenced by three main factors:

1. if there is marked syntax then it is even more likely to be indefinite.
2. it is more likely to be indefinite if the first nominal is definitely marked.
3. it is less likely to be indefinite if the topic is second.

The strong tendency to have the two nominals marked differently is itself influenced by the text factor : the trend being present only when text is presented. The tendency to have the topic nominal definitely marked and the other nominal indefinitely marked is, as one might expect, confined to the text case.

To summarise : there appears to be:-

1. a very strong tendency to mark the previously mentioned nominal with "the" and the other nominal with "a". This is of course confined to the text case.
2. there is a tendency with marked syntax to have the first nominal "the" and the second "a". This is present principally in the no text case and is largely overridden in the text condition by (1), though some residual effect remains.
3. there are a number of other smaller effects including (a) a marked trend towards placing the topic first where subjects are relatively free to choose. (b) a tendency, where an article is inserted, to have the first noun "the" when "behind" is the relational term. This may be due to (2) above and (3(c)) below.
(c) a tendency in some cases to have "behind" rather than "in front of"

with marked syntax.

These results are noticeably different from those of the last experiment in more respects than the greatly increased importance of definiteness in the present experiment. Not one of the results found in the comprehension data of the last experiment reached the criterion level used in the present experiment. In discussion of the last experiment it was accepted that there were five effects in need of explanation:-

1. the relational term x text interaction
2. the syntax x topic position x text interaction
3. the relational term x syntax x text interaction
4. the topic position x first nominal marking x second nominal marking x text interaction
5. the six way interaction.

Effects 3 and 5 were left unexplained.

Effect 1 was explained using the "good reason" interpretation of marking with topic decisions considered to dominate decisions about lexical marking choices. Effect 4 was explained as due to a secondary effect of this : namely a tendency for "behind" to go more easily with cases where the two nouns are differently marked. Effect 2 was said to be due to the possibility of a choice of theme : either the item to be added, or the old item, with marked syntax a way of keeping the old item theme.

The lack of an interaction between text and relational term in the present experiment is easy to explain given the above interpretation. In data set 2 the constraints are such as to prevent this. In set 1 the subject is forced to make definiteness choices because of the insertion of an article. In set 3 there is a mild trend in a similar direction to before. Though the fourth effect is not present in the current experiment related effects are. For example the first nominal is more likely to be "the" if the relational term is "behind", and the relational term is more likely to be "behind" if there is marked syntax. We have already seen that the nominals are more likely to be marked differently if syntax is marked. So "behind" is more frequent where nominals are differently marked. However in the present experiment this is not true for both types of syntax nor more so for the no text case than the text.[3] Again this may be partially due to the greater importance of decisions as to definiteness marking in the present experiment.

The lack of effect 2 is certainly most critical for the explanation given in the last chapter of that effect. However it is partially made up for by the presence of a main effect of topic position in the blank frame data : an effect which further demonstrates the importance of position in the sentence. The lack of an interaction with syntax may well be due to the different task demands of the two experiments. In the verification experiment the main point of interest in the target sentence is likely to be the new item since one has to add this to a "mental array". But in the present experiment this is not so : the subject has plenty of time to study the array and it seems possible that he is more likely to continue to focus on the topic. In that case we would expect it to remain theme for both types of syntax. This explanation is obviously rather weak. However the Clark/Huttenlocher explanation seems to be untenable for the present data since the new item is not necessarily the grammatical subject (nor the old item the reference point) - it depends on syntax, and yet there is clearly an effect of topic position.

Finally a brief comment on the greatly enhanced importance of definiteness marking in the present experiment. It is possible that the nature of the experiment encouraged subjects to think about definiteness marking since they could hardly avoid seeing that it was an experimenter-manipulated parameter. It is also possible that definiteness marking is a more salient parameter in production than in comprehension. This seems unlikely in general but it is probably true to say that it is not a very important feature in the verification task, in that it in no way affects

3. these results are what one would expect from translating shorter RT into greater frequency.

truth. It is conceivable that having subjects write sentences may also enhance the importance of the decision as to which article to use. This problem is to some extent overcome in the next experiment.

truth. It is conceivable that having subjects write sentences may also enhance the importance of the decision as to which article to use. This problem is to some extent overcome in the next experiment.

Experiment 3

Introduction

The last experiment is rather artificial for a number of obvious reasons : people are not normally restricted as to syntactic type, or definiteness of the first or second noun, or position of the noun referring to the previously mentioned object, or relational term. As noted the format of the experiment probably encourages people to think about the form of the sentence more than they would usually do. Moreover only the written mode is used. The method of creating a topic leads to the switch from a passive receptive mode to a more active expressive mode. This must interrupt the natural flow of discourse and disrupt choices relating to discourse cohesion. Furthermore the restriction to the use of definiteness marking as a method of distinguishing new from old information is probably unrealistic : in normal discourse pronouns can, and probably are, used to achieve that effect.

The current experiment tries to avoid these problems by giving subjects simple tasks which are obviously communicative while avoiding, as far as possible, direct stress on the form of the communication. Two of the tasks emphasise the description of the objects in the picture by asking subjects to describe the picture so someone could sketch it. Two other tasks emphasise the nature of the relations amongst objects in the picture by stressing the importance of somebody being able to construct a schema of the picture. One task of each type is designed so that subjects have to start by indicating the object in the middle of the picture. Since the pictures are the same as in the previous experiment this means that the second object mentioned will be in front of the first object half the time and behind it half the time. The use of two tasks in which the starting point is not specified is to examine an hypothesis of De Soto, London and Handel (1965) that subjects prefer to construct displays from front to back.

The importance of this hypothesis is that it provides a possible explanation of the interaction between the text and topic factors in the previous experiment - an explanation not related to the linguistic representation directly. This explanation is as follows. Assume that order of mention of the two objects corresponds to direction of construction. Now in the case where there is no topic unmarked syntax + "in front of" ensures that order of mention and order in the display is the same. This is not so for marked syntax + "in front of" or unmarked syntax + "behind". However it is so for marked syntax + "behind", but this will be used less than unmarked syntax + "in front of" because only a topicalisation reason would justify marked syntax. The situation is quite different when one object has already been mentioned. "Ex hypothesi" if subjects are unconstrained they will mention the first object first so that when the relational sentence comes it will not really matter what the order of mention is. There is the option of saying either that the previously mentioned object is in front of the other object, or of saying that the other object is behind the previously mentioned object. Of course this explanation is not incompatible with the linguistic explanation I have so far accepted : both factors could be operating. But if there is no evidence that people do have such preferences in building displays then the linguistic explanation stands unopposed. One of the purposes of the present experiment is to test this.

1. Subjects

32 first year psychology students fulfilling a requirement for the Introductory Psychology course.

2. Apparatus and Materials.

32 slides most of which depicted model animals (manufacturers Britains Ltd.), though some were also of people. Each slide depicted two animals/people on a wooden board with a completely white background, viewed side on, both facing in the same direction (i.e. to the viewers left or to his right). One animal was always in the middle of the picture, the other to the left or right. Since they were both at the same depth this meant that one object was behind the other, in the "bus queue" (rather than the "depth") sense of "behind" (on this see footnote 8, Chapter 2). Half of the pictures had the element in the centre "at the front of the queue", half the peripheral element. Half of each of these faced left and half faced right. Presentation order was randomised but was the same for all subjects. It was intended that the picture should be easily divisible into three sections : the centre slot was always filled by an animal as was one of the outside slots, the remaining one being blank. Subjects did not always see the picture as so divided, however.

All pictures were taken from a distance of approximately two feet using an Asahi Pentax S1a single lens reflex camera, in daylight. Slides were back projected onto a screen of the same size and at the same distance from the subject as in Experiment 2. They sat at a similar table with the slide projector control and microphone. Their descriptions were recorded using a Revox A700 tape deck. Recordings were of very high quality and little difficulty was encountered in transcription.

3. Procedure

Subjects were sat at a distance of approximately four feet from the screen on which the slides were projected, with a remote slide changer/

focussing device to their left. They were shown how the slide changer worked and an example picture. Further instructions were dependent on condition as follows (the example slide was displayed throughout the instructions) :

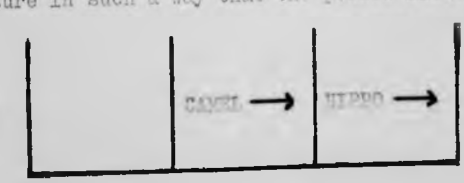
Condition A (First 8 subjects). Subjects were told that they would see a series of pictures and that they were to describe them so that somebody could make a reasonable sketch of the picture using only their description. They were to decide for themselves what level of detail would be required for a sketch to be reasonably said to be of the picture. When they had satisfactorily described the picture they were to use the slide changer to bring on the next picture, there being one blank slide after each picture. Altogether there were 32 pictures and it was stressed these must be completed in one hour at most though they might take less, even considerably less than this.

Finally subjects were told that there was one restriction : they must start their description by describing the object in the centre of the picture.

Condition B : Subjects were given the same instruction as A except that the final instruction restricting their starting point was omitted.

Condition C : Subjects were told that they were going to see 32 pictures which they were to describe so that someone listening to the recordings of their description would have a good idea of what was in the picture. This was to be a fairly abstract idea, not involving minutiae. For example one might describe the example picture in such a way that the person listening could produce the following:

arrows denote direction in which the animal faces.



(This diagram was drawn on the board)

The experimenter always described this, with appropriate pointing using indexical expressions (Morris, 1945) e.g. "The hippo is here facing this way, the camel here facing this way, and there is nothing here". The experimenter did not always start at the front in this way though : he tried to vary his approach as much as possible in order not to unduly influence the subject. The indexical expressions were similarly used in order not to suggest a linguistic strategy.

Subjects were again told that they had a maximum of one hour in which to complete 32 picture-descriptions, but they would probably take less time than this.

Finally they were told that they must start their description using the following from "In the middle of this picture is a -----". This had two purposes : (i) to make them concentrate on the middle item so that they would have to go backwards in the display half the time, and forwards half the time and (ii) to stress the importance of the relation involved (by thematic foregrounding of this element).

Condition D : instructions were as for condition C except that the final constraint on starting point was omitted.

Results [4]

The tape recordings of subjects' descriptions were scored as follows:-

1. the order of mention of the two objects was noted in order to test the hypothesis outlined in the introduction.

2. each sentence containing either "in front of" or "behind" was scored in terms of the following criteria

(a) whether the first nominal was a pronoun, a definite noun phrase or an indefinite noun phrase.

(b)similarly for the second nominal.

(c) whether marked or unmarked syntax was used. In some cases only one noun phrase is present e.g. "In front is a lion", "a tiger is behind". If the relational term comes before the noun phrase this is scored as marked syntax, if not as unmarked.

(d) whether the relational term is "behind" or "in front of"

(e) the position of the noun referring to the previously mentioned object (1st., 2nd., both or neither).

(f) which noun phrase refers to the object referred to in the subsequent sentence (1st., 2nd., both or neither if either there is no subsequent sentence or there is but neither object is mentioned).

The reasons for scoring on this last criterion will become clearer in the discussion.

A good many descriptions did not contain either "behind" or "in front of" (see Footnote 4). In addition there were several subjects who did not pace themselves to complete the task within the hour so these did not complete 32 descriptions. (See Footnote 4). Finally one subject failed to follow instructions in condition A and did not start from the designated place.

Tables 1 - 4 give the sentences which contained "behind" and "in front of" classified according to the above criteria, for Conditions A - D respectively. Fig. 7 graphically represents the number of people producing

4 A note on Egocentricity. The present experiment is similar in some respects

A note on Egocentricity (continued)

to the kind of communication problems studied by Krauss and Glucksberg (1969). One can think of the inability to describe pictures in such a manner as to enable someone to reconstruct it as partially due to inability to adopt the viewpoint of the other.

Even on a very weak criterion of success subjects in the present experiment did not do too well. A minimal requirement is that the description should be such as to enable one to:

(a) know what two objects are depicted

(b) know the direction in which they face

(c) know their relative positions.

Even these three criteria have a lot of hidden assumptions e.g. that only two objects are depicted, (perhaps) that they are seen in profile and (perhaps) that they both face in the same direction.

Subjects always indicate which two animals (or whatever else it is) are depicted. But they do not always indicate either the direction faced or relative position. If one assumes giving the direction which at least one of the animals faces and giving their relative position is enough then one gets the following table:-

Condition A	Neither Given	Only Direction Given	Only Relation Given	Both Given	Missing
S1	3	2	22	5	
S2			1	30	1
S3			2	28	2
S4			20		12
S5	failed to follow instructions				
S6		23	9		
S7	17	6	9		
S8	5		26	1	
Condition B					
S1			3	29	
S2		1	4	24	3
S3			4	28	
S4	32				
S5	3	9	2	18	

A note on Egocentricity (continued)

Condition B (Contd.)	Neither Given	Direction Only Given	Relation Only Given	Both Given	Missing
S6		1		17	14
S7				32	
S8	1			28	3
Condition C					
S1				32	
S2			6	26	
S3		2	3	27	
S4				32	
S5	1			31	
S6		1		31	
S7			31	1	
S8				32	
Condition D					
S1		31		1	
S2				32	
S3	4	5	5	18	
S4				32	
S5		1	[illegible]	25	
S6		[illegible]	[illegible]	26	
S7			2	30	
S8		1		31	

One subject (4 in Condition B) failed to give either the direction faced or the relative positions of the objects for any slide. Subject 6 in Condition A made syntactic errors for the 23 cases of "direction only given". In fact she intended to give direction and succeeded in giving position since she said "to its left" rather than "to the left of it".

Note that the criterion used here assumes the correct interpretation of ambiguous phrases like "to the left/right of it" and "behind/in front of it" will be given. This may not be true though.

Table 8 Experiment 3 : Results, Condition A (Sketch oriented description starting with central object).

SENTENCE TYPE

		Noun 1			Noun 2			Syntax		Relational Term		Prev.	Subs.
Subject	f	Pro	"The"	"A"	Pro	"The"	"A"	Mar.	Unm.	"Beh."	"Front"	Ment.	Ment.
1	9	x					x	x		x		1st.	2nd.
	3	x					x	x		x		1st.	Neither
	2		x				x	x		x		1st.	2nd.
	7	x					x	x			x	1st.	2nd.
2		NO USE OF "BEHIND" OR "IN FRONT OF"											
3	1	x					x	x		x		1st.	2nd.
	2	x					x	x			x	1st.	2nd.
4	2	x					x	x			x	1st.	2nd.
	11	—	—				x	x			x	-	2nd.
	5	—	—				x	x		x		-	2nd.
	1			x	—	—			x	x		-	1st.
5		FAILED TO FOLLOW INSTRUCTIONS											
6	6	x					x	x		x		1st.	2nd.
	3	x					x	x			x	1st.	2nd.
7	1	x					x	x		x		1st.	2nd.
8	1	x					x	x		x		1st.	2nd.

sentences of each type, omitting analysis of previous and subsequent mention. Note the great variation in Condition B, and to a lesser extent Condition D, with rather less variation in Conditions A and C. There are in fact 36 possible sentence types on the four criteria of NP1 marking, NP2 marking, relational term and syntax. Only 23 occur though : 4 in Condition A, 19 in Condition B, 9 in Condition C and 10 in Condition D. Responses from Conditions A and C tend to cluster together, as do those from Conditions B and D.

Below is a more detailed description of the data, condition by condition.

Condition A (sketch oriented description starting with central object) (see Table 8)

Six subjects gave responses using "behind" or "in front of". Of these two gave only one response of this type. Nevertheless it is clear even from the few cases remaining what is the dominant response. This is typically a case of marked syntax with the first nominal a pronoun, the second nominal a noun with the article "a", either relational term (perhaps a slight bias towards "behind"), the first noun phrase referring to an object previously referred to, the second referring to an object not previously referred to. Usually the succeeding sentence is about the newly introduced object.

There are very few exceptions to this typical example apart from another distinct type - see below. Only in two cases is the first noun phrase not a pronoun (both of these it is a noun + "the"). There are no exceptions to the second noun being "a" + noun; none to the first nominal referring to something previously mentioned, and the second to something new; none to the use of marked syntax; and only 3 responses (from one subject) in which the subsequent sentence did not talk about the object referred to by the second noun phrase (in those cases neither object was mentioned in the subsequent sentence). With regard to the relational term there were 23 cases with "behind" (from 5 subjects) and 14 cases with "in front of" (from 4 subjects).

There is a small set of responses not covered by these generalisations :

that is cases in which only one noun is mentioned. These sentences closely resemble the full relational sentence in many respects - examples are "In front is a pig", "A horse is behind" - and can be easily seen to correspond to full relational sentences except for the absence of one of the noun phrases. All these responses were from one subject.

There are two types : one corresponding to the typical full relational sentence described above except that the first noun phrase is missing (so only the new object is mentioned and the succeeding sentence is about this).

Eleven of these had "In front", 5 "behind". The other case was one of unmarked syntax, with the first (and only) noun phrase having "a" + noun, "behind" as the relational term and the subsequent sentence about the object mentioned here.

With regard to noun phrase definiteness marking, this is as follows:

<u>Previously mentioned object.</u>

	Pro.	The	A
One object only previously mentioned	35	2	0
Both objects previously mentioned	0	0	0

<u>Object Not Previously Mentioned.</u>

One object only previously mentioned	0	0	37
Neither object previously mentioned	0	0	17

The data from the subsequent sentence criterion are as follows for the two syntactic types.

	Mentioned in the Subsequent Sentence			
	Neither	First Noun Referent	2nd.N.Ref.	Both
Marked Syntax	3	0	50	0
Unmarked Syntax	0	1	0	0

None of the other conditions show anything like the uniformity of Condition A.

<u>Condition B (sketch oriented description, starting point unspecified)</u>

(see Table 9)

It is apparent that much of the variation in this condition comes from cases in which either both objects are mentioned prior to the relational sentence, or in which neither object is mentioned prior to the relational sentence. The cases where one or the other object was previously mentioned are much more uniform. There are 56 cases of the same sort as the typical type for Condition A, out of a total of 73 cases in which only one object was previously mentioned. Forty five of these have "behind" as the relational term, only eleven "in front of". There are 12 cases of unmarked syntax with only one object previously mentioned. Of these 9 have the previously mentioned object referred to by the second noun phrase. The remaining 5 cases are cases with marked syntax which differ from the typical case in terms of what is mentioned in the subsequent sentence or in definiteness marking. It is noticeable that in all these data the only noun phrase marking order observed with marked syntax are

Pro - the

Pro - a

The - a.

In the cases where marked syntax is used when both objects have been previously mentioned the immediately previously mentioned object is always referred to by the first noun. (It is not possible to pick these cases out from the table). In fact there are 15 cases altogether in which both objects are previously mentioned and marked syntax is used - there are 21 such cases with unmarked syntax. There are 21 cases where neither object has been previously mentioned (including 3 1 - noun phrase sentences) and all of these use unmarked syntax, 9 with "behind" and 12 with "in front of". This complete absence of marked syntax where neither object has been previously mentioned is consistent with the analysis given of that option. However one would have expected a greater predominance of sentences with "in front of".

With regard to definiteness marking of the previously mentioned item,

Table 9 Experiment 3 : Results, Condition B (Sketch oriented description, starting point unspecified.)

		Sentence Type											
		Noun 1			Noun 2			Syntax		Relational Term		Prev.	Subs.
Subject	f	Pro	"The"	"A"	Pro	"The"	"A"	Mar.	Unm.	"Beh."	"Front"	Ment.	Ment.
1	13	x					x	x		x		1st.	2nd.
	8	x					x	x			x	1st.	2nd.
	1	x					x	x		x		1st.	Neither
	1			x	x				x	x		2nd.	Neither
	1			x	x				x	x		2nd.	1st.
	1			x	x				x		x	2nd.	1st.
2	1		x			x			x	x		Both	2nd.
	1		x			x			x	x		Both	1st.
	5		x			x			x		x	Both	Both
	2	x					x	x		x		Both	2nd.
	3		x			x			x		x	Both	Neither
	1		x			x			x		x	Both	1st.
	2		x			x			x		x	Both	2nd.
	1		x		x				x	x		Both	1st.
	1	x					x	x			x	Both	2nd.
	4	x				x		x		x		Both	2nd.
	2		x				x	x		x		Both	2nd.
	2		x			x		x		x		Both	2nd.
	1	x					x	x		x		1st.	2nd.
	1			x	x				x	x		2nd.	1st.
	1			x	x				x	x		2nd.	Neither
	1		x				x		x	x		1st.	Both
	2			x			x		x	x		Neither	1st.
	4			x			x		x	x		Neither	Neith.
	1			x			x		x	x		Neither	Both
	2			x			x		x		x	Neither	Both
	1			x			x		x		x	Neither	Neith.
	2			x			x		x		x	Neither	2nd.
3	1	x					x	x		x		1st.	Both

Table 9 (Continued)

Sentence Type

		Noun 1			Noun 2			Syntax		Relational Term		Prev.	Subs.
Subject	f	Pro	"The"	"A"	Pro	"The"	"A"	Cr.	Unc.	"Beh."	"Front"	Sent.	Sent.
	9	x					x	x		x		1st.	2nd.
	3	x					x	x			x	1st.	2nd.
	1		x				x	x			x	1st.	2nd.
	1			x		x			x	x		2nd.	1st.
	1			x					x		x	-	1st.
	1		x			x			x	x		Both	Neither
	2		x		x				x	x		Both	2nd.
	1		x			x			x		x	Both	Neither
	1	x				x		x			x	Both	2nd.
4		NO USE OF "BEHIND" OR "IN FRONT OF"											
5	1	x				x			x		x	Both	1st.
	1	x				x			x		x	Both	2nd.
	2	x					x	x		x		Both	2nd.
	1	x				x		x		x		Both	Neither
	1		x			x			x	x		Both	Neither
	1		x				x		x		x	1st.	2nd.
	2	x					x	x		x		1st.	2nd.
	1	x					x		x	x		1st.	Both
	1		x			x		x		x		1st.	2nd.
6	1			x	x				x	x		2nd.	? Both
	2			x					x	x		-	? Both
	1			x					x		x	1---	Both
7	3	x					x	x		x		1st.	2nd.
	1		x				x	x		x		1st.	2nd.
8	5			x			x		x		x	-	Both
	17	x					x	x		x		1st.	2nd.
	1		x		x				x	x		2nd.	Neither
	1			x	x				x	x		2nd.	1st.

this is as follows:-

	Pro	The	A
One object only previously mentioned	67	6	-
Both objects previously mentioned	16	48	7

The object not previously mentioned is marked as follows:

	Pro	the	a
One object only previously mentioned	-	4	69
Neither object previously mentioned	-	12	28

So it appears that previously mentioned objects are occasionally marked with "a" but only when both objects were previously mentioned. This never happens when only one was previously mentioned. On the other hand objects not previously mentioned are occasionally referred to with "the".

There are quite clear differences between the two syntactic types in terms of what people go on to talk of in the subsequent sentence.

	Neither	First Noun Referent	Second Noun Referent	Both
Marked Syntax	2	0	68	1
Unmarked Syntax	14	12	13	19

The very limited freedom of occurrence of the marked syntactic type (tends to occur only when one object has been previously mentioned and is followed almost entirely by sentences in which only the object referred to by the second noun are mentioned) is consistent with this being a marked item. But this does not prevent it from being the most frequently used type in the context of the present experimental condition.

Condition C (schematic description starting with central object) (see Table 10)

This condition produced a slightly more restricted set of sentence/context type than Condition B.

Of a total of 199 responses using "behind" or "in front of" 101 were of the 'typical' class of Condition A. A further 31 differed from it only in terms of the subsequent sentence criterion.

Table 10 Experiment 3 : Results. Condition C (Schematic Description, starting with central object).

Sentence Type

Subject	f	Noun 1 Pro	Noun 1 "The"	Noun 1 "A"	Noun 2 Pro	Noun 2 "The"	Noun 2 "A"	Syntax Mar.	Syntax Unm.	Relational Term "Beh.	Relational Term "Front"	Prev. Ment.	Subs. Ment.
1	19	x					x	x		x		1st.	2nd.
	4	x					x	x		x		1st.	Both
	2	x					x	x			x	1st.	Neither
	4	x					x	x			x	1st.	Both
	9	x					x	x			x	1st.	2nd.
	1			x	x				x		x	2nd.	1st.
2	23		x				x	x		x		1st.	Neither
	4			x		x			x	x		2nd.	Neither
	4	x					x	x		x		1st.	Neither
	10	x					x	x			x	1st.	Neither
	24		x				x	x			x	1st.	Neither
	1		x		x			x			x	2nd.	Neither
	4			x		x			x		x	2nd.	Neither
	2		x		x				x		x	2nd.	Neither
3	2	x					x	x		x		1st.	2nd.
4	15	x					x	x		x		1st.	2nd.
	17	x					x	x			x	1st.	2nd.
5	6	x					x	x		x		1st.	Neither
	5	x					x	x		x		1st.	2nd.
	1	x					x	x		x		1st.	1st.
	3	x					x	x		x		1st.	Both
	6	x					x	x			x	1st.	2nd.
	1	x					x	x			x	1st.	1st.
	2	x					x	x			x	1st.	Both
6		NO USE OF "BEHIND" OR "IN FRONT OF"											
7	1	x					x	x		x		1st.	2nd.
8	18	x					x	x			x	1st.	2nd.
	11	x					x	x		x		1st.	2nd.

A further 47 (all from one subject)[5] differed from it in having a noun + "the" as the first noun phrase, as well as on the subsequent sentence criterion. In fact altogether there were only 11 responses with unmarked syntax, 7 of these with "in front of" and 4 with "behind". The 188 marked syntax responses consisted of 92 with "behind" and 96 with "in front of". Looking at definiteness marking of the previously mentioned item produces the following results.

	Pro	the	a
One object only previously mentioned	146	63	-

The item not previously mentioned is marked as follows:

	Pro	the	a
One object only previously mentioned	-	4	195

As in the other conditions marked syntax is almost invariably followed by a sentence in which the object referred to by the second noun phrase of the relational sentence is mentioned:-

	Object Referred to in Subsequent Sentence			
	Neither	1st. Noun Referent	2nd. Noun Referent	Both
Marked Syntax	8[a]	0	105	13
Unmarked Syntax	0[b]	1	0	0

a Excludes 62 not really classifiable.

b Excludes 10 not really classifiable.

5 All these sentences are from Subject 2 who talked a lot of the time about shadows etc. in the picture. Hence the very large number of responses. The subsequent sentence criterion proved very hard to apply to this subject (mainly because of talk of "its shadow", "the shadow" etc.) and so are arbitrarily classed as neither.

A further 47 (all from one subject)[5] differed from it in having a noun + "the" as the first noun phrase, as well as on the subsequent sentence criterion. In fact altogether there were only 11 responses with unmarked syntax, 7 of these with "in front of" and 4 with "behind". The 188 marked syntax responses consisted of 92 with "behind" and 96 with "in front of". Looking at definiteness marking of the previously mentioned item produces the following results.

	Pro	the	a
One object only previously mentioned	146	63	-

The item not previously mentioned is marked as follows:

	Pro	the	a
One object only previously mentioned	-	4	195

As in the other conditions marked syntax is almost invariably followed by a sentence in which the object referred to by the second noun phrase of the relational sentence is mentioned:-

	Object Referred to in Subsequent Sentence			
	Neither	1st. Noun Referent	2nd. Noun Referent	Both
Marked Syntax	8[a]	0	105	13
Unmarked Syntax	0[b]	1	0	0

a Excludes 62 not really classifiable.

b Excludes 10 not really classifiable.

5 All these sentences are from Subject 2 who talked a lot of the time about shadows etc. in the picture. Hence the very large number of responses. The subsequent sentence criterion proved very hard to apply to this subject (mainly because of talk of "its shadow", "the shadow" etc.) and so are arbitrarily classed as neither.

Condition D (schematic description, starting point unspecified)

(See Table 11 which presents the full set of data from this condition).

This condition followed condition B in the relatively low frequency of marked syntax. Altogether 57 responses out of 106 used marked syntax, of these 30 (all but one from one subject) were of the "typical" class of Condition A. A further one case differed from that only in terms of the subsequent sentence criterion. All the remainder had "the" + noun as the first noun phrase instead of a pronoun. All cases of marked syntax have the previously mentioned object referred to first, and only one object previously mentioned.

This last fact is by no means true of cases of unmarked syntax. In 38 cases both objects were mentioned prior to the relational sentence, in 2 cases neither was mentioned, in 9 cases the object mentioned second in the relational sentence was the only object previously mentioned, and in no cases was the object mentioned first the only object previously mentioned.

Marking of the object previously mentioned was as follows:-

	Pro	the	a
One object only previously mentioned	35	31	-
Both objects previously mentioned	11	65	-

Marking of the object not previously mentioned was as follows:-

	Pro	the	a
One object only previously mentioned	-	-	66
Neither object previously mentioned	-	-	4

These results are very clear : the object previously mentioned is never marked with "a" and the object not previously mentioned is never marked with "the" or "replaced" by a pronoun. There appears to be a much greater tendency to refer to the objects with "the" + noun than with a pronoun in the case where both have been previously referred to.

The subsequent mention criterion shows a pattern very similar to condition B.

Table 11 Experiment 3 : Results, Condition D (Schematic Description, starting point unspecified).

Sentence Type

		Noun 1			Noun 2			Syntax		Relational Term		Prev.	Subs.
Subject	f	Pro	"The"	"A"	Pro	"The"	"A"	Var.	Unm.	"Beh."	"Front"	Ment.	Ment.
1					NO USE OF "BEHIND" OR "IN FRONT OF"								
2					NO USE OF "BEHIND" OR "IN FRONT OF"								
3	1			x	x				x	x		2nd.	Neither
	1			x	x				x	x		2nd.	1st.
	1			x	x				x	x		2nd.	2nd.
	1			x		x			x	x		2nd.	2nd.
4					NO USE OF "BEHIND" OR "IN FRONT OF"								
5	4		x			x			x	x		Both	Neither
	2		x			x			x	x		Both	Both
	1		x			x			x	x		Both	1st.
	1		x			x			x	x		Both	2nd.
	1			x			x		x	x		-	Both
	1	x					x	x		x		1st.	Neither
	1	x				x			x	x		Both	Both
	1		x		x				x	x		Both	Both
	1		x			x			x		x	Both	2nd.
	1		x		x				x		x	Both	Neither
6	4		x			x			x	x		Both	2nd.
	5		x			x			x	x		Both	Both
	1		x			x			x	x		Both	1st.
	3		x			x			x	x		Both	Neither
	3		x		x				x	x		Both	1st.
	1		x		x				x	x		Both	Neither
	1		x		x				x	x		Both	Both
	1	x				x			x	x		Both	Both
	1	x					x	x		x		1st.	2nd.
	1			x			x		x	x		Neither	2nd.
	1		x			x			x		x	Both	Both
	1		x			x			x		x	Both	Neither
	1		x		x				x		x	Both	Both
7	16		x				x	x		x		1st.	Both
	7		x				x	x		x		1st.	2nd.

Table 11 (continued)

Sentence Type

		Noun 1			Noun 2			Syntax		Relational Term		Prev.	Subs.
Subject	f	Pro	"The"	"A"	Pro	"The"	"A"	Mar.	Unm.	"Beh."	"Front"	Ment.	Ment.
7 (contd.)													
	2		x				x	x		x		1st.	1st.
	1		x				x	x		x		1st.	Neither
	2			x		x			x	x		2nd.	1st.
	1			x		x			x	x		2nd.	2nd.
	2		x		x				x	x		Both	1st.
	1		x			x			x	x		Both	2nd.
	1		x			x			x	x		Both	Neither
8	1			x	x				x	x		2nd.	1st.
	29	x					x	x		x		1st.	2nd.
	1			x		x			x	x		2nd.	Both

Object Referred to in Subsequent Sentence

	Neither	1st. Noun Referent	2nd. Noun Referent	Both
Marked Syntax	2	2	37	16
Unmarked Syntax	12	11	11	15

Finally what of the hypothesis suggested initially that the lexical marking results are due to a preferred homomorphism between order of mention and order in the display? The simplest way to test this hypothesis is to consider those cases in Conditions B and D (where subjects were not given a starting point) where "behind" or "in front of" are not used. This eliminates any linguistic bias so we can examine any other bias by simply looking at order of mention. This gives the following table.

Condition B

	Front Object Mentioned First	Rear Object Mentioned First	
S1	0	5	-
S2	2	0	+
S3	14	0	+
S4	15	17	-
S5	10	9	+
S6	13	1	+
S7	15	13	+
S8	5	0	+
	$\bar{x}$ 9.25	$\bar{x}$ 5.62	

Condition D

S1	20	12	+
S2	22	10	+
S3	23	5	+
S4	32	0	+
S5	8	10	-
S6	8	1	+
S7	0	0	
S8	1	0	+
	$\bar{x}$ 16.28	$\bar{x}$ 5.43	

The probability of only 3 out of 15 beginning with the second object is quite small - in fact p = 0.036. So there does appear to be a bias.

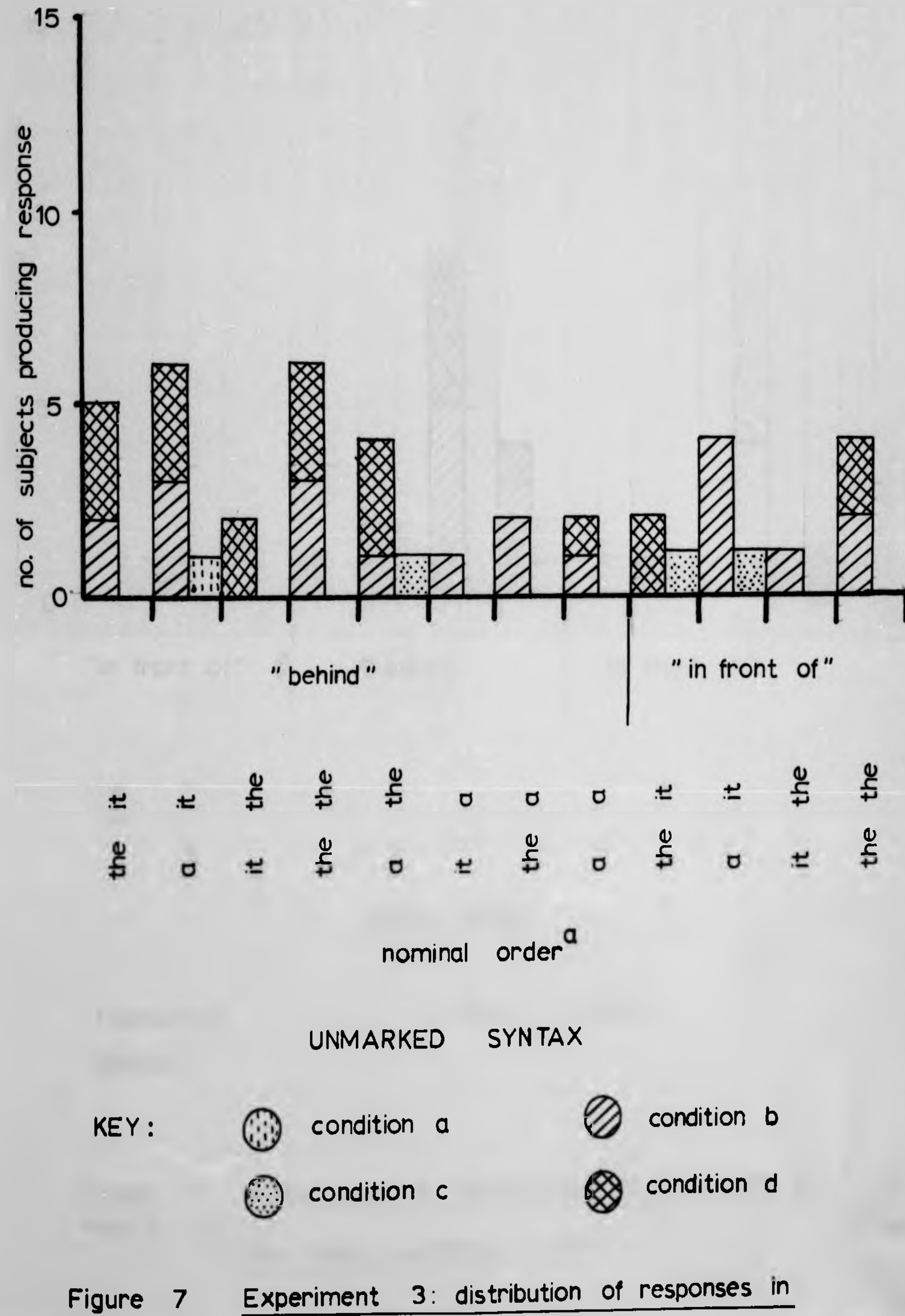


Figure 7 Part 1 Experiment 3: distribution of responses in the four conditions

[a] "it" denotes any pronoun

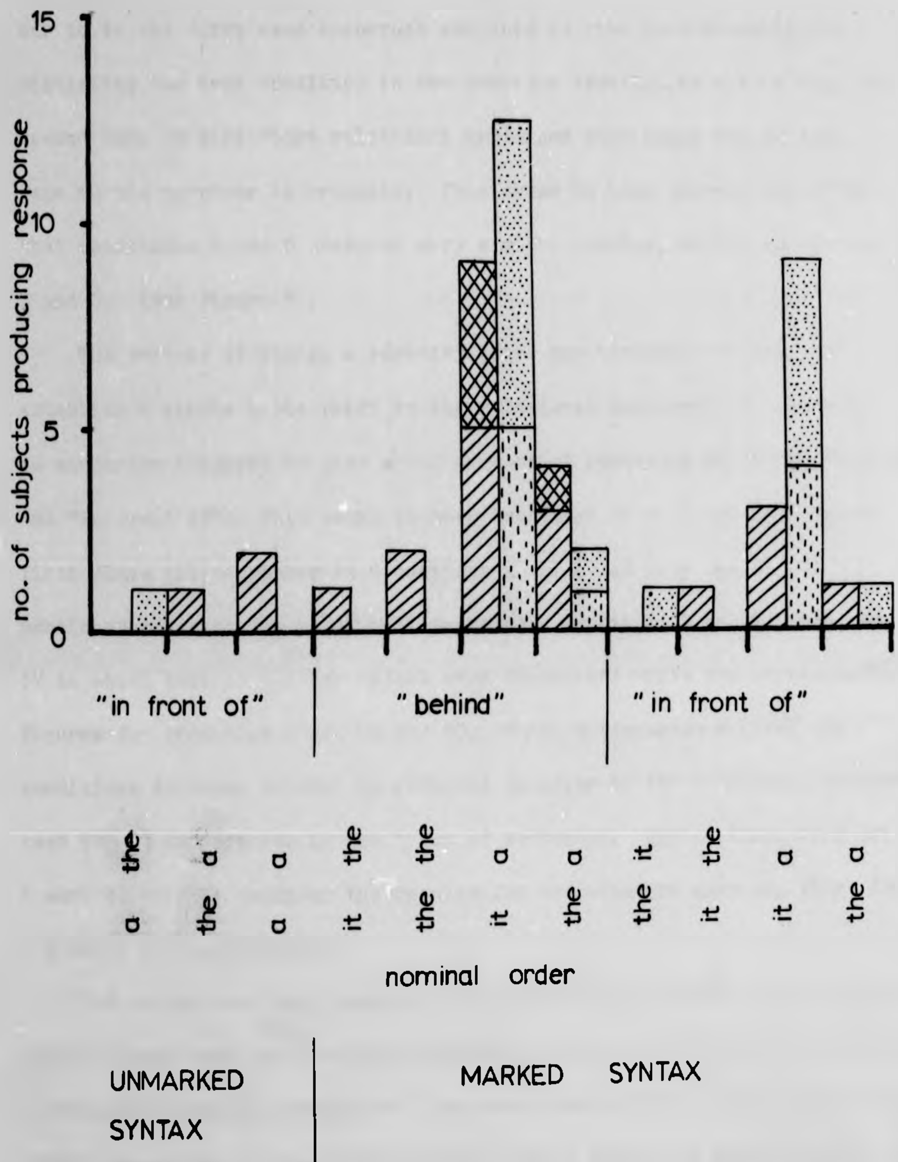


Figure 7 Part 2 Experiment 3 : distribution of responses in the four conditions (cont'd.)

Discussion

The purpose of giving instructions to either give detailed descriptions or simple abstract descriptions of position and direction was to in the first case encourage subjects to give long descriptions simulating the text condition in the previous experiments and in the second case to give short relational sentences simulating the no text case in the previous experiments. This seems to have largely failed in that conditions A and C produced very similar results, as did conditions B and D (see Figure 7).

The purpose of giving a starting point was twofold. Firstly to establish a single topic prior to the relational sentence; and secondly to encourage subjects to give equal numbers of sentences involving "behind" and "in front of". This seems to have succeeded on both counts. In the first place all responses in Conditions A and C had only one object mentioned prior to the relational sentence. Condition B had 73 such, and 57 in which both or neither object were mentioned; while the corresponding figures for condition D are 66 and 80. These differences between the conditions in terms of what is referred to prior to the relational sentences lead to big differences in the types of sentences. Before discussing this I want to briefly consider the results for definiteness marking, since these are quite straightforward.

The results are very similar to those of Grieve (1974) in that people almost always mark the previously mentioned object as definite and the newly introduced object as indefinite. Out of a total of 532 nominals there are only 7 exceptions to the first of these rules, and out of 436 examples there are only 20 exceptions to the second. These figures of 1% and 4.5% are not dissimilar to the figures of 0% and 8% from Grieve's Experiment 2. There seems to be no influence of any other factors on definiteness marking than previous mention - though there is perhaps a slight tendency to assume something was previously mentioned when it was not. This is in sharp contrast to the earlier experiment where a number of other factors were demonstrated to have an effect. In fact though if the variables had been defined in terms of position of the nominal in the sentence we would have found a main effect of both position variables and possibly higher order effects involving both topic position and syntax. This is because of an interaction with syntax which we will discuss later. To anticipate that discussion a little it may be that subjects in the previous experiment were torn between responding on the basis of previous mention and responding on the basis of the probability of certain configurations of definiteness marking with the two syntactic types. On the kind of functional account I am advocating selection of the marked syntax option is not independent of topicalisation (and hence in this case previous mention). This would partly explain the complex interactions produced by the last two experiments.

One important respect in which the present results do differ noticeably from Grieve's is in the use of pronouns rather than nominals with "the" as a way of indicating both definiteness and previous mention. In fact the vast majority of nominals referring to objects already referred to are pronouns : 310 as against 215 nominals with noun + "the" and 7 with noun + "a". What is more interesting is the different ratios in the case where only one noun was previously mentioned, compared to the case where both were. In the former case the figures are 283, 102 and 0; in the latter they are 27, 113 and 7 respectively. To express this differently : where only one object has been referred to there is a 3 in 5 chance that this will be referred to in the relational sentence by means of a pronoun; where both have been referred to there is only a 1 in 5 chance that <u>either</u> will be referred to by a pronoun. There are obvious reasons of clarity for this result. What is important for the present is

(1) the obviously large role played by pronouns in cohesion

(2) the possibility that some of the sentences in the comprehension experiment may have been made harder to understand because of the use of the relatively unlikely "the" + noun instead of a pronoun.

(3) the distinct possibility that (2) is not a random effect. In particular it may have had a much greater effect on the marked syntactic case since this is particularly likely to have a pronoun in the present experiment. (cf. the "typical" sentence of Condition A and the fact that marked syntax is very rarely used when both nouns have been previously mentioned - the case where "the" is most likely to occur). This is yet another reason for refusing to take the apparent main effect of syntax in Experiment 1 at face value.

Turning now to the position of the topic in the relational sentence : there is only one case of the previously mentioned object being second with marked syntax, 18 cases of both objects being previously mentioned, and 288 cases of the previously mentioned object being referred to first. Compared to these figures of 288 - 1 for the case where only one object is previously mentioned the figures for unmarked syntax are 3 cases of the topic first nominal and 29 cases of the topic second nominal. These figures provide very clear support for the topic position x syntax x text interaction found in the comprehension data (assuming, as seems reasonable, that cases where one object is previously mentioned are analogous to the text condition).

The only result wholly original to the present experiment concerns the classification in terms of subsequent mention. From a glance at the tables one can see the following (apart from the striking fact that unmarked syntax was only used on one occasion in each of Conditions A and C):-

(1) when unmarked syntax is used there seems to be a roughly equal probability that the next sentence will be about (a) the object referred to by the first nominal, (b) the object referred to by the second nominal, (c) both, or (d) neither.

(2) in contrast when marked syntax is used there is roughly an 8 in 9 chance that the succeeding sentence will be about the object referred to by the second nominal.

These results are quite dramatic especially when considered in the light of the parallel restriction of marked syntax to cases where only one object has been previously referred to, while unmarked syntax is not at all restricted in this manner. This is a very clear demonstration that (1) choices of discourse structure are involved in making choices within the sentence and (2) the syntactic form with the locative phrase first is a marked form, because of its restricted range of applicability. However the much greater frequency of use of the marked form in the present experiment clearly demonstrates that "marked" does not mean necessarily more difficult. In terms of a functional analysis "marked" means simply "having more complex entry requirements", but once these are satisfied it may be the simplest form.

The use of marked syntax in the present kind of communication task serves to build a kind of rhythmic structure in the discourse as a whole - a structure it is not possible to construct with unmarked syntax. It allows one to maintain a pattern of moving constantly from old to new information without having two successive chunks of new information. By placing the locative phrase with the topic noun first in the sentence one maintains an unmarked given/new structure and allows the listener to move steadily from what is familiar to him to what is unfamiliar. One directs his attention to a point in his knowledge structure and then gives the new information to be attached at that point. The use of unmarked syntax in the same situation is less satisfactory because one either has to put the new information first, which breaks the rhythmic structure, or else proceed to give the position of the object one has already referred to, and which may constitute the only reference point in the picture, by reference to something new. With regard to the rhythmic information structure the present situation differs from the verification experiment : there the structure of the task is clear and when the final sentence comes one naturally focusses on the new object; here the uttering of the relational sentence is not constrained in the same way and the speaker has to lead the listener from the present focus of attention to what he wants him to focus on next.

Finally, the De Soto et al. (1965) hypothesis about preferred directions of building displays: this is clearly substantiated by the evidence. It may explain : (1) the strong preference for "behind" with marked syntax shown in the present experiment, and weakly evident in the previous experiment; (2) the bias found in Experiment 1 towards faster RTs with "behind" when the nominals are differently marked for definiteness, but with "in front of" when they are similarly marked. With one object previously mentioned "behind" is consistent with the hypothesis (so long as the previously mentioned object is in the locative phrase), but with neither or both previously mentioned "in front of" is consistent with it (as long as unmarked syntax is used).

If one accepts this hypothesis one seems forced to reject the lexical marking interpretation of the superior performance with "in front of" in the no context case, and with "behind" in the context case. However I can think of no way of comparing the two hypotheses. It is of course possible that both are correct.

Chapter 4 : The role of pronouns in cohesion, with special reference to the three term series problem.

General Introduction to the Two Experiments.

The two experiments reported in the present chapter attempt to examine the use of pronouns. The first experiment asks "what is the difference between pronouns, names, definite and indefinite descriptions in terms of reaction time to understand sentences involving them?" This is looked at only for the text-embedded case and not for the single sentence case. No attempt is made to systematically vary any parameter other than the means of cross-reference between sentences.

The second experiment is rather more complex and seeks to examine the role of pronouns in a difficult "reasoning" task. It also examines lexical marking again, using a broad range of marked/unmarked pairs. The task used is the so-called "three term series problem". This is an interesting task in its own right, but is made even more so because of the work done on it by people working in other areas examined in the present thesis (in particular Clark, Huttenlocher and Johnson-Laird). In some cases they have tried to relate strategies shown to occur in other tasks to the three term series problem. This is also done here - emphasis being placed on the integration of information added in the second premise (new information) to that already expressed by the first premise, as well as the means of crossreferring from one premise to the other.

Experiment 4 : A very preliminary investigation of four methods of making anaphoric reference.

Introduction

The last experiment showed the greater probability of anaphoric reference (see Bolinger, 1972, Halliday and Hasan, 1976) across sentence boundaries being made by means of pronouns than by nominals containing "the". There was there some indication of a possible interaction with syntax. The present experiment sets aside the problem of such an interaction instead asking simply "can we show any effect on RT of using different means of making an anaphoric reference in a comprehension task?". Four methods of reference are used:-

Pronoun, name, noun + "the", noun + "a". These four methods differ in a number of respects - in fact two of them are not even clearly anaphoric. Taking them in turn:-

Pronouns can be clearly anaphoric : that is they can be used to refer back to something already referred to in the discourse. They can be used also cataphorically to refer forward to something which is clearly specified later and has not been specified previously. This use would appear to be marginal though. They can also be used, of course, to refer directly to something in the present situation (with accompanying gesture). Often a pronoun may be used to refer both endophorically (to another part of the text) and exophorically (to the situation).

Names are strictly speaking not anaphoric as they always refer directly to an object. However in a case where the hearer's only warrant that the object exists is the speaker's reference to an object by name, later reference to the object by name bears similarities at one level to anaphoric reference. On the other hand one might suppose that the reference is carried through in a simpler fashion by using the name than by using a pronoun since it would not seem necessary to locate the antecedent expression.

Definite Descriptions can be both cataphoric and anaphoric. (The term "definite description" is here used simply to denote noun phrases in which the noun is modified by the definite article and not with any intention to evoke the connotations that phrase has in philosophical circles - on that see Russell (1904) and Strawson (1975)).

Indefinite Descriptions are clearly not usually anaphoric - they are used to set up a reference rather than to refer to an object whose reference is specified more fully elsewhere. However it seems possible that people occasionally interpret them anaphorically.

It is obviously of some importance whether these expressions are interpreted anaphorically or not, so the present task is designed both to measure RT and to assess whether a co-referential or non co-referential interpretation was made.

Method

1. Subjects. 25 subjects, 13 male and 12 female, fulfilling a course requirement for an introductory psychology course. Mean age approximately 19 years.

2. Apparatus and Materials.

Subjects sat in a quiet chamber with a tape recorder (Revox A77) and microphone and a GT40 visual display linked to a PDP 11/45 computer. They had a single control button.

Materials consisted of 16 sentence sets each composed of an introductory sentence and the four different possible target sentences. Each first sentence was of the form : "Here is a x called y" where x and y were a common noun and name respectively. It was essential to introduce the name in the first sentence in order to use that as a means of cross-reference in the second sentence. A full list of the sentence sets is given in Appendix B.

3. Design and Procedure

Subjects were brought in and sat down in front of the GT40 and given the following instructions:-

"This experiment is concerned with the ability of people to paraphrase sentences. You will be presented with a series of sentences, one sentence at a time. When you think you have understood each one press the button on your right. The next sentence will appear immediately after you have pressed the button.

The sentences will be in groups of two. After each group there will be a pause of 9 seconds in which time you should speak aloud your paraphrase of the previous two sentences. At the end of the 9 seconds the first member of the next group of two will appear. You must complete your paraphrase in this time.

Try to work as quickly as possible, making especially sure you press the button as soon as you have understood the sentence being presented. Try to avoid rehearsing a paraphrase until after you have pressed the button indicating you have understood the second member of the pair".

In addition it was orally stressed to subjects that they must try to work as quickly as possible.

They were also told what a paraphrase is, if they did not understand the term.

It was made very explicit that they could paraphrase the two sentences either together or separately and that they did not have to follow a consistent strategy throughout the experiment.

As you can see from the Instructions the two sentences were presented one at a time with (subjectively) no delay between them, the second sentence coming up as soon as the first had finished. Subjects were encouraged to believe it was a paraphrase task and we were interested in the nature of the paraphrase. They had 9 seconds in which to give the paraphrase before the next set began.

Each set of materials consisted of a starter sentence and four second sentences : one each with the four different referring expressions. Second sentences were identical except for the referring expression. Each second sentence had only one nominal, and that was in first position.

Subjects had 16 trials - 4 of each type of referring expression. The order of material sets was randomised separately for each subject, as was the order of sentence types which was randomised independently of the material sets. The only constraint on this was that the sentence types were randomised in blocks of four.

The task lasted altogether about 10 minutes and subjects generally found it quite entertaining.

Results

The Computer measured RTs and printed them out, while Ss. paraphrases were scored from the tapes to see whether they had adopted the co-referential or non-co-referential interpretation.

Three subjects results were thrown out (two males and one female) because of exceptionally long times in one or more condition (10 seconds mean for a condition was the criterion - even this seems a little long).

The median times for the four conditions are as follows (medians are used because of the skewed nature of the distributions and are about 140 msec. lower than corresponding means):-

	Pronouns		Names		Definite Descs.		Indefinite Descs.
msec.	2553		2658		2786		3182
difference msec.		105		128		396	

The following differences are significant on sign tests:-

1. Pronouns - Names. 16 Ss show longer RTs for names $p < 0.05$
2. Pronouns - Definite Descriptions. 17 Ss show longer RTs for definite descriptions $p < 0.01$.
3. Pronouns - Indefinite Descriptions. 20 Ss show longer RTs for indefinite descriptions $p < 0.0001$.
4. Definite Descriptions - Indefinite Descriptions. 16 Ss show longer RTs for indefinite descriptions $p < 0.05$.

The difference between Names and Definite Descriptions is not significant ($p > 0.1$: 8 subjects having longer RTs for names), nor is that between Names and Indefinite Descriptions ($p > 0.1$: also with 8 subjects having longer RTs for names).

Despite this last result the following ordering seems entirely justified:-

Pronouns are responded to faster than Names or Definite Descriptions, which probably do not differ from one another, but are in turn responded to faster than Indefinite Descriptions.

This result is despite the fact that almost all subjects opt for the co-referential interpretation of the indefinitely marked referring expression. In fact only four subjects decide on the non co-referential interpretation : two of them on all 4 occasions, one three times and one only once. Of these four, three produced the non-co-referential interpretation of the definite descriptions - the first two 4 times and twice, and the third twice.

Discussion

The reaction time results could scarcely be more clear cut in one respect : pronouns definitely assist sentence comprehension where a co-referential interpretation is required. It seems extremely doubtful, to say the least, that they would do so in the single sentence case, though quite what the relationship would be between the other three types is impossible to predict.

These results seem to clearly rule out any theory which states that access to the representation of a referent is less direct when a pronoun is used - on the contrary it would appear to be more direct. Of course pronouns are shorter than the other three types and so perhaps takes less time to read. But any explanation in terms of reading time alone is doomed to failure : definite descriptions are longer than indefinite (just!), and yet take almost 400 msec. less time. In addition definite descriptions are generally longer than names - yet subjects do not take significantly longer to press the button. Thirdly, pronouns are scarcely any shorter than names, and yet RTs to them are significantly shorter. So it seems that the benefit accruing from using pronouns to make anaphoric reference must be accounted for at some less peripheral stage.

So far as I have been able to discover there is nothing at all in the literature to suggest just what the value of pronouns is, and why we should use them rather than names or definite descriptions. It is of course perfectly clear that they serve to cross refer but so do these other methods. They are certainly shorter in all languages than the average name but this in itself does not seem sufficient explanation. To explain their high frequency of use on the basis of their brevity would almost certainly be to put the cart before the horse : brevity is more likely to be the result of their high frequency of use - and frequency is in this case not an explanatory concept; it is what gives rise to high frequency that we are searching for. In the light of the evidence presented already in this thesis I would suggest that pronouns are especially useful as an additional way of indicating the distribution of new and old information in the sentence. Both names and definite descriptions do this to some extent but pronouns have the additional feature that they habitually relate to what was in the immediately previous discourse - something not true of the other two methods. They tell the listener not just "this is something you already know" but "this is something which is currently at the centre of your attention" - no search is required through memory to find the representation of the referent of the expression. Names and definite descriptions only imply that the listener is familiar with the referent. Indefinite descriptions on the other hand suggest that the listener is not familiar with the referent - a suggestion which most subjects in the present experiment clearly reject. Possibly this is because of the nature of the task : instructions to paraphrase probably encourage people to "integrate" sentences (in the Barclay and Bransford sense). Then again the obligation to refer to a previously mentioned object by means of some form which makes it clear that the object was previously mentioned is paralleled by another obligation. In some cases one needs to refer to another member of the same class as an object recently referred to. Instead of using the class name with the indefinite article, one then uses some term which makes it clear that another object is being referred to : usually, as repeatedly in the last few sentences here, "another". Thus whatever the interpretation intended for the "a" + noun cases, it should have been made clearer, so that the bias towards integration is perhaps more understandable.

This does not help us explain a rather odd result of the present experiment : namely a tendency in a few of the subjects not to integrate in cases where a definite description was used. This is rather odd. It may be due to the fact that the name of the object was given in the first sentence so that it seems odd to then go on to use a more general term such as a definite description. This explanation implies that there is an ordering between these two terms : the proper name supposing more knowledge than the generic noun. This may be true for a number of cases but it is obviously not true in general because of the conventions we adopt about just what classes of object are appropriately called by proper names. In addition, though it seems possible to extend this ordering back towards the less familiar it does not seem reasonable to suppose that pronouns are distinguished from the others solely on this basis, as was noted above. Indeed a pronoun may be used where one is not familiar enough with someone to know their name. This is because the warrant for the use of a pronoun is either the immediate situation (as in exophoric reference) or the immediately preceding discourse, whereas with names and definite descriptions there are other possible warrants as well.

Experiment 5 : The Three Term Series problem re-examined.

General Introduction to Experiment 5.

This experiment seeks to extend the findings of the last few experiments by comparing pronouns with names in a complex task in which several other parameters are varied also. The last experiment showed a clear difference in reaction time to pronouns and names in a simple paraphrase task. The superior performance with pronouns was attributed to the clearer division into new and old information which is achieved with pronouns together with the additional fact that pronouns give a clear indication that the object referred to was in the immediately preceding sentence - something which is not necessarily true of names. If the locus of the pronoun effect is in helping subjects to divide sentences into new and old information then it is obviously not unrelated to the topicalisation devices investigated in the first three experiments.

The present experiment seeks to make a beginning to investigating this as well as to cast more light on the processes involved in a much studied task : the three term series problem.

The Literature; Part One : Data

This section is concerned with a particular class of inferential problem in which the answer depends upon the relationships between items in the premises. These relational inferences can involve any number of premises, and a variety of different types of relation. An example given by Johnson-Laird (1972) in his review of the topic is the following:

> John stood in the last local elections in Camden.
>
> Camden is a borough of London.
>
> London had its annual borough elections on Tuesday.
>
> Therefore, John stood in the elections on Tuesday.

The particular concern here is with a special class of problems with only two premises each of which contains a comparative term (either the same comparative term in both premises or else the comparative term in one premise and its "converse" in the other).

e.g. Frank is taller than Jane.
Jane is taller than Ernie.
Who is tallest?

As Johnson-Laird points out the answer to these problems is not, strictly, a valid deduction. The answer however follows directly given a knowledge of English.

This type of problem is known as a 'linear syllogism' (something of a misnomer as they aren't syllogisms at all) or '3-term series problem'. There is a very large number of such problems : broadly they can be classified on the basis of the premise combinations and the types of questions. The premise combinations fall into six classes distinguishable on the basis of two dimensions : the type of relational term (either <u>comparative</u> or <u>negative equative</u>), and the ordering possibilities for the items (they may be strictly orderable in which case all items can be distinguished with respect to all other items; failing this they may be partially orderable, in which case at least one item can be distinguished from all the others, but there are also at least two items which cannot be distinguished from one another; finally the premises may be contradictory : here no item can be placed with respect to any other). This gives rise to the six possibilities illustrated with examples in Table 1. There are 8 members of each of the strictly orderable and partially orderable classes but only four members of the two contradictory classes, giving a total of 40 possible premise combinations.

The questions do not constitute such a neatly classifiable group, although there are more questions overall. The three basic questions simply ask one to specify the individual occupying a particular place in the series e.g. Which is greatest? Which is middle? Which is least? In addition there are 48 questions asking about a particular comparison e.g. Is x > y? Are x and y > z? Is x or y > z? Is x > z or y? Is x > z and y? etc. Not all of the questions are applicable to all of the premise combinations : for example the question '"Which is middle?' does not arise for contradictory premises which have only two terms. Furthermore the answers one gives vary too : 'can't tell', 'neither', 'yes', 'no' or the specification of an individual are all possible right answers to one or more of the questions, though they may also be nonsense if given in reply to some of the other questions.

Faced with such a bewildering variety of problems experimenters have, not surprisingly, been selective about the ones they have chosen to investigate. Generally they have experimented with strictly orderable comparatives, though Clark has also done some work on partially orderable comparatives and some on negative equatives (Clark, 1969b). Table 2 presents the results from four of the major published studies which have examined all of the strictly orderable comparative problems (from here on I will concentrate almost exclusively on this group of problems. For results from one study examining indeterminate problems see Table 3). Johnson-Laird (1972) has pointed out the potential importance of the number of trials a subject has in determining the results : there being evidence from an unpublished study by Wood (1969) that subjects switch strategies

Table 1

	Strictly orderable	Partially orderable	Contradictory
Comparative.	e.g. $A > B$ $B > C$	e.g. $A > B$ $C < A$	e.g. $A > B$ $B > A$
Neg. Eq.	e.g. A not as great as B. B not as great as C	e.g. A not as great as B C not as small as A	e.g. A not as great as B A not as small as B.

if they perform a large number of trials. The number of problems which subjects solved (including practice trials) is therefore also given in Table 2. As can be seen this varies quite considerably : from 32 for the Huttenlocher study to 104 for the experiment reported in Clark's Psychological Review paper (Clark, 1969a). Curiously enough there is no corresponding variation in the number of relational terms used : all studies used only one! This comes as something of a surprise when one remembers the degree of generality which is usually claimed for the results, not to mention the heated controversy between Huttenlocher and Clark over their respective models. Amazingly I have been unable to detect any comment in the literature to the effect that the repeated use of one relational term might itself both encourage the formation of a specialist strategy and dictate the nature of that strategy, though these suggestions seem highly plausible. Furthermore Clark's use of the pair better/worse and Huttenlocher's use of taller/shorter seem especially likely to encourage subjects to use a linguistic and imagery strategy, respectively.

The four studies whose results are summarised in Table 2 differ quite considerably both in the measures they used and in the way the experiments were performed. Clark's (1969b) study and De Soto et al.'s study both use error rates as the chief measure. They presented subjects with the problems on cards forcing subjects to answer within ten seconds in order to produce more errors. Clark's problems, however, use one of the two superlative questions "Who is best?" or "Who is worst?" and list the three possibilities with the subject having to underline one (in fact since this study also used problems in which the three elements were not strictly orderable Clark included a fourth response category : "Can't tell"). De Soto et al. on the other hand required a verbal response ("Yes" or "No") to the four questions comparing the two end members of the series (e.g. "Is A better than C?"). The overall error rates are very much higher for the De Soto et al. study than for Clark's study - even though the total number of trials is similar.

The third study using error rates is that of Huttenlocher (1968), but

Table 2 : Four Studies of the Three Term Series Problem.

Study : Premise Combinations	Clark (1969a) RT in Centiseconds	Huttenlocher (1968) RT in Centiseconds	Huttenlocher (1968) % Error	De Soto et al. (1965) % Error	Clark (1969b) % Error
1 > 2 3 > 1	52	135	11	47	12
1 > 2 2 > 3	57	155	17	40	22
1 > 2 3 < 2	53	141	10	38	9
1 > 2 1 < 3	58	157	19	62	38
1 < 2 3 > 2	53	142	9	43	11
1 < 2 1 > 3	55	157	18	50	23
1 < 2 3 < 1	64	142	8	50	21
1 < 2 2 < 3	55	161	14	58	46
Number of Problems:	104	32		64	80
Number of Subjects:	13	48		117	100

All four studies use only one relational term pair.

Table 3 : Clark's data on 8 problems in which it is not possible to produce a strict ordering of the elements.

Premise Combinations	Data from Clark (1969b) % Error
1 > 2 3 > 2	10
1 > 2 1 > 3	8
1 < 2 3 < 2	17
1 < 2 1 < 3	10
1 > 2 2 < 3	42
1 < 2 3 > 1	41
1 > 2 3 < 1	33
1 < 2 2 > 3	38

here subjects were not constrained by time : that is they were allowed as much time as they needed though they were encouraged to go rapidly. Huttenlocher measured both errors and reaction times and not surprisingly the error rates she found were much lower than in either of the other two studies. She presented the problems orally and asked subjects one of the two superlative questions : "Who is tallest?" or "Who is smallest?". Subjects answered orally. One major difference between this study and the others is that Huttenlocher asked subjects the two possible comparative questions after the first premise in order to check that subjects had understood it. This method seems likely to minimise the contribution of linguistic factors, especially those operating on the first premise, to the final result. It probably encourages subjects to code the information in a markedness-free form, since they have to access information contained in the first premise with both a marked and an unmarked question.

The fourth study summarised in Table 2 is again by Clark (1969a) and used only reaction times (which he seems to regard as less precise than error rates, though it is unclear why). The method was substantially as in his other study : the main difference being that all the problems were determinate so that subjects had no can't tell option.

One important point to note about the four studies is their mode of presentation of the problems : in the De Soto et al. and Clark studies the whole problem (two premises and the question) is presented, written on a single card and all parts are visible simultaneously. This is not the case with the Huttenlocher study where by the very nature of oral presentation the three parts are not available together, this being further exaggerated by her asking subjects about the first premise before presenting the second premise and the question. Huttenlocher's is the only study where it is impossible for the subject to process the question prior to processing the premises. Subjects in her experiment have to process both premises sufficiently to hold them in memory before processing the question and it may well benefit them to integrate the two into a single representation. In the other studies the subject does not have to integrate to the same extent : he simply has to scan each premise for whichever individual satisfies the superlative predicate. There is the further point that in Huttenlocher's study interest is focussed on the new item when the second premise is presented and the subject has to fit this into the representation he already has. But in De Soto et al. and Clark's studies subjects can read the second premise first, so that it is not surprising that Clark feels that the position of the new item in the second premise is, of itself, unimportant (though Clark's model does in fact lead one to predict an overall effect of the position of the new item - see Table 9).

It is rather surprising that all the models extant in the literature are designed such that subjects are always assumed to process the sentences in the order in which they are written (or spoken) namely first premise, second premise, question. This appears to be a somewhat gratuitous assumption for all the data except Huttenlocher's (and this includes the study by Hunter (1957) not included here because it does not cover all strictly orderable premise combinations). The design of Huttenlocher's study guarantees that subjects at least receive the problem in the assumed order - although there is no guarantee that they process it to any significant depth in that order (except, presumably, for the first premise).

The Literature; Part Two : Models.

As we have already seen the two major current models are the so-called "linguistic" model of Clark and the "image" model due to De Soto et al. and to Huttenlocher. Figure 1 presents the linguistic model as formalised by Johnson-Laird, with a small amendment by myself. I have omitted those sections of the model only necessary for dealing with negative equatives and concentrated on the more straightforward comparative problems. In addition I have added the instruction "change less to least" to box number 8 : without this small change the model gets into a loop (between boxes

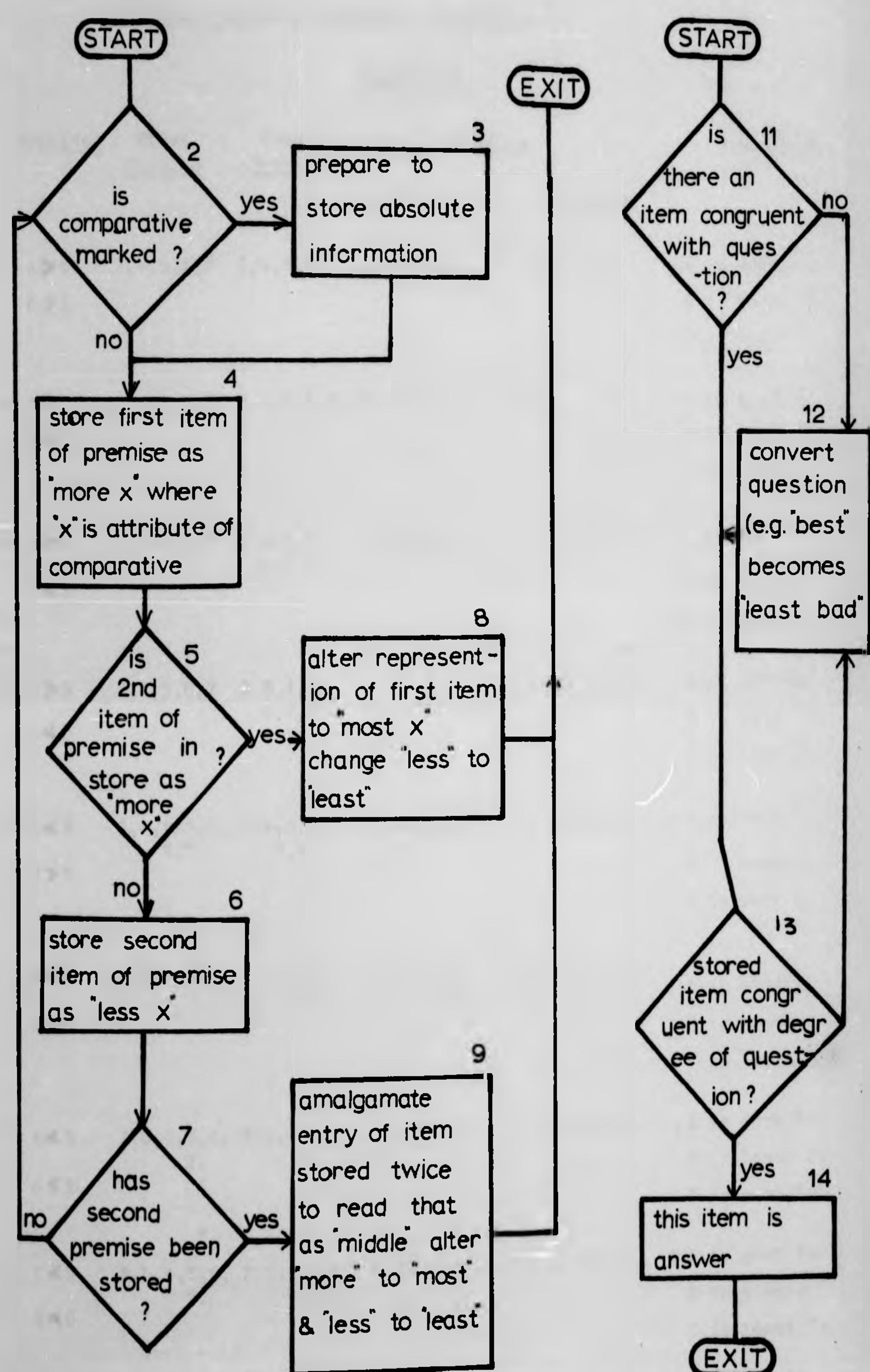


FIGURE 1 Johnson-Laird's formulation of Clark's linguistic model

Table 4 : A slightly revised version of Johnson-Laird's formulation of Clark's model : predicted operations.[a]

Premises	First Premise	Second Premise	Question: marked	Question: unmarked	Base Forms
1. A > B C > A	2,4,5,6,7	2,4,5,8	11,12,13,14	11,13,14	A is more x B is least x C is most x
2. A > B B > C	2,4,5,6,7	2,4,5,6,7,9	11,12,13,14	11,13,14	A is most x B is middle C is least x
3. A > B C < B	2,4,5,6,7	2,3,4,5,6,7,9	11,13,14	11,13,14	A is most x B is middle C is most 7x
4. A > B A < C	2,4,5,6,7	2,3,4,5,6,7,9	11,13,12,13,14	11,13,12,13,14.	A is middle B is least x C is least 7x
5. A < B C > B	2,3,4,5,6,7.	2,4,5,6,7,9.	11,13,14	11,13,14	A is most 7x B is middle C is most x
6. A < B A > C	2,3,4,5,6,7.	2,4,5,6,7,9.	11,13,12,13,14.	11,13,12,13,14.	A is middle B is least 7x C is least x
7. A < B C < A	2,3,4,5,6,7.	2,3,4,5,8.	11,13,14.	11,12,13,14.	A is more 7x B is least 7x C is most 7x
8. A < B B < C	2,3,4,5,6,7.	2,3,4,5,6,7,9.	11,13,14.	11,12,13,14.	A is most 7x B is middle C is least 7x

a Conventions: ">" denotes an unmarked term, "<" denotes a marked term in the premise schemata. In the base string schemata "x" denotes an unmarked and "7x" a marked term. These operations are derived from the flow chart presented in Figure 1.

Table 5 : An extremely crude method of comparing the models of Huttenlocher and Clark with the data from the four studies

Problem[a]	Actual[b] Order	No. of Clark[c] Operations	Clark Order	No. of Huttenlocher Operations	Huttenlocher Order
1. A > B C > A	7	12.5	8	9	3
2. A > B B > C	5	14.5	6.5	8	6
3. A > B C < B	8	15	4.5	7	8
4. A > B A < C	1	17	1.5	10	1
5. A < B C > B	6	15	4.5	8	6
6. A < B A > C	3	17	1.5	9	3
7. A < B C < A	4	14.5	6.5	8	6
8. A < B B < C	2	16.5	3	9	3

N.B. The rank ordering of actual difficulty is derived from an average of the four studies' order of difficulty for each problem (using the RT data from the Huttenlocher study). See Text.

a Conventions as to premise type as in Table 4.

b Orders are given from hardest to easiest : thus a lower rank denotes a harder problem.

c This is derived from an average of marked and unmarked questions.

12 and 13) on problem 1 with a marked question and on problem 7 with an unmarked question (see Table 4). Despite its rather mixed pedigree I will refer to this model as Clark's model since most of the ideas are his. Clark may not in fact agree (and may never have agreed) with all the details of the model, however. Table 4 gives the operations which would need to be performed for each premise combination and for both superlative questions. Each number in Table 4 corresponds to one of the numbered operations in Figure 1. If we make the simplifying - but theoretically unmotivated - assumption that each of these operations is equally complex then we can add them all up to yield a total number of operations for each problem type and this will give us a crude basis on which to compare the different problems.[1] At the same time it is possible to get a crude estimate of the actual difficulty of the various problems by rank ordering the problems on the basis of the results in Table 2 (ranking the problems within each set of data and then averaging these ranks for each problem and so deriving an overall rank ordering of the problems).[2]

1 In fact this simplifying assumption is not nearly as gross as it at first sight appears : so many of the operations cancel out across problems that very little hangs on this assumption. In the first premise all but operation 3 cancel out - making the four problems with this operation harder than the rest. In the second premise only operations 3,6,7,8 and 9 do not cancel out and the only barrier to forming a strict ordering is the relative complexity of operation 8 on the one hand and the sum of operations 6,7 and 9 on the other. Finally with regard to the question the problems fall into three categories : those with neither operation 12 nor operation 13, those with operation 12 only and those with both operation 12 and operation 13.

2 The figures which follow are based on the RT data from the Huttenlocher study. To simplify I have simply averaged over the marked and unmarked question for each premise combination.

According to Clark's model, then, problems 4 and 6 should be hardest with 17 operations : in fact problem 4 is hardest and problem 6 third hardest. Next should be problem 8 with 16.5 operations : in fact it is second hardest. These three are clearly harder than the next group of problems in terms of number of Clark operations : actually 1.5 operations more than the next premise combination. In fact problems 3 and 5 should be next hardest by Clark's model but they are easiest and third easiest respectively (i.e. sixth and eighth hardest instead of having a rank of 4.5). Joint sixth hardest by number of operations are problems 2 and 7, but they are in fact fifth and fourth respectively. Finally problem 1 is supposed to be easiest, but it is in fact next to easiest. Considering the fact that these scores are derived from an extremely crude averaging system over several different paradigms the degree of fit is not at all bad. In fact the Spearman correlation co-efficient between the two orderings is approximately 0.69 ($p < 0.05$ with 6df). On the other hand though the fit to Clark's own data only produces the same correlation co-efficient for his error data and a very low one indeed ($p = 0.15$) for his RT data.

Of course this is an extraordinarily crude method of comparison but the image model, if anything performs slightly better than Clark's model. Johnson-Laird's formalisation of this model is presented in Figure 2 and the operations for each premise combination are presented in Table 6. One can predict from this model that problem 4 will be hardest, problem 6 third hardest, problem 5 sixth hardest and problem 3 easiest: all of these predictions are correct. On the other hand problem 8 is predicted to be third hardest, when it is second hardest, problem 7 to be sixth, when it is fourth, problem 2 to be sixth also (but it is fifth) and problem 1 to be third, whereas in fact it is seventh. This yields an overall rank order correlation coefficient between the two ordering of $p = 0.76$ (6df, $p < 0.05$). This is marginally better than Clark's model and is achieved with a much simpler model. However, the image model

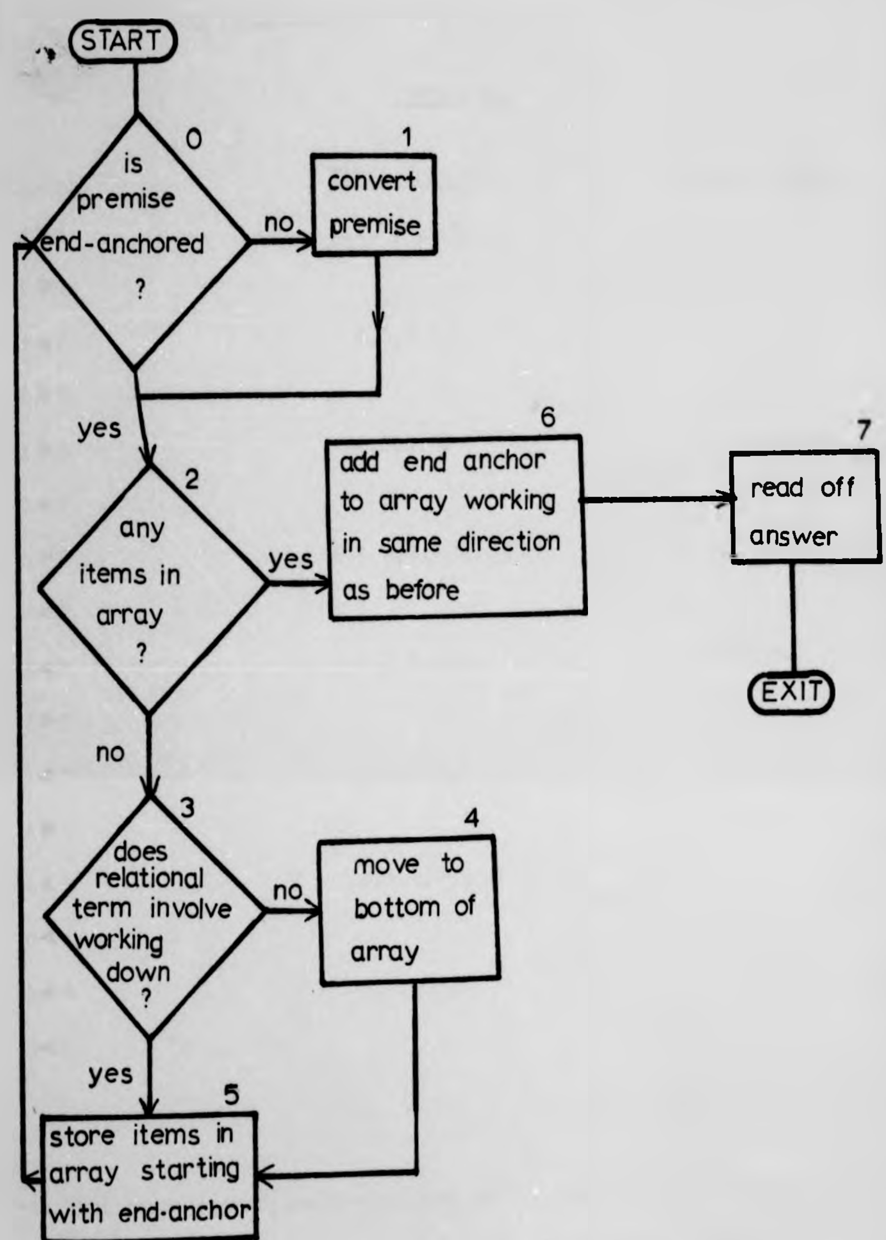


FIGURE 2 Johnson-Laird s formulation of the image model

Table 6 Johnson-Laird's formulation of Huttenlocher's model : predicted operations.

	Operations	
Premises	First Premise	Second Premise
1. A > B C > A	0,1,2,3,4,5.	0,2,6.
2. A > B B > C	0,2,3,5.	0,1,2,6.
3. A > B C < B	0,2,3,5.	0,2,6.
4. A > B A < C	0,1,2,3,4,5.	0,1,2,6.
5. A < B C > B	0,2,3,4,5.	0,2,6.
6. A < B A > C	0,1,2,3,5.	0,1,2,6.
7. A < B C < A	0,1,2,3,5.	0,2,6.
8. A < B B < C	0,2,3,4,5.	0,1,2,6.

Note that these operations are carried out at the integration stage and the first premise operations make no prediction about first premise times as measured in the present experiment.

Convention: '>' denotes an unmarked and '<' a marked term. See also Fig.2.

Table 6 Johnson-Laird's formulation of Huttenlocher's model : predicted operations.

	Operations	
Premises	First Premise	Second Premise
1. A > B C > A	0,1,2,3,4,5.	0,2,6.
2. A > B B > C	0,2,3,5.	0,1,2,6.
3. A > B C < B	0,2,3,5.	0,2,6.
4. A > B A < C	0,1,2,3,4,5.	0,1,2,6.
5. A < B C > B	0,2,3,4,5.	0,2,6.
6. A < B A > C	0,1,2,3,5.	0,1,2,6.
7. A < B C < A	0,1,2,3,5.	0,2,6.
8. A < B B < C	0,2,3,4,5.	0,1,2,6.

Note that these operations are carried out at the integration stage and the first premise operations make no prediction about first premise times as measured in the present experiment.

Convention: '>' denotes an unmarked and '<' a marked term. See also Fig.2.

provides no account of differences between a premise combination with the marked question and one with the unmarked, whereas Clark's model does. Furthermore the image model only allows one to distinguish four levels of difficulty whereas Clark's model distinguishes five; this inevitably makes the latter more attractive.

Overall though there is really little reason for choosing one model rather than the other : they both show a moderately good fit to the data and both appear to exhibit some of the important structure of the language involved. On the other hand one cannot happily leave it at that : the two theories present very different interpretations of what is going on when somebody is trying to solve one of these problems, and they are not compatible. Johnson-Laird accepts both emphasising that subjects' strategies vary both between individuals and, perhaps more importantly, within individuals depending upon experience with the problems. He suggests people are likely to start using a strategy similar to the one embodied in Huttenlocher's model but, with practice, are likely to switch to a strategy more closely approximating Clark's model. The emphasis Johnson-Laird places upon strategies (as opposed to some deeper level, less easily manipulated set of procedures) seems appropriate to the experiments reported in the literature for reasons already noted above, namely the large number of trials used and the use of only one pair of relational terms throughout each experiment. However one must wonder at the relevance of such studies to relational inferences in general. In everyday life we presumably make such inferences quite frequently (though rarely in blocks!) and with quite a variety of different relations. It would be surprising if the general properties of English sentences were not relevant to performance on these problems. Both Clark and Huttenlocher would undoubtedly agree with this - witness Clark's (1974) generalised model and Huttenlocher's earlier work with children on the importance of the relationship between a statement and the situation it describes (Huttenlocher, Eisenberg and Strauss, 1968; Huttenlocher and Strauss, 1968).

But Johnson-Laird's claim that the two models are merely formalisations of different strategies implicitly contradicts this : it suggests that these are properties of English we <u>can</u> exploit if need be, but that they are not properties which we <u>habitually</u> exploit in understanding English sentences.

In the experiment reported in this chapter I have attempted to re-examine the three term series problem from the point of view of a more general theory of language (SG) and in particular with emphasis on the role of cohesive devices (chiefly pronouns) in assisting the subject to integrate the two premises. In addition this experiment is better suited to testing models of the sort presented by Clark and Huttenlocher (as formalised by Johnson-Laird) in that it allows us to examine processing of each of the premises and of the conclusion separately, so permitting fairly rigorous testing of the different parts of the models. The experimental technique is such that it would be unreasonable to suppose that subjects process the three sentences in any order other than first premise, second premise, question, so that the data satisfies one of the major assumptions of the models.

<u>The role of cohesion and systemic choices in the three term series problem</u>.

For the sake of simplicity I have until now assumed that the first sentence in a discourse will not display any deliberate thematic choice[3] : there will always be a thematic structure, by definition, but this will always be the least marked type. This simplifying assumption, though it seems likely to hold in many cases, is undoubtedly not justified for the whole range of possible cases. Partly this is because of the possibility of multiple themes including adjunct themes (e.g. "Once upon a time, in a

3 Of course isolated single sentences, as used in some of the experiments already presented, are simply a special case of this.

land far, far away, there lived a beautiful maiden ....". On this see Chapter 6 and Halliday and Hasan, 1976). Partly also it is because of the possibility of various stylistic reasons dictating a marked theme at the beginning of a discourse (e.g. the desire to instil in a reader the idea he is entering an ongoing drama).[4]

The relevance of this for the present discussion is as follows. If a marked lexical form is used in the first sentence of a three term series problem (or any other "discourse") it is unclear to the listener whether this is due to a thematic choice on the part of the speaker or whether it is due to a desire to convey information about the absolute position of the objects on the relevant dimension. It is impossible for the hearer to tell whether it is thematic or not until he hears what follows the first sentence. But in any case it seems it must also be intended as a way of conveying absolute information. This is so because if the second premise has the same theme as the first then (1) if it has a marked relational term the object has twice been compared unfavourably to other objects and so the marked option looks like a way of conveying absolute information[5]; (2) if the second premise is unmarked then what could be the reason for ordering the premises this way rather than with the unmarked first, other than the desire to convey absolute information?[6] So, whether the choice is thematic or not, it has to convey absolute information. We may therefore expect it to take longer

4 In fact marked theme - other than adjunct - is very rare at the beginning of a discourse. The most common method of giving the impression that the reader is entering into the middle of things is to begin a novel with either a name or a pronoun, both of these being extremely common in the modern novel.

5 Note that this combination of premises only allows a partial ordering.

6 This argument does not seem to apply to the same extent where more than two premises are involved as topicalisation choices in the third premise might be relevant (i.e. in the N-term series problem where $N > 3$).

to encode.[7] In this I follow both the linguistic and the image model, though protagonists of the latter model take this view because of a belief about the way images are built up which neither Clark nor myself believe to be general enough to account for the effect.[8]

In understanding the second premise the subject has one major task to perform : the integration of the new item into the representation formed from the previous premise. This task has several components : recognition of the new item, identification of the old item and its association with the correct item to the representation. In addition there may be some attempt to assess the absolute position on the dimension of the three objects. However if the thematic choice can be seen to make sense in some way then the marking of the second premise is, of itself, irrelevant and should be seen as purely thematic. In that case marking of the first

7 Elsewhere I oscillate between two interpretations of marking : (1) that it is a purely thematic choice and the marked form only leads to longer RTs etc. if this choice is unjustifiable in the hearer's eyes; (2) that it is both a thematic choice and in some cases, where the thematic choice interpretation seems unlikely, a way of conveying absolute information. Here I am taking the second view.

8 Both here and in other work on this subject the problem of syncategorematicity is inadequately dealt with. This is the phenomenon whereby the usual markedness designations are reversed due to their collocation with other terms. Thus "tall" is usually considered unmarked and "short" marked, but we can have a "tall dwarf" and a "rather short skyscraper." This is an important problem in that it brings in questions about the independence of linguistic knowledge from general knowledge of the world. It has implications also for the way one constructs materials for experiments like the present one. I will touch on it again but the dimensions of the problem preclude thorough treatment in the present thesis.

premise will dominate encoding of the second (i.e. the second premise will be coded the same as the first since the new item is assumed to be simply added to the representation formed from the first premise). This may not be the case if subjects employ some special local strategy to solve the problem and do not use any normal process of linguistic integration.

A major parameter likely to affect all of this is the use of pronouns. The oral descriptions of pictures study showed that they are preferred to the definite article as a means of referring to objects also referred to in the preceding sentence. The last study showed that, for a set of simple sentences, they lead to more rapid comprehension when they refer to the object also referred to in the previous sentence. They seem to be the natural way to make anaphoric references between adjacent sentences (whereas definite descriptions seem to be a way of making an anaphoric reference over a greater interval). They lead also to a very clear distinction between new and old information. This is usually indicated by intonation but this means is not available in writing, so that pronouns probably have a greater part to play in indicating new and old information in written than in spoken text. This seems likely to be even more so in cases where written sentences have a marked given/new structure - i.e. where "given" information does not precede new information in the sentence.

In the three term series problem the subject faced with the second premise will be looking to "solve for" the new item : it should therefore help if the new item is thematic. But this is contrary to the usual given/new structure so that two conflicting tendencies are set up : to assume that the theme has been previously mentioned, and to assume that it will be the new item. This conflict is resolved if the sentence is clearly divided into new and old information by means of a pronoun. There is another factor to be considered here though : namely whether the first or second noun phrase in the first premise is pronominalised in the second premise. The third experiment showed that the pronoun is rarely the grammatical subject of the sentence. In addition there was a strong tendency in Experiment 3 for subsequent sentences to be about the grammatical subject of the sentence preceding them (the first noun with unmarked syntax but the second noun with marked syntax). That there is such a tendency can be easily seen from a couple of examples in which one controls for pragmatic expectations:-

e.g. 1 ..... He showed us a picture he took on the Serengeti Plains. A lion was standing beside a cheetah. It looked incredibly fierce.

e.g. 2 ..... Guess what I saw on the way up here! A silver Ferrari chasing a red Alfa Romeo. It was being driven by a guy with a turban on.

One might expect this tendency to play a part in integrating the premises in three term series problems if cross-reference is achieved through using a pronoun. It is possible that a similar effect could occur with names but this seems less likely[9] - in fact one reason for using names would seem to be the desire to talk about something other than the grammatical subject of the previous sentence.

So far I have not touched on the process of question answering. Huttenlocher, as we have noted, fails to give any account of the process. Clark's account is complex in its detailed application though quite simple to formulate. Basically it hinges on two processes : (1) finding a base congruent with the question (recall that Clark's model has us store the information from the two premises as three separate nuclear propositions - see Table 4); (2) if this succeeds finding an item congruent with the degree of the question (positive, comparative or superlative). If either of these fails the question is converted e.g. from "best" to "least bad". A list of some of the predictions which can be derived from Clark's model is presented in Table 10. These will be considered more fully in the Discussion.

9 i.e. superiority when the name in sentence N + 1 refers to the same object as the grammatical subject rather than the grammatical object of sentence N.

The present theory is not directly concerned with the question answering process. However it does allow one to make some predictions. Firstly there should be no effect of lexical marking in the question since the decision as to which lexical item to use does not depend on anything other than what the speaker is interested in and hence neither possibility carries more information than its alternative. If the two premises are integrated into a unitary representation there should be no effects at all of any factor - except perhaps that questions should be answered quicker if the new item is the answer since this is likely to be at the forefront of the subject's attention. What strategy will be adopted when the two premises are not fully integrated it is impossible to say on the basis of the present theory. It does however allow one to predict that integration will be slower with names and hence these will show more interactions with other factors than will pronouns.

Method

1. Subjects

48 first year undergraduate psychology students fulfilling a course requirement.

They were divided into 4 groups : two groups attempted problems in which cross reference between the premises was by means of pronouns, and two received problems in which it was by means of names (see Materials). All groups were composed of equal numbers of males and females.

2. Materials.

Materials were constructed on the basis of 40 bipolar adjective pairs and 120 different first names (80 male names and 40 female).

Eight of the pairs of words were used on the 3 practice trials. These are shown in Appendix C.

The members of all 40 pairs were classed as marked or unmarked in a rather informal fashion. The experimenter attempted to pick out the unmarked member of any pair by (1)intuitions as to which adjective would be used if one wanted to ask a neutral question; (2) which of the abstract nouns associated with the two adjectives correctly denoted the dimension as a whole. This categorisation was checked by asking three other native speakers of British English. Disputed or undecided cases were decided by reference to frequency : the more frequent member of the pair was assumed unmarked. This appears a rather arbitrary strategy and no doubt several proponents of semantic feature theory would dispute its correctness arguing, perhaps, that many of the pairs consist of members both of which are marked. To this there are two replies : (1) if one restricts one's designation of marked/unmarked to those pairs which satisfy all the classical criteria, then since the number of such pairs is extremely small, the range of applicability of the concept is so small as to make it uninteresting; (2) Greenberg, in his review, points out that there is often no clear basis other than frequency for categorising the two members of a pair as marked/unmarked, but that we should neverthe-

Method

1. Subjects

48 First year undergraduate psychology students fulfilling a course requirement.

They were divided into 4 groups : two groups attempted problems in which cross reference between the premises was by means of pronouns, and two received problems in which it was by means of names (see Materials). All groups were composed of equal numbers of males and females.

2. Materials.

Materials were constructed on the basis of 40 bipolar adjective pairs and 120 different first names (80 male names and 40 female).

Eight of the pairs of words were used on the 8 practice trials. These are shown in Appendix C.

The members of all 40 pairs were classed as marked or unmarked in a rather informal fashion. The experimenter attempted to pick out the unmarked member of any pair by (1)intuitions as to which adjective would be used if one wanted to ask a neutral question; (2) which of the abstract nouns associated with the two adjectives correctly denoted the dimension as a whole. This categorisation was checked by asking three other native speakers of British English. Disputed or undecided cases were decided by reference to frequency : the more frequent member of the pair was assumed unmarked. This appears a rather arbitrary strategy and no doubt several proponents of semantic feature theory would dispute its correctness arguing, perhaps, that many of the pairs consist of members both of which are marked. To this there are two replies : (1) if one restricts one's designation of marked/unmarked to those pairs which satisfy all the classical criteria, then since the number of such pairs is extremely small, the range of applicability of the concept is so small as to make it uninteresting; (2) Greenberg, in his review, points out that there is often no clear basis other than frequency for categorising the two members of a pair as marked/unmarked, but that we should nevertheless preserve the distinction because of relevance to so many different linguistic phenomena. Table 7 lists all the 32 pairs used in the experimental trials together with their designation (marked/unmarked), the mean RT to them in the first premise and their Lorge Magazine Count frequency (Thorndike and Lorge, 1944). Of the 32 pairs 24 have the unmarked member more frequent than the marked. The correlation of the frequency of marked and unmarked members of each pair is $\rho = 0.48$ (df 30, $p < 0.01$).

The materials were constructed in the following way. Firstly two male names and one female name were assigned to each of the 32 adjective pairs at random. Next a suitable noun was found to go with the adjective if the adjective could not be applied to people. Thus there was no need for an extra noun for big/small since people can be said to be bigger than/smaller than one another; but long/short are not appropriately applied to people so "garden" was used to give, in one case, "Neville's garden is longer than Helen's. Peter's is longer than Neville's. Whose is longest?" Because of the need to have an unambiguous pronoun reference in the second premise the first premises always mention one male and one female (the person referred to for the first time in the second premise was, arbitrarily, always male).

Having constructed 32 sets of name, adjective, (additional noun) combinations 4 problems were made up for each of the 8 premise combinations from these - assignment of problem type to the name etc. sets being at random. Half of the problems for each premise combination received the marked question, half the unmarked.

Next a second list was constructed using the same 32 name, adjective, (additional noun) sets. Again problems were constructed by assigning premise combinations randomly to these, but with the restriction that if any combination had had a marked first premise in the first set it must have an unmarked first premise in the second set.

Finally two more lists were made, the same as lists 1 and 2 only instead of having names throughout, lists 3 and 4 had pronouns to cross-refer between sentences.

The 8 practice problems were constructed from a further 8 adjective, name, (additional noun) sets and consisted of a random selection of 8 of the 16 possible problems, half of them with pronouns. This set was used for all subjects.

3. Apparatus and Procedure

Problems were presented on a CRT screen (GT40 visual display) controlled by a PDP 11/45 computer. Each subject received the 8 practice problems followed after a short break (about 50 seconds) by the 32 experimental problems of one of the four sets. Order of presentation of both practice and experimental trials was randomised separately for each subject. There was an intertrial interval of ten seconds. A trial started independently of S who was instructed to press a button when he understood the first premise. As soon as he did so this was deleted and replaced immediately with the second premise. This was replaced by the question as soon as S had pressed the button to indicate he had understood it. When subjects thought they had the answer to the question they said it aloud at the same time pressing the button for the third time. Oral responses were recorded. Subjectively, there is no delay between pressing the button and the sentence being replaced by the next sentence (or a blank if it is the question). The subject sat alone in a cubicle with the CRT screen and a microphone throughout the experiment. Reaction times for all three responses were taken by the computer from the button-pressing. Accuracy information was added from the audio tape.

Subjects had the experiment explained to them by the following instructions:

"In this experiment you have to solve a number of logical problems involving transitivity. A simple example of such a problem is : $A > B$, $A < C$; which is greatest? In this experiment however all the problems are fully written out in ordinary English.

You will be presented with the premises one at a time. When you think you understand each one press the button to your right. After you have read both premises and indicated that you have understood them, you will be presented with a question. Answer this orally at the same time pressing the button to your right again. Press the button as soon as you start to speak.

It is important that you should try to work as quickly as possible, while trying also to avoid making any errors. You will inevitably make some errors on the side of going too fast and making more errors rather than going too slow and making fewer errors.

You will be presented with two blocks of trials : a short one and a long one. There will be a gap of about one minute between blocks.

Once again : it is important that you should try to go as quickly as possible". In addition they were strongly encouraged to work as fast as possible by the experimenter.

Table 7. Adjectives used together with the mean RT (in msec.) to each of them in the first premise and their Lorge Magazine Count Frequency.

	Unmarked			Marked		
	Adjective	Lorge Frequency	Mean RT[a]	Adjective	Lorge Frequency	Mean RT[a]
1	longer	5362	3555	shorter	887	3582
2	farther	1835	3624	nearer	1338	3766
3	happier	1449	2578	sadder	202	2909
4	faster	514	3224	slower	434	3572
5	tidier	42	2668	sloppier	20	2797
6	wider	593	4499	narrower	391	4066
7	fatter	512	2608	thinner	646	2744
8	brighter	645	2829	duller	289	2908
9	lighter	2387	3203	darker	1005	3515
10	cleaner	781	3641	dirtier	221	4038
11	hotter	1006	3166	colder	1092	3598
12	smoother	346	4056	rougher	294	4050
13	sharper	324	3253	blunter	26	3456
14	wetter	319	3730	drier	592	3502
15	fresher	551	3499	staler	46	4220
16	wiser	420	2767	stupider	144	3250
17	tighter	264	4128	looser	274	3791
18	sweeter	679	4010	sourer	102	3790
19	crisper	154	4362	soggier	13	4272
20	clearer	587	3599	cloudier	20	3502
21	richer	656	3108	poorer	837	2769
22	stronger	770	3228	weaker	276	3570
23	deeper	881	3766	shallower	104	3369
24	heavier	680	3743	lighter	2387	3370
25	healthier	207	3136	sicker	615	2845
26	harder	1909	3823	softer	549	3292
27	dearer	1326	3817	cheaper	327	3385
28	subtler	770	3320	cruder	276	3020
29	louder	214	3647	softer	549	4240
30	commoner	568	3775	rarer	172	3587
31	harder	1909	4071	easier	1077	3802
32	tauter	36	4576	slacker	30	4467

Unmarked Adjectives 1 - 16 from Groups 1 and 3

a units are milliseconds.

(cont'd.)

Marked Adjectives 17 - 32 from Groups 1 and 3

Unmarked Adjectives 17 - 32 from Groups 2 and 4

Marked Adjectives 1 - 16 from Groups 2 and 4.

Groups 2 and 4 appear to be slower than Groups 1 and 3.

Results

Results are in four sections:-

1. Reaction times to the first premise.
2. Reaction times to the second premise.
3. Reaction times to the question.
4. Errors.

All means are given in Table 8.

As usual in this thesis all reaction times are used - no correction being made for errors.

1. The First Premise.

The mean times for the different adjectives are given in Table 7, along with the frequency for each word in the Lorge Magazine Count. Each mean is based upon 24 responses : 12 each from the pronoun problems and the name problems. Results are given on the basis of the adjectives rather than the conditions as the latter are not evident to subjects at this point. Overall means for the marked and unmarked adjectives are identical : 3533 milliseconds. This is despite the overall higher frequency of the unmarked items (for 24 of the 32 pairs the unmarked member is more frequent. This is highly significant on a χ^2, $p < 0.01$). Not surprisingly therefore there is no correlation of reaction time with frequency ($r = -0.01$ for the marked and $r = -0.10$ for the unmarked). However there is a strong correlation between the reaction time for the unmarked terms and their corresponding marked terms : $\rho = 0.72$, $p < 0.001$. The frequency of marked terms correlates with the frequency of the corresponding unmarked : $\rho = 0.48$, $p < 0.01$.

2. The Second Premise.

These reaction times were analysed by means of a four factor analysis of variance with marking of the first premise, marking of the second premise and position of the new item in the second premise (theme or rheme) as within subjects factors and pronoun/name as a between subjects factor. The full ANOVA results are given in Table 9 and a list of significant effects

Table [illegible] : Three Term Series Problem Experiment : Means in msec.

Premise Combinations.[a]

2nd. Premise Results.		A > B C > A	A < B C > B	A > B C < B	A < B C < A	A > B B > C	A < B A > C	A > B A < C	A < B B < C
Pronouns:		3087	4433	4001	3436	4008	4680	4329	4631
Names:		3634	4384	3924	3736	3571	4484	4095	4048
Question Results.	Question								
Pronouns:	unmarked	1794	2492	2344	2682	2913	3140	2714	2595
	marked	2185	2410	2531	2205	2179	2505	3207	2634
Names:	unmarked	1985	2059	2316	2586	2427	2955	2712	2608
	marked	2086	2747	2437	2032	1956	2360	3227	2368
Total Times.									
Pronouns:	unmarked	8278	10853	9823	9878	10072	11251	10882	12053
	marked	8947	10249	10353	9173	9895	10233	11071	10750
Names:	unmarked	8902	11314	9478	9496	9208	10705	10370	9767
	marked	9869	10389	9727	9576	8955	9680	11022	10330
Errors									
Pronouns:	unmarked	.041	.583	.5	.25	.416	.666	.583	1.083
	marked	.083	.333	.208	.25	.208	.5	.708	.666
Names:	unmarked	.083	.208	.583	.541	.5	.583	.666	.666
	marked	.167	.541	.041	.292	.125	.333	.75	.916

a As elsewhere '>' denotes an unmarked term, '<' a marked term.

Table 9 : Three Term Series Problem : Anovas.

	df	2nd. Premise RT	Question RT	Total RT	Errors
Subjects	47,46	-	-	-	-
A Pronoun	1,46	1.23	0.52	0.33	0.93
B Question Marking.	1,46	-	*5.54	1.00	0.34
C 1st.Prem. Marking.	1,46	***14.27	*7.22	**11.08	**10.57
D 2nd. Prem. Marking.	1,46	0.01	2.25	* 4.47	***15.51
E New Item Position.	1,46	**9.66	***16.97	***16.85	***38.29
AB	1,46	-	0.58	1.70	2.60
AC	1,46	0.15	0.99	0.30	0.02
AD	1,46	0.43	0.37	2.19	2.69
AE	1,46	*4.42	3.59	** 9.35	0.06
BC	1,46	-	* 6.42	** 7.72	0.98
BD	1,46	-	3.26	1.83	2.08
BE	1,46	-	0.54	1.47	0.57
CD	1,46	***24.69	***18.84	***20.90	***14.98
CE	1,46	0.40	* 5.31	0.72	0.96
DE	1,46	0.94	1.17	** 8.14	**12.00
ABC	1,46	-	0.20	0.46	1.21
ABD	1,46	-	0.00	2.29	0.93
ABE	1,46	-	0.58	0.61	0.88
ACD	1,46	0.21	1.60	0.12	2.56
ACE	1,46	0.02	0.03	0.31	0.24
ADE	1,46	0.01	3.24	1.06	2.52
BCD	1,46	-	1.68	0.73	0.62
BCE	1,46	-	1.12	0.35	1.38
BDE	1,46	-	**10.93	0.97	***19.94
CDE	1,46	3.68	0.02	1.80	0.46
ABCD	1,46	-	0.00	2.09	0.85
ABCE	1,46	-	0.14	0.19	1.05
ABDE	1,46	-	0.87	0.53	* 4.72
ACDE	1,46	3.91	0.11	2.69	0.07
BCDE	1,46	-	0.72	0.89	1.12
ABCDE	1,46	-	0.02	0.04	0.85

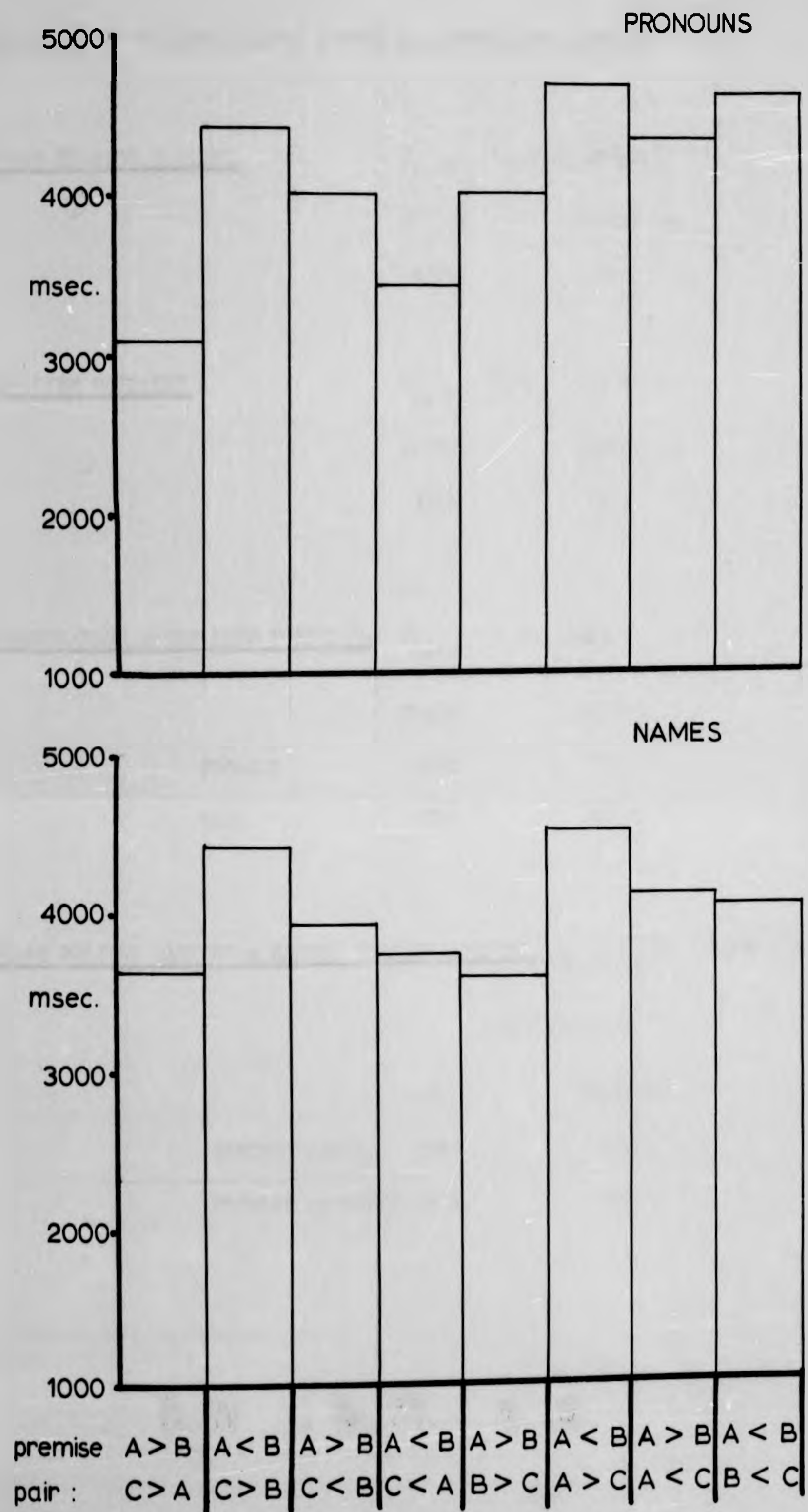


Figure 3 Experiment 5: reaction times to the second premise

Table 10 : Second Premise : List of Significant Effects. (Figures are msec.)

FIRST PREMISE MARKING. $F_{1,46} = 14.27$, $p < 0.001$

MARKED	UNMARKED
4229	3831

NEW ITEM POSITION $F_{1,46} = 9.66$, $p < 0.01$

FIRST	SECOND
3329	4230

PORNOUN/NAME x NEW ITEM POSITION. $F_{1,46} = 4.42$, $p < 0.05$

	FIRST	SECOND
PRONOUN	3739	4412
NAME	3919	4049

FIRST PREMISE MARKING x SECOND PREMISE MARKING. $F_{1,46} = 24.69$, $p < 0.001$

	FIRST PREMISE	
	MARKED	UNMARKED
SECOND MARKED	3963	4087
PREMISE UNMARKED	4495	3575

with means is given in Table 10. The following effects reached significance:-

(a) Position of the New Item. RTs are significantly shorter if the new item is theme of the second premise ($F_{1,46} = 9.66$, $p < 0.01$). This is similar to effects reported by both Huttenlocher and Clark. In the present case mean times for New Item First and New Item Second are 3829 msec. and 4230 msec., a difference of 401 msec. This result is confounded by an interaction between the new item position factor and the pronoun factor ($F_{1,46} = 4.42$, $p < 0.05$). Although there is a 673 msec. advantage of having the new item first where a pronoun is used (3739 msec. vs. 4412 msec.), with names this is reduced to only 130 msec. (3910 msec. vs. 4040 msec.)

(b) First Premise Marking. RTs are substantially lower if the first premise is unmarked : 3831 msec. vs. 4229 msec., a difference of 398 msec. ($F_{1,46} = 14.27$, $p < 0.001$). This is confounded by a considerable interaction of the first premise marking and second premise marking factors. If the second premise is marked then there is little difference between a marked first premise and an unmarked one (4087 msec. if unmarked, 3962 msec. if marked). However if the second premise is unmarked there is a considerable difference between marked and unmarked first premises, in the opposite direction (4495 msec. and 3575 msec. respectively).

Only these four results reached significance though there were two other effects almost reaching significance:-

(i) Position of the New Item x First Premise marking x Second Premise marking. $F_{1,46} = 3.68$, $p < 0.1$.

(ii) Pronoun/Name x Position of the New Item x First Premise marking x Second Premise marking. $F_{1,46} = 3.91$, $p < 0.1$.

Given the small size of these effects, as well as the fairly large number of F tests carried out, no discussion will be given of these results.

3. The Question

These reaction times were analysed by means of a five factor analysis

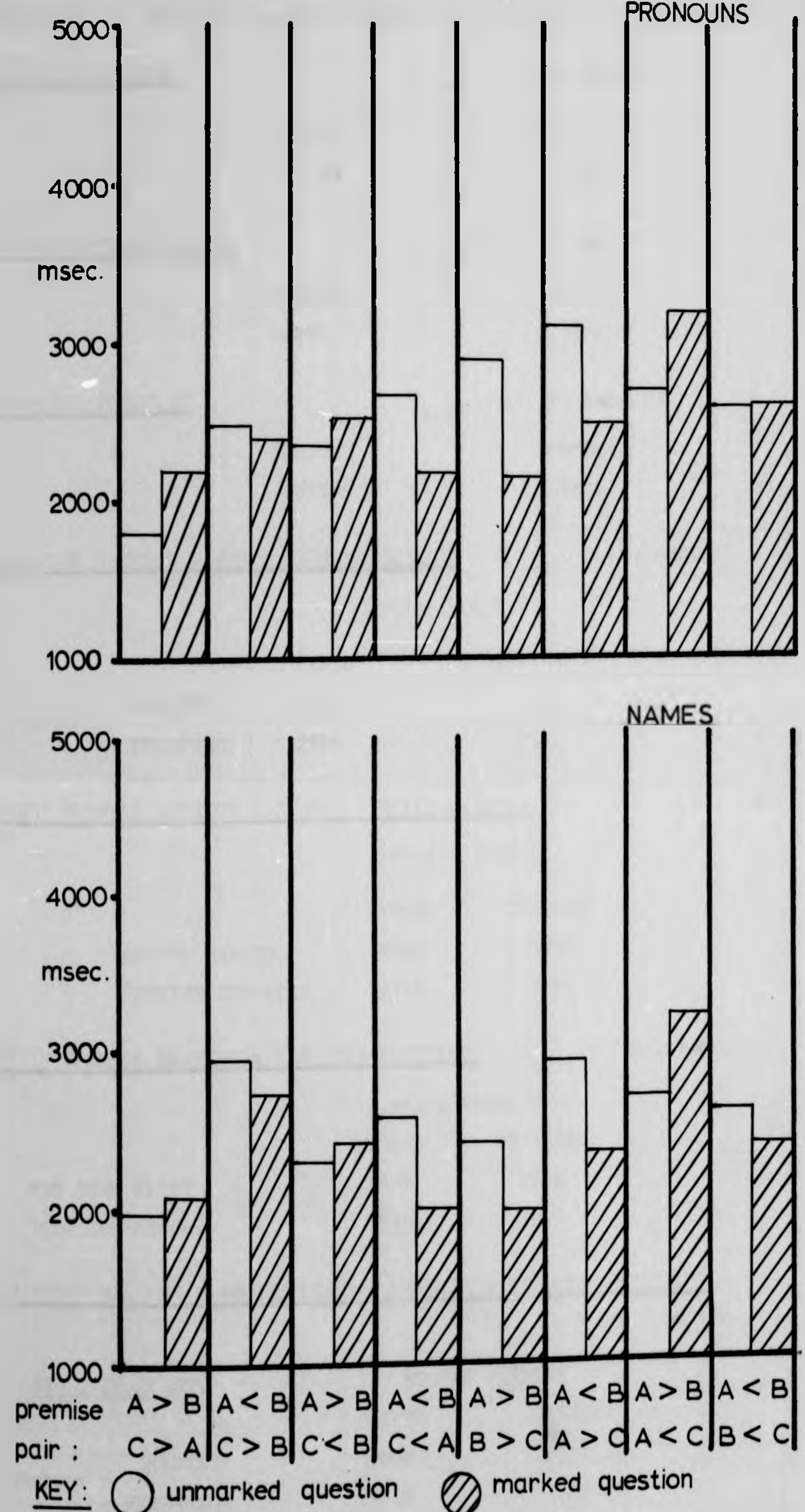


Figure 4 Experiment 5: reaction times to the question

Table 11 : Questions : List of Significant Effects. (Figures are msec.).

QUESTION MARKING $F_{1,46} = 5.54$, $p < 0.05$

MARKED	UNMARKED
2442	2593

FIRST PREMISE MARKING $F_{1,46} = 7.22$, $p < 0.05$

MARKED	UNMARKED
2507	2438

NEW ITEM POSITION $F_{1,46} = 16.97$, $p < 0.001$

FIRST	SECOND
2378	2656

QUESTION MARKING x FIRST PREMISE MARKING $F_{1,46} = 6.42$, $p < 0.05$

		FIRST PREMISE	
		MARKED	UNMARKED
QUESTION	MARKED	2417	2476
	UNMARKED	2785	2401

FIRST PREMISE MARKING x SECOND PREMISE MARKING $F_{1,46} = 18.84$, $p < 0.001$

	FIRST PREMISE	
	MARKED	UNMARKED
SECOND MARKED	2464	2686
PREMISE UNMARKED	2729	2190

FIRST PREMISE MARKING x NEW ITEM POSITION $F_{1,46} = 5.31$, $p < 0.05$

	FIRST PREMISE	
	MARKED	UNMARKED
NEW ITEM FIRST	2547	2209
POSITION SECOND	2646	2667

QUESTION MARKING x SECOND PREMISE MARKING x NEW ITEM POSITION

$F_{1,46} = 10.93$, $p < 0.01$.

NEW ITEM FIRST

		SECOND PREMISE	
		MARKED	UNMARKED
QUESTION	MARKED	2302	2357
	UNMARKED	2482	2373

NEW ITEM SECOND

		SECOND PREMISE	
		MARKED	UNMARKED
QUESTION	MARKED	2859	2250
	UNMARKED	2658	2859

of variance with question marking, marking of the first premise, marking of the second premise and position of the new item in the second premise as within subjects factors and the means of cross referring between premises as a between subjects factor. The full ANOVA results are given in Table [illegible] and a list of significant effects with means is given in Table 11. The following results reached significance:-

(a) Question Marking. Reaction times are significantly shorter if the question is marked ($F_{1,46} = 5.54$, $p < 0.05$). In fact this result appears to be partly due to an interaction between Question Marking and First Premise Marking ($F_{1,46} = 6.42$, $p < 0.05$) which seems to be due to the fact that the combination marked first premise, unmarked question takes over 300 msec. longer than any of the other three possibilities which are all quite similar.

(b) First Premise Marking. RTs are shorter when the first premise is unmarked (2597 msec. vs. 2438 msec.) ($F_{1,46} = 7.22$, $p < 0.05$). Again this result is partly due to the interaction mentioned in (a).

(c) New Item Position. RTs are very much shorter when the new item is first in the second premise ($F_{1,46} = 16.97$, $p < 0.001$). This difference is greatly reduced when the first premise is marked (only 97 msec. difference between new item first and new item second) compared to when it is unmarked (where there is a difference of 458 msec.). This gives rise to a significant interaction between First Premise Marking and New Item Position ($F_{1,46} = 5.31$, $p < 0.05$).

(d) First Premise Marking x Second Premise Marking. This is a very substantial effect ($F_{1,46} = 18.84$, $p < 0.001$). RTs are lower when both premises are marked the same (2190 msec. for both unmarked, 2464 msec. for both marked compared with 2686 msec. for unmarked/marked and 2729 msec. for marked/unmarked).

(e) Question Marking x Second Premise Marking x New Item Position. This result is significant at the $p < 0.01$ level ($F_{1,46} = 10.93$). The advantage of having the new item first is restricted chiefly to the problems where

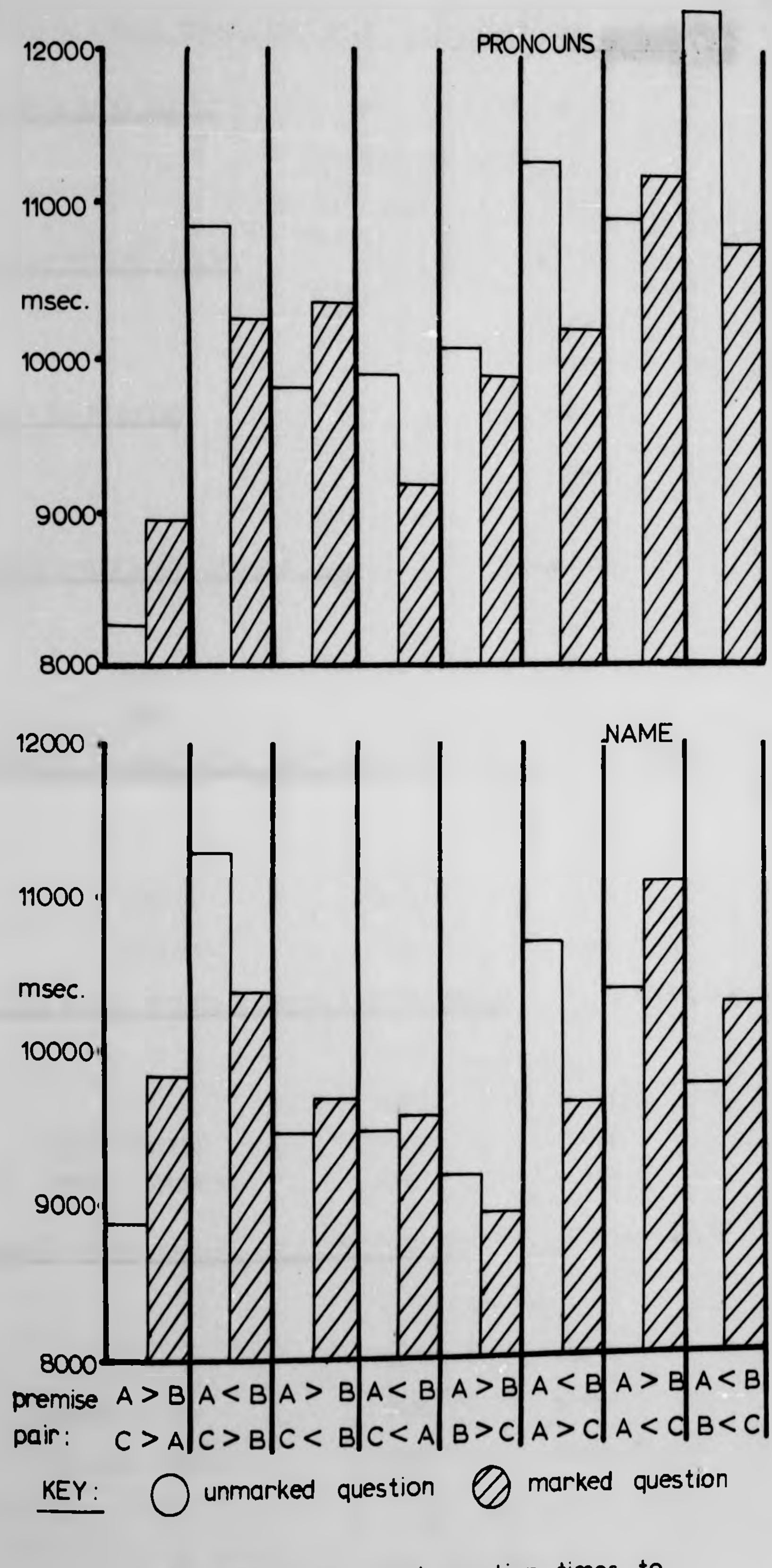


Figure 5 Experiment 5: total reaction times to each problem

Table 12 : Total Times : List of Significant Effects. (Figures are msec.)

FIRST PREMISE MARKING — $F_{1,46} = 11.08$, $p < 0.01$

MARKED	UNMARKED
10355	9803

SECOND PREMISE MARKING — $F_{1,46} = 4.47$, $p < 0.05$

MARKED	UNMARKED
10233	9925

NEW ITEM POSITION — $F_{1,46} = 16.85$, $p < 0.001$

FIRST	SECOND
9768	10390

PRONOUN/NAME x NEW ITEM POSITION — $F_{1,46} = 9.35$, $p < 0.01$

	FIRST	SECOND
PRONOUN	9669	10776
NAME	9867	10390

QUESTION MARKING x FIRST PREMISE MARKING — $F_{1,46} = 7.72$, $p < 0.01$

		FIRST PREMISE	
		MARKED	UNMARKED
QUESTION	MARKED	10047	9979
	UNMARKED	10663	9627

FIRST PREMISE MARKING x SECOND PREMISE MARKING — $F_{1,46} = 20.90$, $p < 0.001$

	FIRST PREMISE	
	MARKED	UNMARKED
SECOND PREMISE MARKED	10126	10341
SECOND PREMISE UNMARKED	10584	9266

SECOND PREMISE MARKING x NEW ITEM POSITION — $F_{1,46} = 8.14$, $p < 0.01$

	NEW ITEM	
	FIRST	SECOND
SECOND PREMISE MARKED	9686	10780
SECOND PREMISE UNMARKED	9850	10000

marking of the question and the second premise is congruent (in which case the effect is about 500 msec.). Where these are not congruent the effect seems to disappear (2419 msec. for new first, 2454 msec. for new second).[10] It is perhaps worth bearing in mind that whenever the second premise and the question are congruent and the new item is first then it is the answer. But whenever they are not congruent and the new item is first, it is not the answer. Similarly whenever the new item is second and they are congruent then the new item is not the answer but whenever the new item is second and they are not congruent then the new item is the answer. It is possible there are separate effects of position of the new item in the second premise and whether or not it is the answer.

4. Total Times.

These were analysed with the same factors as the question times. Full results are presented in Table 8 and a list of significant effects with means is given in Table 12. The following effects reached significance:-

(a) First Premise Marking. There was a mean advantage of [illegible] msec. of having the first premise unmarked. This was highly significant ($F_{1,46} = 11.08$, $p < 0.01$).

(b) Second Premise Marking. There was a mean advantage of [illegible] msec. of having the second premise unmarked ($F_{1,46} = 4.47$, $p < 0.05$).

(c) New Item Position. RTs were 622 msec. shorter on average when the new item was first ($F_{1,46} = 16.85$, $p < 0.001$).

(d) Pronoun/Name x New Item Position. The advantage of having the new item first is much greater with pronouns (1107 msec.) than with names (523 msec.) ($F_{1,46} = 9.35$, $p < 0.01$).

(e) Question marking x First Premise marking. The superiority with the first premise unmarked is much reduced if the question is marked (only 68 msec. compared with 1036 msec. when the question is unmarked).

10. This may well be an oversimplification : see the relevant table.

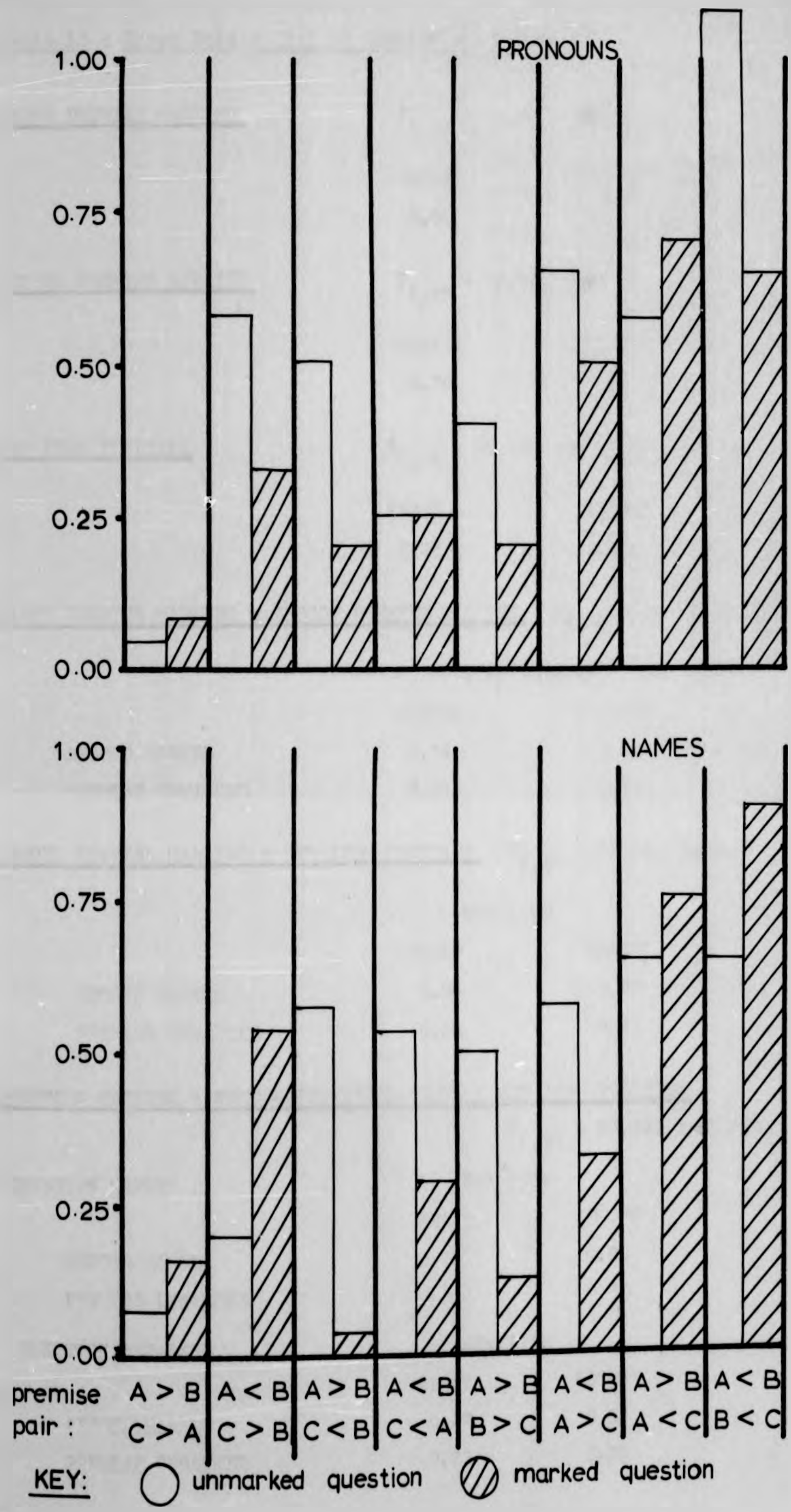


Figure 6 Experiment 5: mean number of errors for each problem type

Table 13 : Error Data : List of Significant Effects.

FIRST PREMISE MARKING $F_{1,46} = 10.57,\ p < 0.01$

MARKED	UNMARKED
0.51	0.38

SECOND PREMISE MARKING $F_{1,46} = 15.51,\ p < 0.001$

MARKED	UNMARKED
0.54	0.35

NEW ITEM POSITION $F_{1,46} = 38.29,\ p < 0.001$

FIRST	SECOND
0.32	0.57

FIRST PREMISE MARKING x SECOND PREMISE MARKING $F_{1,46} = 14.98,\ p < 0.001$

	FIRST PREMISE	
	MARKED	UNMARKED
SECOND MARKED	0.54	0.49
PREMISE UNMARKED	0.55	0.22

SECOND PREMISE MARKING x NEW ITEM POSITION $F_{1,46} = 12.00,\ p < 0.01$

	NEW ITEM	
	FIRST	SECOND
SECOND MARKED	0.36	0.72
PREMISE UNMARKED	0.29	0.42

QUESTION MARKING x SECOND PREMISE MARKING x NEW ITEM POSITION

$F_{1,46} = 19.94,\ p < 0.001$

QUESTION MARKED

	NEW ITEM	
	FIRST	SECOND
SECOND MARKED	0.26	0.83
PREMISE UNMARKED	0.34	0.30

QUESTION UNMARKED

	NEW ITEM	
	FIRST	SECOND
SECOND MARKED	0.47	0.61
PREMISE UNMARKED	0.23	0.54

Errors (cont'd.)

PRONOUN/NAME x QUESTION MARKING x SECOND PREMISE MARKING x NEW ITEM POSITION

$F_{1,48} = 4.72$, $p < 0.05$.

PRONOUN

QUESTION MARKED

	NEW ITEM FIRST	NEW ITEM SECOND
SECOND MARKED	0.23	0.69
PREMISE UNMARKED	0.21	0.35

QUESTION UNMARKED

	NEW ITEM FIRST	NEW ITEM SECOND
SECOND MARKED	0.37	0.56
PREMISE UNMARKED	0.31	0.54

NAME

QUESTION MARKED

	NEW ITEM FIRST	NEW ITEM SECOND
SECOND MARKED	0.29	0.98
PREMISE UNMARKED	0.48	0.25

QUESTION UNMARKED

	NEW ITEM FIRST	NEW ITEM SECOND
SECOND MARKED	0.56	0.67
PREMISE UNMARKED	0.15	0.54

(f) First Premise Marking x Second Premise Marking. Having both premises marked the same leads to shorter RTs - though more so if both are unmarked. This two way interaction is highly significant ($F_{1,48} = 20.90$, $p < 0.001$).

(g) Second Premise Marking x New Item Position. If the new item is first there is little difference between marked and unmarked second premises (actually 164 msec. in favour of the marked). If it is second then RTs are much shorter to the unmarked (by 780 msec.).

5. Errors.

Overall the error rate was 22% : quite high but comparable with earlier results. The errors were analysed in the same way as the question RTs. This meant that only 3 possible scores could be derived from each subject for each problem (viz. 0,1 and 2). This gives odd distributions and makes the ANOVA rather inaccurate. However I present it in Table [illegible] for the sake of completeness. It in fact gives little or no extra information, results being more or less as one would expect from a direct relationship between error rate and RT. I will therefore not go into these data any further. (A list of significant effects with means is given in Table 13).

Discussion

The First Premise

The results from the first premise times reveal no difference between marked and unmarked terms. This is in accord with Huttenlocher's model : she predicts an average of two operations for both marked and unmarked first premises. However the model itself makes no real sense if we consider it in this way for the very first test step in the model asks "Is premise end anchored?" But it is impossible to tell this until the second premise has been encoded. It is possible that, given this, subjects simply store the first premise in its surface form until the second premise is coded. However Huttenlocher's experimental technique, in which she asks subjects questions about the first premise before presenting the second premise, makes it rather unlikely that subjects would hold the first premise in surface form - at least in her experiment. Furthermore, even given the second premise, this first test either (i) assumes that the whole operation which we are here trying to model has already been carried out, since the only fully accurate method of discovering whether a premise is end-anchored or not is to figure out where all the objects go in one's imaginary display; or (ii) fails to always work if one simply uses the method of seeing whether the subject of the first sentence is referred to in the second as a means of testing for end-anchoring (if it occurs in both then the first is not end-anchored, if not, it is). (This does not work because it cannot cope with partially ordered sets). Even if one can avoid these problems with the Image model this first premise data hardly constitutes strong support for it since it is here merely predicting no difference between the two sets of adjectives.

The failure to find an effect clearly goes contrary to Clark's theory : his model predicts four operations for each lexically marked first premise but only three for premises with unmarked relational terms. The results provide no support for such an hypothesis. This may be because the items do not constitute true marked/unmarked pairs. To this one can merely restate the point made already that although there are dubious cases, unless one can find a fairly broad sample of marked/unmarked pairs, the whole distinction is rather pointless. In addition, of course, the results for the second premise and question do show markedness effects. So why not here? The answer may lie in the method used : the comparison of any marked word with its unmarked partner is a between subjects one comparing Groups 1 and 2 with Groups 3 and 4. As it happens Groups 1 and 2 are consistently faster and any effect of marking would be easily swamped by this apparently random effect. Certainly the failure to find a marking effect goes contrary to the present theory as well as Clark's and since it only provides the weakest possible confirmation of Huttenlocher's account it may not be unreasonable to blame the failure to produce an effect on the weakness of the experimental design at this point. The design was used in order to discourage strategy formation by never using any relation more than once. An experiment which risked this possibility in order to explore more fully lexical marking effects in the first premise would seem to be necessary though.

The Second Premise

Table 14 gives some selected predictions from both the Clark and Huttenlocher models, as formulated by Johnson-Laird, in terms of the four way analysis of variance performed on the second premise times. Both models correctly predict that RTs would be shorter when the new item is theme in the second premise. Both theories predict a complex three way interaction involving new item position, first premise marking and second premise marking, but their predictions are rather different. This interaction, as we have seen, just fails to reach significance. Clark's model comes much nearer to predicting the results than Huttenlocher's predicting the direction of the effects when the new item is first quite accurately though incorrectly predicting a simple deficit when the second premise is marked and the new item second. All in all the results look more complex than the model predicts. In particular, it looks as though

Table 14 : Predictions from Johnson-Laird's versions of the Linguistic and Image models for the second premise times.

Figures are hypothesised numbers of operations - totals, not means .

First Premise Marking

		Marked	Unmarked
Model	Linguistic	24	24
	Image	34	34

New Item Position

		First	Second
Model	Linguistic	22	26
	Image	32	36

Second Premise Marking

		Marked	Unmarked
Model	Linguistic	26	22
	Image	34	34

First Premise Marking x Second Premise Marking

	First Premise :	Marked		Unmarked	
	Second Premise :	Marked	Unmarked	Marked	Unmarked
Model	Linguistic	12	12	14	10
	Image	17	17	17	17

First Premise Marking x New Item Position x Second Premise Marking

	First Premise :	Marked		Unmarked	
	Second Premise :	Marked	Unmarked	Marked	Unmarked
New First	Linguistic	5	6	7	4
	Image	8	8	7	9
New Second	Linguistic	7	6	7	6
	Image	9	9	10	8

the use of a pronoun or name to crossrefer between premises interacts with all these three other factors (although the result is only significant at the $p < 0.1$ level). As the two models are formulated by Johnson-Laird they operate in precisely the same fashion whether reference is by means of names or pronouns so that this result provides counter-evidence to both models as formulated by him. However I will not discuss this result further since it is so small and probably unreliable, save to note that if it were to prove reliable both models would need extending to distinguish between these two methods of cross-referring.

Only Clark's model predicts a substantial interaction between the first and second premise marking factors (the image model predicting no effect due to either factor as well as no interaction - see Table 14). His model predicts that both premises unmarked ought to be easier than both marked as well as first marked, second unmarked (which ought to be about the same as one another), while these two ought to be easier than first unmarked, second marked. In fact the first of these averages about 400 msec. faster than the second and fourth (which are about equal) which are in turn about 400 msec. faster than the third. In other words instead of the first unmarked, second marked case being hardest the first marked, second unmarked case is hardest. This is clearly quite a difference.

Along with this interaction there is a significant main effect of first premise marking with those problems having a marked first premise taking significantly longer than those having an unmarked first premise. The predictions I have derived from Johnson-Laird's formulations of the linguistic and image models state that there should be no simple effect of this factor. Both of them therefore prove inadequate on this count. As already noted it is not possible to predict from either model any effects of varying the method of cross-referring - effects which are clearly present in these data. They are clearly in need of modification if they are to cope with this.

If we concentrate on the process of integration of the two premises and the reasons for using the various linguistic devices which are available then we can begin to make more sense of the data. In the introduction to this chapter we stressed the importance for the subject of finding where the third item goes - in Huttenlocher's terminology "solving for" that item. Because the subject is focussing his attention on the new item it helps him if this is thematic - but only, we said, if he can easily see that the usual distribution of new and old information in the sentence does not hold. Pronouns make identification of new and old information much easier - particularly with sequentially presented sentences where it is not possible to look back and match the names against those in the previous sentence. Accordingly we would expect to find the advantage of having the new item first to be much greater when pronouns are used to cross-refer. This is clearly true in the present data : the advantage is only 130 msec. with names, but 673 msec. with pronouns.

A second factor is whether the pronoun in the second premise refers to the subject or object of the first premise. On the basis of the oral description of pictures study it seems likely that when the pronoun refers to the same object as the subject of the previous premise RTs will be shorter than when it refers to the same thing as the object (since the latter is a much less frequent collocation). Though this is not tested explicitly in the analysis it appears to be confirmed by the data : when the pronoun is coreferential with the subject noun phrase of the first premise mean RT is 3883 msec. compared with 4361 msec. when it is coreferential with the object noun phrase (a difference of 478 msec.). The corresponding figures for the name data are 3938 msec. and 3985 msec. This is a rather convincing difference.

The present theory also leads one to predict a simple effect of first premise marking since this cannot be (or rather is unlikely to be) due to a topicalisation choice and is most likely a way of conveying information

Table 15 Selected predictions from Clark's question answering model as formulated by Johnson-Laird.

[Figures are total number of operations].

First Premise Marking

Marked	Unmarked
30	30

Question Marking

Marked	Unmarked
30	30

New Item Position

First	Second
26	34

First Premise Marking x Second Premise Marking

		First Premise	
		Marked	Unmarked
Second	Marked	14	16
Premise	Unmarked	16	14

First Premise Marking x Question Marking

		First Premise	
		Marked	Unmarked
Question	Marked	14	16
	Unmarked	16	14

First Premise Marking x New Item Position

		First Premise	
		Marked	Unmarked
New	First	13	13
Item	Second	17	17

First Premise Marking x Question Marking x New Item Position

New Item First

		First Premise	
		Marked	Unmarked
Question	Marked	6	7
	Unmarked	7	6

New Item Second

		First Premise	
		Marked	Unmarked
Question	Marked	8	9
	Unmarked	9	8

about the absolute positions of the various objects on the relevant dimension. Choice of the marked or unmarked adjective in the second premise is assumed always to be thematic, so that there should be no simple effect of second premise marking. Both these predictions are opposite to those of the linguistic model and both are supported (though obviously support for the latter is rather weak). The present "model" does not enable one to predict the interaction which is seen to occur between first premise marking and second premise marking, but this is not clear evidence against it. If one assumes that having the two premises marked differently leads to longer RTs then this, together with the assumption that first premise marking dominates coding and coding takes longer if in the marked form, gives the observed results. Only the first of these two assumptions is post hoc.

The Question

The reaction time results from the question are remarkably similar to those from the second premise. All of the effects which proved significant in the analysis of the second premise data were again significant in the question data with the sole exception of the interaction between the new item position and the method of crossreferring which falls short of significance. The same effect is still manifest though : the superiority with the new item first is still much greater with pronouns (413 msec.) than with names (184 msec.). The only effect present in the question times which could have been present in the second premise comprehension times but was not, is a tendency for problems with the first premise unmarked to be solved more quickly than those with it marked, but only if the new item is first in the second premise. I can think of no explanation for this result and am inclined to question its reliability (this result incidentally would not be predicted on the basis of Clark's model - see Table 15).

There are three significant results in the question times which could not have been present in the analysis of the second premise times :

a tendency for the marked question to be responded to quicker than the unmarked; a tendency for RTs to be longer if the first premise is marked and the question unmarked (all the other three combinations of these two factors yield roughly similar times); thirdly, a very complex interaction between second premise marking, question marking and new item position. When the new item is first in the second premise there seems to be little effect of question or second premise marking. When it is second RTs are noticeably longer when the second premise and question are similarly marked than when they are differently marked with the second premise unmarked, question marked case being noticeably short.

Huttenlocher's model of course provides no account of question answering. Clark's model is much more explicit and allows us to compare the number of operations which it indicates should be performed, with RTs. Table 10 gives the total number of operations performed for each of the results which is significant on the analysis of variance. The model gives correct predictions for the effect of new item position in the second premise and also for the interaction between first premise marking and second premise marking (longer RTs if they are marked differently). All the other predictions are clearly falsified. The model predicts no main effect of first premise marking or question marking but in fact marked questions and unmarked first premises lead to shorter RTs. The model predicts an interaction between question marking and first premise marking similar to that between first premise marking and second premise marking. In fact, as already noted, only the question unmarked, first premise marked case takes longer. There is an interaction between first premise marking and new item position not predicted by the model (as we have already seen) and finally the pattern of results defined by the new item first premise marking and question marking factors with the new item second in the second premise is very different indeed from the prediction.

As repeatedly stressed, the approach which I am advocating does not attempt to give a complete model of the question answering process. The three term series problem is viewed as a "test-bed" for a more general model and no attempt is made to provide a complete model for this particular task. We are interested in how some fairly general properties of language work in this problem - not so much in the problem itself. Despite this some suggestions were made as to what one might expect from the question times. In particular the greater ease of integration when the new item is thematic in the second premise was expected to lead to a more easily addressable representation and so to faster reaction times to the question. This is clearly the case from the presence of a strong new item position effect. A second prediction was that less marking effects would be evident when the new item is first in the second premise because once the integrated representation is stored it is as easily addressable by the marked as the unmarked question. This prediction is partly borne out by the fact that there is a 204 msec. difference between the marked and unmarked question when the new item is second in the second premise but only a 99 msec. difference between the two question forms when the new item is first. However this is only weak confirmation. The final prediction was that responses would be appreciably quicker if the new item was the answer as this is the item at the centre of the subjects' attention. In addition this effect was predicted to be greater when the new item is second in the second premise as subjects would then have difficulty integrating it into the representation formed for the first premise and might still be focussing on its position while processing the question. Both these predictions were supported : subjects respond faster when the new item is the answer, and this effect is substantially larger when the new item is second in the second premise (it is 115 msec. with the new item first and 405 msec. with the new item second). The remaining effects in the data were not predicted and are not easily explained on the basis of the present model.

The data on total reaction times as well as the error data alter the overall picture so little that in the interests of economy I will avoid discussion of them. Suffice it to say that all results significant on both the other analyses (i.e. second premise times and question times) are significant on both of these and the two odd results of the question time analysis (namely superiority with the marked question and the peculiar first premise marking x new item position interaction) are present in neither. In addition there is an effect present in both which is not present in either of the two earlier analyses. This is a tendency for second premise marking to make no difference to RT if the new item is first, but for the marked case to be much longer if the new item is second. This is interpretable in terms of the good reason topicalisation principle : if there is an obvious reason for having a particular item as theme then disregard lexical marking, if not then take it into consideration.

Conclusions

These data provide very little support for the image model. This is despite the fact that the content of the premises is, for the most part, easily imagined and quite "concrete". The use of a wide range of different terms ought to have prevented strategy formation to a certain extent. Also the number of trials used is far from great compared with some earlier experiments (see Introduction to this experiment). Both these facts ought to lead to data relatively amenable to treatment by the image model, according to Johnson-Laird. They clearly do not. Only the importance of the position of the new item in the second premise provides clear support for the image model, and this also supports both alternatives. Furthermore the use of pronouns to cross-refer has a noticeable effect - and one which is not the same across all premise combinations. In particular the presence of an interaction between the method of cross-referring and the position of the new item in the second premise is evidence against Huttenlocher's notion of an analogue between perceived actor and subject : pronouns ought to make no difference.

Clark's model, on the other hand, is not entirely discredited by the present experiment though there appears to be more involved than his model can account for. It would have been possible to alter his model to account for a simple effect of the means of cross-referring : for example if pronouns had been found to produce a simple facilitation across the board this could have been attributed to easier merging of the two premises with the same subject. However this would not really be in the spirit of Clark's model. In any case there is the interaction between position of the new noun and method of crossreference, and this seems impossible to fit into Clark's model. In addition to this a good many results go clearly against the predictions of this model as formulated by Johnson-Laird.

The results generally provide support for the kind of account I have been putting forward. There is clear evidence that position in the sentence is important and that this is related to the distribution of new and old information in the sentence. The superior performance when the new item is first in the second premise is dependent, it seems, on the subjects ability to distinguish new from old information - something which is obviously greatly aided by the use of pronouns (especially in the present situation where different names are used in every problem and subjects have difficulty remembering them). There also appears to be support for the observation first made in the oral description of pictures study that the modal subject of a sentence is the noun more likely to be pronominalised in the subsequent sentence (from the fact that RTs are much longer when the pronoun refers to the same object as the second noun of the previous premise).

If we are correct in assuming that having the new item first leads to more rapid integration, then there is a little evidence to suggest that the integrated representation is neutral in that question marking makes very little difference there, but a big difference in what looks like the un-integrated case.

Finally, a note of caution. In discussing the various models I have tended to assume that their purpose is to provide a generalised account of the three term series problem in all its forms and have therefore judged them on that basis. On that basis both the image and linguistic models are plainly inadequate. But the evidence presented in this chapter could as easily be taken in another way. It may be only evidence that people behave differently in solving three-term series problem depending on the variety of adjectives, how practised they are, how "imageable" the objects described are, and so on. From this point of view (which in essence is Johnson-Laird's) the image and linguistic models are simply two strategies for different situations. The materials in the present experiment were much more varied than those in the experiments which lead to the formulation of those models. Hence neither strategy applied here and it is not surprising so little evidence was found to support the models. On the other hand, of course, I am claiming that the "model" I have given of performance in the present experiment is rather more than an account of a local strategy. It is based upon a general model of how information is structured in sentences. As it stands it is incomplete as an account of performance in three term series experiments since so little attention is paid to local strategies and undoubtedly some attention must be paid to these. It is to be hoped that the more general nature of the "model" compensates for this deficiency.

have given of performance in the present experiment is rather more than an account of a local strategy. It is based upon a general model of how information is structured in sentences. As it stands it is incomplete as an account of performance in three term series experiments since so little attention is paid to local strategies and undoubtedly some attention must be paid to these. It is to be hoped that the more general nature of the "model" compensates for this deficiency.

Chapter 5 : Three Experiments Involving Verification of Transitive Declarative Sentences

Almost all the experiments reported up to now have been concerned with the comprehension and production of copular sentences. This chapter has 2 primary aims: (1) to extend the research to cover some aspects of a rather more complicated set of sentences : those involving "transitive" verbs (2) to add yet another measurement technique to the set of those used in the experiments presented above.

To take the second point first : although error scores, comprehension latencies, judgement latencies, and error rates provide a useful battery of instruments for measuring difficulty in understanding sentences, they tend to be excessively gross. One has to make rather a lot of assumptions about what a subject is doing in the verification task for example - assumptions which are only restricted in a rather loose way by the amount of time a subject takes to respond. The problem is essentially one of introducing more 'grain' into the behaviour one observes. By this I mean we have to tie the theory to behaviour at rather more points than at present, and in order to do this we need to develop a more detailed picture of what is actually occurring - we cannot be content to fill in the holes in an ad hoc way by means of a dash of introspection here and a sprinkling of computer terminology there.

The first step in that direction was the separation of comprehension from judgement times. By increasing the sampling of the time course of the verification process a much richer picture of the possible mental processes involved was suggested. The series of experiments reported in this chapter revert to a single latency measurement of the verification process, but seek to obtain more information about the details of the way in which the subject scans the picture - or, in Margaret Donaldson's terminology, asks questions of the picture -, by recording subjects' eye movements during the verification process. The method used is not especially fine grained, but it is very simple to use and did produce some unusual data - data which did not fit at all simply into the picture which the reaction time data collected simultaneously produced.

Transitive Sentences

Transitive sentences form an altogether richer set than the simple relational sentences used in the earlier experiments (Halliday 1967-68, Quirk et al. 1972). A much larger range of systemic options is available within the set of transitive verbs; for example the voice option (active or passive) is not available within relational sentences. In fact the exact scope of transitivity phenomena is a matter of considerable debate (see Halliday 1967, Pt.2; Quirk et al., 1972; Svartvik, 1966; Anderson 1974) and there is little doubt that there are a number of different classes of transitive verbs each with its own peculiarities (Svartvik, 1966). However I don't propose to go into the linguistics involved here, simply noting that the four verbs used in the experiments reported below are comparatively clear-cut cases of transitive verbs. In any case they all allow passivisation which is the major linguistic innovation in the following experiments compared with those reported above. According to Halliday the use of the passive voice enables one to have the patient of the sentence in thematic position without at the same time having to have a marked thematic or informational structure. So for example we have corresponding to the active 1, the passive 2.

1. A boy shot the man.

2. The man was shot by a boy.

In (1) the unmarked voice option is selected, but this leads to a marked information structure with new information in the earlier part of the clause, with old information later. In (2) on the other hand the information structure is unmarked but the marked voice has been selected. Halliday suggests that voice is a method of varying theme without the excessive emphasis that would be involved with a marked theme, such as in 3 (corresponding to 2).

3. The man, a boy shot.

[On this see Halliday 1967, pp216-217]. In this respect voice performs a similar function to the selection of a marked relational term which, as we have already seen allows the theme to be varied without the necessity of the kind of marked syntax seen in the sentence scheme of 4.

4. In front of the x is a y.

In the experiments which follow no attempt is made to investigate the use of marked theme in transitive sentences as the options involved are extremely complex. Instead the experiments are kept rather simple restricting the manipulations to a comparison of context with no context, active with passive and old information = theme with new information = theme. Before considering the experiments I want to point out some of the main results in the literature and describe the set of experiments on which the present ones are based.

The Literature.[1]

Historically voice has been one of the most thoroughly investigated areas of psycholinguistics. Initial investigators attempted to test (as they thought) some of the assumptions of Chomsky's 1957 model. Perhaps the most important early experiment was that reported by Miller and M^cKean (1964). They looked at people's ability to transform sentences from active to passive (and the reverse), and negative to positive (and vice versa). Their major pertinent finding was that (when reading time had been suitably eliminated) it took more than twice as long to transform or detransform the passive than the negative. What is more there was a suggestion that the RT for the two transformations might be additive. This was a rather unsubtle experiment in that its whole structure was such as to encourage people to see the relationship between the different sentences as a question of two distinct transformations.

Savin & Perchonock (1965) using a much subtler memory technique designed to give a measure of the amount of immediate memory space taken

1 The early part of this review is heavily indebted to Greene, 1972.

up by various syntactic constructions, estimated that negatives take up more memory space than passives. It's possible that this result is in conflict with that of Miller and M[c]Kean because Savin & Perchonock's experiment might only be measuring one's ability to store a surface structure - that one need not compute the deep structure in order to perform successfully. There are 3 objections to this line of thought : (1) it seems perfectly possible to perform in Miller's experiment with a few simple rules for transforming surface structure (without being able to compute a deeper structure); (2) if only surface structure were involved questions ought to take up no more memory space than S.A.A.D.'s[a] - but in fact they do, which suggests Ss do compute some kind of deeper information (3) passives are longer than active negatives anyway.

Further evidence that negatives are harder to process than passives comes from Gough's work on sentence verification. His data, like Miller and M[c]Kean's, are chronometric, but it points in the opposite direction from theirs! There can be no doubt that Ss do have to compute a deep structure in this task (if they ever do), so that the initial temptation is to conclude that the negative transformation is harder to process than the passive. In fact though the interaction observed between truth value and whether the sentence is positive or negative strongly suggests that the psycholinguistic 'performance' version of Chomsky's 1957 model (on which the experiment was based) was incorrect in separating syntax and semantics so rigidly. This conclusion is further reinforced by Slobin's (1966) evidence that non-reversible passives (where only one of the referents of the substantives could perform the action) are easier to process than reversibles (where both could).

Despite radical modifications of the 1957 model published in 1964 (Katz and Postal) and 1965 (Chomsky), Miller eventually came to believe that the essential psycholinguistic problems were semantic rather than syntactic and that Ss' performance was very strongly influenced by task demands etc. rather than by the intrinsic difficulty of different syntactic

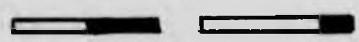
[a] simple affirmative active declarative

constructions. As Judith Greene (1972, p116) points out psycholinguistic research now went in 2 directions : (1) on the one hand some attempts were made to burrow deeper into the parameters involved in the various laboratory tasks; (2) on the other, some researchers attempted to make the laboratory tasks more naturalistic. Groups working on the laboratory problems tended to assume (in line with the initial research of Miller and his co-workers) that SAADs were easiest and ask what Ss did to the sentences with other transformations to translate them into SAADs plus additional markers. (Eventually the terminology of "canonical forms" and "operators" replaced the early Chomskyan language, but the change in substance was minimal.)

Wason and later Johnson-Laird worked on putting the sentence in context, and being within the British tradition inevitably asked the Darwinian question 'What is it good for?' If passives, negatives etc. are harder to understand why has the process of linguistic natural selection not eliminated them altogether? Wason's (1965) answer was that in some situations the negative at least is as easy to process as the affirmative. The negative, it appears, is used to deny something which might reasonably be expected, or is known to be expected or believed. It is used not, as it were, to assert "negative facts", but rather to deny that something holds in a particular case.

If there is a problem of accounting for why negatives survived (as Wason seems to have thought) then the problem is infinitely greater for passives. Negatives at least express something which could not be expressed by using only affirmatives. Passives, on the other hand, appear to be truth-conditionally redundant; a given passive would seem to be true if and only if its corresponding active is true. (As Chomsky (1957) himself pointed out this does not seem to be strictly true when quantifiers are involved, but we will set this aside.) What, then, is the use of passives, and are they always harder to understand than actives?

Johnson-Laird in his (1968a) paper on the subject, suggests that the main function of the passive is to stress the logical object. He claims that special importance is attached to first position in the English clause (a point which, as we have seen, Halliday expands considerably from a linguistic viewpoint, and one which had already been made in the psycholinguistic literature by, amongst others, John Morton (1966)). In his (1963a) paper experiment he asked people to colour in strips of paper so that other people would be able to match them to sentences e.g. "Red is preceded by blue". He found a consistent tendency to make the area corresponding to the subject of the sentence larger than that corresponding to the object. This tendency was significantly greater with passives than actives, so confirming both predictions (that first position in the sentence has a special significance, and that placing the deep object first is even more special).

Johnson-Laird followed up this study with a rather more sophisticated one in which he presented subjects with 4 sentences (2 active and 2 passive, one of each 'normal' and one 'inverted' - for explanation see below), and 2 strips of paper each coloured in 2 colours, one 50/50 and one much more of one colour than the other. For example: 

Inverted active	1. There is a blue area that a red area precedes.
Passive	2. There is a blue area that is preceded by a red area.
Active	3. There is a red area that precedes a blue area.
Inverted Passive	4. There is a red area that a blue area is preceded by.

Subjects had to either describe the 50/50 paper or the other one and they had to rate the 4 sentences in order of preference as ways of picking out the correct one of the 2 pieces of paper for someone else who they were told would read them. The results were somewhat equivocal for the case where subjects described the 50/50 paper, but for the other piece results were relatively clear cut. When the 'object' is the bigger subjects choose the passive as their first choice almost always. 2nd. choice is the inverted active, third and fourth choices are as likely to be inverted passive as normal active. When the subject is the larger area subjects tend to prefer the inverted passive, with active second choice and third and fourth choices as likely to be 'normal' passive as inverted active.

Johnson-Laird's conclusion is that voice, in itself, is unimportant; the thing which matters is word order : subjects tend to make the bigger area correspond to the first noun in the sentence. There are two criticisms of this. The first is general and applies to Johnson-Laird's first experiment too : the situation is exceedingly artificial. Although Johnson-Laird is trying to show how the passive can be contextualised, he has only succeeded in doing this by producing a situation which is unique to this experiment - and unique not just in detail but in overall structure. Who knows what relevance this experiment has to natural linguistic usage? The second criticism is specific to Johnson-Laird's second experiment : Johnson-Laird claims that only word order is important. This is obviously false : both with the subject and object corresponding to the larger area people choose the passive for preference : in one case with the more highly marked inverted active as second choice, in the other with the less highly marked normal active. I have no explanation for why this is the case, but it is so (as Johnson-Laird himself points out) and certainly indicates that more than word-order is involved.

Two further points about these experiments - the second supplementary to the first :

(1) the constructions used here are rather special in their transformational complexity since they are in fact complex sentences - not simple actives and passives. Since two clauses are involved, on a Hallidayan account there are also two themes. This makes any explanation exceedingly difficult. Johnson-Laird clearly thinks of theme in terms of position in the matrix sentence, whereas systemic accounts usually define it in terms of position in the clause.

(2) the so-called 'inverted' sentences are extremely complex in other respects too. In particular their information structure in the embedded clause is unusual in that it has new information as theme. The moral of

these last two criticisms is that it would be as well to investigate voice in simple sentences before attempting to cope with the quagmire of multiple-clause sentences.

Although the fine detail of Johnson-Laird's data is rather difficult to disentangle several points do emerge from it. Firstly it is clear that the passive is concerned with emphasising the "logical object" of a sentence, and that in some situations people do indeed want to emphasise this element. Secondly it seems to follow from this that actives and passives are not synonymous. (As noted above a point already recognised in 1957 by Chomsky for the special case in which quantifiers are involved - see Chomsky (1957) and Johnson-Laird (1968a).) Thirdly people do not choose the transformationally simpler way of expressing something as a matter of course. This last point suggests the possibility that the transformationally simpler version may not always be psychologically simpler.

Further support for this conclusion comes from the paper by Clark and Clark (1968) referred to in the Introduction. This deals with memory for sentences describing the temporal order of events. They show that people have a bias towards recalling these sentences with the subordinate clause second, a result which would be predicted on the basis of transformational complexity, but in addition, people seem to prefer to order the clauses so that their order of utterance corresponds to the temporal order in which they occurred. This latter result cannot be explained by a transformational account. The Clarks attempt to explain both results in terms of markedness (see Greenberg, 1966) suggesting that there is a tendency to recall the marked case as unmarked : the marked cases being where the subordinate clause is first (as most linguists would agree) and the prior event second (an innovation of their own). Their data also show that the subjects accuracy in recalling the sense of a sentence is not related at all closely to transformational complexity. In addition an earlier paper of Clark's in which he presents the work he did for his master's thesis (Clark, 1965) presents evidence that people's performance in producing written sentences fits closer to a model which assumes left to right production of sentences rather than the derivation of passives from actives. The argument is not very convincing - nevertheless it's true that a systemic description is consistent with evidence for left-right planning of a sentence and this is not true of a transformational description.

More support for Johnson-Laird's primary conclusion (that passives serve to emphasise the "patient" or "logical object" and actives the "actor" or "logical subject") comes from a series of experiments by Turner and Rommetveit (1967a,b, 1968). In the experiment reported in 1968 they used a memory technique presenting active and passive sentences to children for later recall. (There were five groups of children aged approximately 4,5, 6,7 and 8 years.) Subjects were presented with a picture of (i) the whole scene described in the sentence, (ii) the actor only, or (iii) the patient only, at both first presentation and recall, all combinations of pictures being used. Sentences were both reversible and non-reversible. A very large number of effects was found, but those of most interest were a significant tendency to recall non-reversibles more accurately than reversibles, a significant bias towards producing actives, a much larger effect of retrieval picture than presentation picture; finally the main effect of relevance here was a retrieval picture x sentence voice interaction. This showed that with passives the presentation of a picture of the patient significantly improved recall over presentation of either the total picture or the actor only; with actives presentation of both the total picture and the actor only significantly improved performance over presentation of the patient. Johnson-Laird's conclusion is borne out by the errors too : 72% of the incorrect responses had as the subject of the sentence the object depicted. One should be wary of seeing these results as straightforward confirmation of Johnson-Laird's conclusion, though. The fact that presentation of the total picture led to production of actives, rather than passives, shows the versatility of the active. It can be used in order to stress the first noun, but its use does not necessarily imply this : it may also be neutral. The passive on the other hand is not neutral. (This, of course, is almost the defining characteristic of markedness - see Greenberg 1966.) This may explain why the subjects in one of Johnson-Laird's experiments (1968b) tended to choose the passive when they wanted to single out one section of the picture : it is less equivocal.

The papers discussed up until now focus on the <u>nature</u> of the passive - when it tends to be used, what it is good for, and so on. These contrast with the work reported in a number of other papers, where the processing of sentences in laboratory situations has been the main focus of interest. Inevitably some work lies on the border : attempting to assess how passives are processed or even simply whether they are harder to process, while at the same time providing some kind of assessment of contextual effects. One such paper is by Tannenbaum & Williams (1968). They looked at the speed with which 11 year old children could generate active and passive sentences to describe a picture. Pictures were preceded by a preamble of 6 sentences describing either the actor or patient. They found an interaction between the element described and voice : if the patient was described then passives were speeded, if the actor then actives. This is the kind of result one might expect having seen the papers reported in the last section. However they also found a main effect of voice : passives, it appeared, are intrinsically harder. Even with patient-focus they take as long as actives. In a way we are back with this result to the early work of Miller and his associates - one is tempted to look again at transformational complexity. However there are several alternative explanations. Firstly it is possible these subjects had not adequately mastered the passive. This seems unlikely in view of the Turner and Rommetveit study with 4-8 year olds. Secondly as noted in Chapter One it is possible there is a general response bias towards actives. This would be supported by the greater probability of producing actives in the Turner and Rommetveit study. However this seems an unduly vague formulation and fails to capture the structure of the observations. Thirdly, it seems possible the preambles used were not as suitable as they might have been to focus the subject's attention on either actor or patient - if this were so, as seen above, one would expect the neutral active. There is good reason to think they might not have been suitable : the pictures were as pictures must be (cf. Berkeley) of particular objects (e.g. a car being hit by a train), whereas the preambles were very general (e.g. about the opening up of America with the advent of the railways). Many experiments show that priming the class facilitates recognition and processing of its member (e.g. Collins and Quillian, 1969) but it seems likely that the psychological break from such broad generalities to such narrow specificities is too severe to really facilitate the passive to any great extent.

A paper by Olson and Filby lies more clearly in the process-oriented group. They attempted to examine subject's reaction times for processing passives in a verification experiment. This was combined with an attempt to focus the subject's attention on either the actor or the patient by means of various different methods - in the main experiment (Experiment 3) by presenting a picture of either actor or patient and then following this with a picture of the overall action prior to presentation of the sentence. Of the many results those of primary interest here were tendencies for actives to be responded to faster than passives, and true sentences faster than false and a tendency for passives to be facilitated if the picture sequence foregrounded the patient, and actives to be facilitated if it foregrounded the actor. This last result was also affected by truth value though : a mild tendency to the reverse of this result being found with false sentences.

The Olson and Filby explanation assumes that picture-codings can be "voiced". Quite what this means is unclear but they appear to believe that picture representations are held in a form very similar to the deep structures of sentences, though at the same time they seem to believe that these deep structures are similar to surface structures in, for example, noun phrase orderings. As noted in Chapter One Clark holds a similar view, with the exception that he tends to the view that pictures will always be coded "in the active voice", as it were, if the sentence succeeds rather than precedes the picture. This difference leads Olson and Filby to rather different predictions from either Clark or Trabasso. Firstly they predict that passives will tend to take slightly longer for three reasons: (1) passives take longer to read (this is supported by Garrod and Trabasso, 1973. We will see in the third experiment presented here that they also take longer to say), (2) they believe it may take longer to produce a passive coding since such codings are developmentally late (they cite Beilin, 1960, a paper I have not been able to trace), (3) passives are less frequent. Secondly Olson and Filby predict that in some conditions passives will be at least as easy to understand as actives, despite their overall greater difficulty. This prediction is confirmed and is not surprising given, for example, Johnson-Laird's results. Olson and Filby draw three conclusions from this result : (i) comprehension of the passive does not necessarily require transformation into the active, (ii) passives are not invariably more difficult to comprehend than actives, (iii) the short term memory code appears to retain surface structure word order. The first two of these conclusions are supported by their results and fit in with evidence reviewed above. The third seems to go beyond their data in that they do not present two surface structures for comparison. It is, however, supported by the results of experiments by Gough (1966) and Wright (1969) reviewed in Chapter One. Gough found traditional transformation effects even after a delay and Wright found that people answered questions faster if the sentence about which the question is asked is in the same voice as the question. As noted in Chapter One these results tend to argue against reduction to a canonical form, and hence against the 1957 version of Chomsky's theory where the passive is treated as an <u>optional</u>

transformation on the kernel. It does not argue against the 1965 version of Chomsky's theory in which voice is an obligatory transformation marked in the base and not added to the "kernel". It does not argue either against SG in which no distinction is made between deep and surface ordering.

Greene (1972) has attempted to apply the models of Clark and Chase (1972) and Trabasso, Rollins and Shaughnessy (1972) to passives. She appears to believe that, in general, the "true"model works best for actives, but the conversion model best for passives (p.132-133). Greene shows that only by translating the passive into an "active" base string with an additional passive affix can the true model be made to work for passives. But the evidence noted above strongly points to the opposite conclusion : namely that subjects do not typically recode passives into a canonical form, although they may sometimes use a conversion strategy in which they will only have to solve "binary" problems (problems in which $\emptyset$ (x,y) is false only if $\emptyset$ (y,x) is true).

Neither the Clark and Chase (1972) nor the Trabasso et al. (1972) papers concern themselves directly with the passive; Glucksberg, Trabasso and Wald (1973) do, though. They suggest that a sentence may not be encoded as a single structure but as a series of unordered propositions (cf. Clark's model of the three-term series problem). For example "x $\emptyset$ y" might be stored as (x)

(y)

$\emptyset$ (x,y)

They seem, in fact to oscillate between this position and one based on Fillmore's case grammar. In their experiments they use both presentation orders (sentence then picture and picture then sentence) and assume that Ss need not encode all the picture information in the picture second case since they can selectively scan it. They also assume that Ss must encode all the picture information in the picture first case, as they do not know what is relevant until they hear the sentence.[2] They so arrange the

2 I confess to being totally mystified (as was Donaldson, 1974, p.7) by the

(Footnote continued from previous page)

idea that people could code all the information in a picture. It seems to derive, as Donaldson points out, from a linguistic view of vision which seems to me incoherent.

sentences that only the subject, only the object, or only the verb differs from the scene in the picture. Sentences were both reversible and non-reversible. Subjects received either sentence or picture and when they were ready pressed a button to bring on the other slide. The first time is taken to be an encoding time, the second a verification time. No effects were found of sentence voice or reversibility on encoding of the sentence. This contrasts with most studies and is interesting in the light of the study of relational sentences in Chapter Two which shows that it most certainly is possible to obtain effects in encoding times (but the RTs in that experiment were about three times as long as those here). Glucksberg et al. did find longer RTs for picture encoding when the situation is reversible. They suggest that Slobin's (1966) experiment might have been tapping picture coding as well as sentence processing. On the other hand Slobin used actual sentences, Glucksberg et al. did not (their sentences being, for example, "truck passes car"), and it is not clear what effect this might have.

The analysis of verification times reveals that for false sentences (sentence → picture order) times were shortest if the verb mismatched, next shortest if the actor mismatched and longest if the patient mismatched. Actives and passives did not differ. With the picture → sentence case there were no effects with actives, whereas with passives mismatches in the patient took longer than in either the actor or verb. This result may relate to the fact that full passives are relatively uncommon in English (outside the laboratory!) - a point noted by Clark (1965) in his study of transitional probabilities. In that study he points out that the actor constrains the patient and verb relatively little in passives, while they constrain it a lot - an asymmetry not found in actives.

Glucksberg et al. conclude that comparison operations appear consistent with a serial, self-terminating search in the order verb, actor, patient. Neither a Chomsky-style deep structure nor surface structure model fits their data. Before commenting on this conclusion I want to consider the results for true sentences. With these sentences reversible sentences were reacted to more slowly than non-reversible and passives more slowly than actives in both the sentence → picture and the picture → sentence conditions. The voice effect was much smaller with non-reversibles in the sentence → picture condition. These results are explained as follows. Taking first the sentence → picture case. Having found that all three elements match one then checks to see if one noun is "potent" (Agent, Instrument or Force in Fillmore's case grammar) and one not. If so then processing terminates. If not then S has to decide, using voice information which noun is agent. Since nouns are stored in surface order he has to check twice against the picture in the passive. Therefore it is only with reversibles that passives take longer. The procedures are very different for the picture first condition. The first noun in the sentence is compared with the first noun in the picture representation, then with the second noun in the picture representation if the first matches (otherwise the sentence cannot be true and the test exits "false"), and it is only compared with the noun not already matched. If it matches then potency is tested for. The upshot of all this is that mismatches take one extra operation if there is a patient mismatch in the passive. For true sentences passives always have one extra operation because the agent is not the first noun in the sentence representation and pictures are always coded "in the active". (This result would not occur, as they acknowledge, if one could induce passive coding of the picture)

I have gone into this model in great depth because it takes into account both the work of Clark (on information processing models) and that of Olson and Filby, and in the opinion of Glucksberg et al. taken together with these other papers "provides an information processing account of all the major 'linguistic' variables studied in sentence verification tasks as tests of the 1957 version of Chomsky's transformational grammar" (p365). A number of points about the model require emphasis : (1) the picture first and sentence first models are very different; (2) the model predicts a combination of deep and surface structure effects; (3) the authors are very firm in stressing the importance of task demands in influencing subject's strategies and would not wish to claim great generality for their model alone. Their view seems to be that task demands are so influential that it is impossible to talk in terms of a "natural language mode". This view appears to me to be excessively pessimistic and unjustified when one considers that (i) only Olson and Filby of the authors mentioned above used actual sentences, (ii) no one has used aural presentation of sentences and (iii) subjects in the Trabasso and Clark studies were all very highly practised indeed; (iv) the model as it stands cannot deal with several kinds of problems. Two examples will suffice. Firstly if one noun is "potent" and one not then if the model gets to the point of testing for potency it will automatically return "true". Hence it does not distinguish "The car hit the telegraph pole" from "The telegraph pole hit the car". Secondly it cannot cope with sentences which are false in a binary fashion. It in fact always treats these as true - assuming false passives to be true actives and false actives to be true passives; (v) it is not at all clear quite how one is to use the Olson and Filby, Clark and Glucksberg et al. models in order to obtain the generality claimed in their remark quoted above. Presumably Glucksberg et al. would wish to state that the models are simply different strategies used in different experimental situations. Certainly it is true that they cannot all handle all the various kinds of problem (e.g. the inadequacy of the Glucksberg model in the binary situation). This would not matter if there were some common thread running through them all which made them all hang together in some way. But the only similarity of this sort is the additive processing stage assumption, which is surely too weak a commonality. Clark attempts in his work to examine what he suggests might be fundamental processes. Glucksberg and Trabasso, in adding a much needed note of caution, have gone too far in the other direction.

Rationale of the Present Series of Experiments.

The experiments shortly to be described are intended to yield results of greater generality than Glucksberg et al.'s. In this respect they fall much more clearly within the tradition represented by, for example, Johnson-Laird's work, than that represented by Trabasso's. At the same time the explanations given, though heavily reliant, as with all the work in this thesis, on a particular form of linguistic description (Systemic grammar) are nevertheless still quite process-oriented. Furthermore it will become clear that, as Glucksberg et al. emphasise, the nature of the task does have a considerable influence on the relative difficulty of the various sentence-schemata.

The research presented here gets its immediate impetus from a series of experiments reported by Hall (1975). She used a simple method of recording eye movements to assess the ability of subjects of different ages groups[3] to scan pictures in order to verify sentences with different syntactic constructions. She scored number of eye movements, "appropriate" first eye movements and pattern of eye movements (efficient, semi-efficient, and non-efficient[4]). Simultaneously she took a more standard measure of difficulty - reaction time. The latter revealed main effects of age, truth value and voice, in the expected directions (younger children, false sentences, and passive voice all take longer). In addition Hall found interactions indicating that the youngest children found both false and passive sentences to be very much harder than the other groups. The number of eye movements measure revealed the same three main effects, and a voice

3 Aged : 4½, 7½, 14½, and 25 years.

4 An appropriate first eye movement is said to occur when S looks to the spot where the object would be if the sentence were true (sometimes no object was there, sometimes a different one from the one mentioned depending on whether the sentence was false in a binary or non-binary fashion). (Cont'd.)

Footnote (continued)

Curiously Hall found no effects whatever with this measure (this result will recur in the experiments reported below). An efficient search was defined as a pattern in which only the relevant pictures were fixated. As semi-efficient search was defined as one in which the relevant pictures were fixated more than the irrelevant ones. A non-efficient search type 1 was defined as a search in which the relevant + irrelevant pictures were fixated equally. A non efficient search type 2 was a search in which irrelevant pictures were fixated more often.

x truth value interaction which indicated that true sentences are reacted to much faster than false for the active voice, but not for the passive. This result appears to contradict findings of Gough. No significant effects were found on the appropriate first eye movement measure (see below). Older subjects were found to produce more efficient search patterns, however.

To understand this experiment one needs a fairly thorough grasp of the procedure and some of the fine detail. Subjects looked into a box in which pictures of objects would be hung. E showed S the picture of what would be the actor in the sentence to be presented, and hung this on a peg in the centre of the screen. She then uttered the test sentence. The lights were then extinguished and E placed the other picture in position. The test sentence is now uttered twice more by E and as she finishes the second repetition the screen is illuminated. Reaction time and eye movements are recorded from this point. One important aspect of her verbal material needs to be borne in mind : the actor was always accompanied by the definite article, and the patient by the indefinite article. This, coupled with always having the actor in the middle, would appear to strongly topicalise the actor as opposed to the patient. On the basis of work such as Johnson-Laird's reported above, this should make the passive inappropriate, hence one would expect a voice effect. Consequently it is not surprising to find one reported, but once again one has to beware of interpreting this result as due to any intrinsic difficulty of processing the passive.

The present series of studies was designed to:-

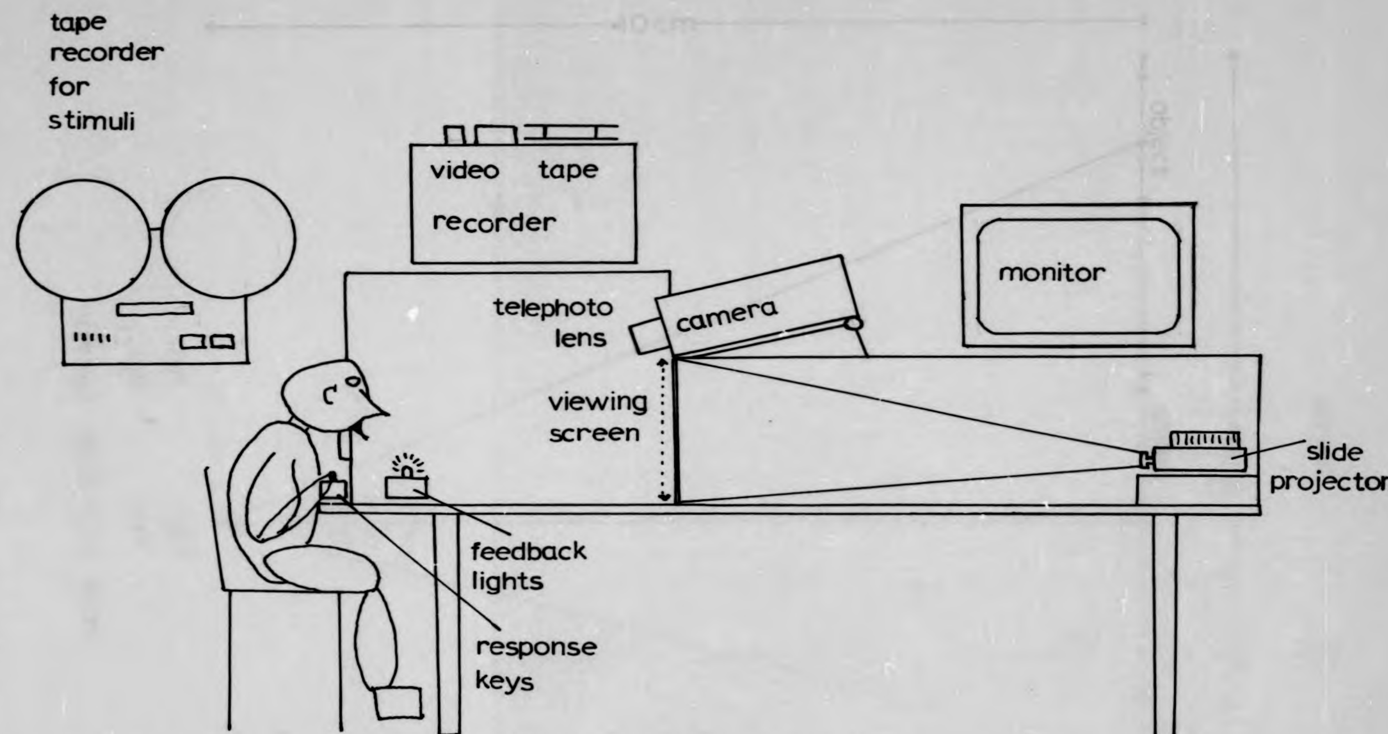
(1) remedy this last defect in Hall's study by varying the definiteness marking of the nouns in the sentence.

(2) compare the one-sentence case directly with a case in which a verbal preamble topicalising one element in the sentence is given.

(3) remedy the strong topicalisation of the actor by not showing the subject either of the objects prior to presentation of the sentence.

(4) to try to separate out the importance of voice and position in the sentence as indications of focus of interest, and to assess the suggestion that actives are neutral with regard to topicalisation.

(5) to further investigate how performance is affected by aural presentation of the verbal stimuli (something which Hall also had done).



box is totally enclosed except for viewing hole – see text

Ss eye is approx. 40cm. from the screen

(not to scale)

Figure 1 Apparatus for the eye movement experiments

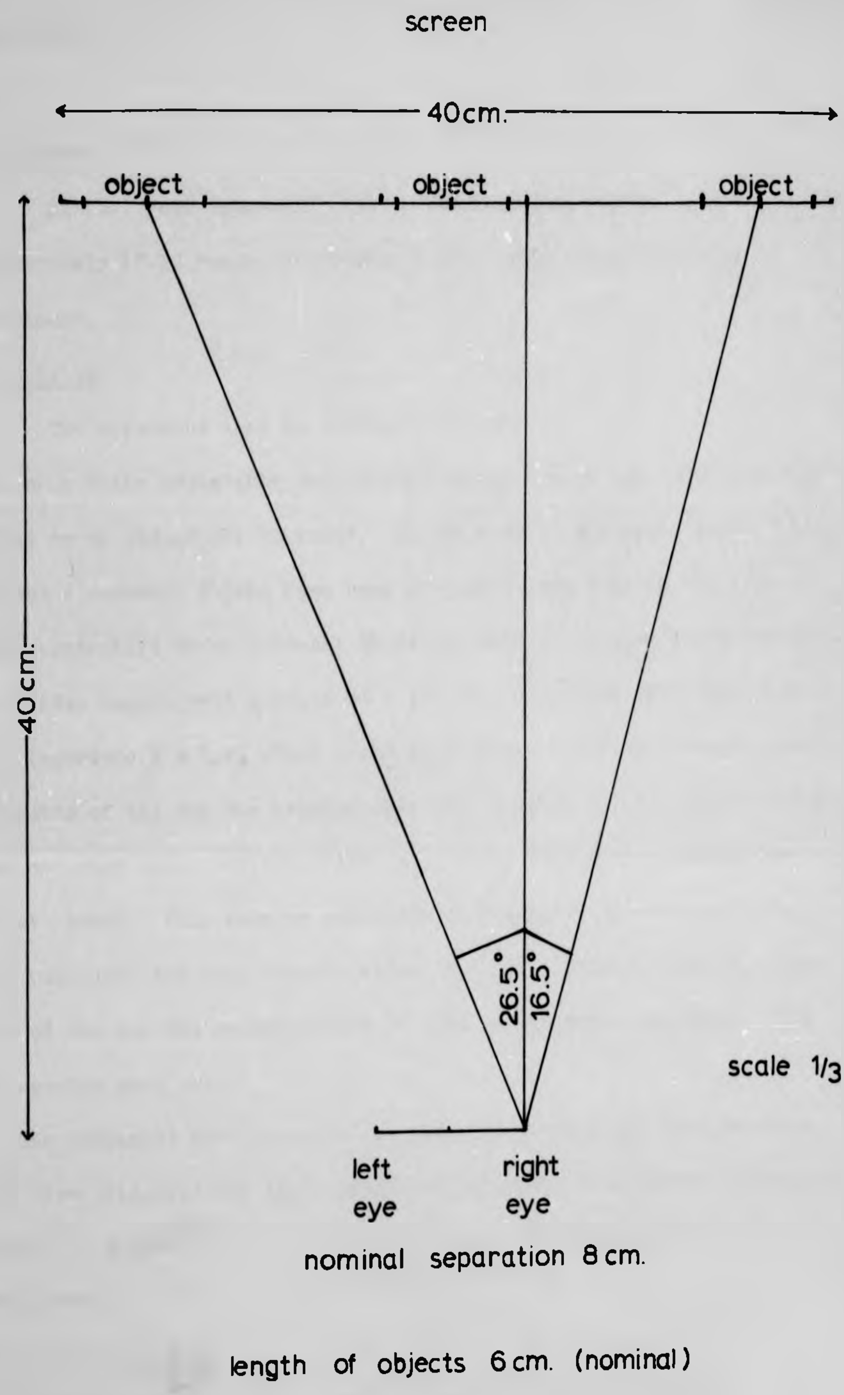


Figure 2 Plan geometry of the eye movement situation

Experiment 6

Method

1. Subjects

22 1st year undergraduates at Stirling University, aged approximately 17-18 years, 16 female, 6 male, fulfilling a course requirement.

2. Apparatus

The apparatus used is sketched in figure 1.
S sat on a chair adjustable for height looking into a box, with his chin resting on an adjustable chinrest. At the back of the box, approx. 40 cm. away was a screen. Slides were back projected onto this by means of a remote controlled Kodak Carousel Slide projector. Mounted above the screen was a video camera with a Canon 80 - 210 mm. Zoom Lens with extension tubes (aperture F = 1.8), which could be adjusted to focus on either eye. The inside of the box was painted black and blanked off so that the only light entering came from the slide (save for a little from around the subjects head.) This gave an extremely high quality picture using Sony video recorders and high density video tape and a Philips monitor. The image of the eye was approximately 7" wide on the monitor screen. This made scoring very easy.

The sentences were presented by means of a Revox A77 tape recorder which also triggered the slide projector, which in turn started the clock by means of a photocell mounted in the box. (For more details see 'procedure').

3. Materials

Slides were taken using Asahi Pentax Sla and Praktica LLC cameras in artificial light. They were of models of animals and people, manufactured by Britains Ltd. They were in colour on Kodak High Speed Ektachrome film. The background of the slides was 100% white. The objects on the slides were arranged one in the centre and the other to either the left or the right of the picture. All were near the top of the picture :

this enabled E to get a much better picture of the eye without the eyelid occluding it. They always faced the same way, namely right. Fig. 2 shows the approximate relationship between the objects and the viewer's eye. (The sketch is based on someone with a pupil separation of 8 cm.). The objects varied somewhat in both height and width but they were always so arranged that the object in the centre had its middle running through the centre line, and the objects at the periphery had the part of them nearest the centre 7 cm. from the edge of the slide. This means that the eye has to travel through a 12.5° arc to the beginning of an object on the same side as it, and through approximately 22.5° to an object on the opposite side. These figures are from the straight ahead position; moving from observing the object in the centre increases the first figure and decreases the second slightly (depending on the size of the object). In any event the deflection required was more than enough to make scoring very straightforward.

The descriptions presented on tape were similar to those used in the written presentation of the relational sentences described in Chapter 3. For example:

> In the middle of this picture is a cowboy.
>
> He has a big blue hat and a brown jacket.
>
> He wears a little moustache.

Following this the target sentence was presented.
The object in the centre was always topicalised in this way.
No attempt was made to equate the preamble sentences for length as it was feared this might encourage Ss to develop a rhythm.

The four kinds of target sentence for the above preamble set would be:

> An Indian is shooting the cowboy.
>
> The cowboy is shooting an Indian.
>
> An Indian is getting shot by the cowboy.
>
> The cowboy is getting shot by an Indian.

There were two slides for each pair so that all four sentences could this enabled E to get a much better picture of the eye without the eyelid occluding it. They always faced the same way, namely right. Fig. 2 shows the approximate relationship between the objects and the viewer's eye. (The sketch is based on someone with a pupil separation of 8 cm.). The objects varied somewhat in both height and width but they were always so arranged that the object in the centre had its middle running through the centre line, and the objects at the periphery had the part of them nearest the centre 7 cm. from the edge of the slide. This means that the eye has to travel through a 12.5° arc to the beginning of an object on the same side as it, and through approximately 22.5° to an object on the opposite side. These figures are from the straight ahead position; moving from observing the object in the centre increases the first figure and decreases the second slightly (depending on the size of the object). In any event the deflection required was more than enough to make scoring very straightforward.

The descriptions presented on tape were similar to those used in the written presentation of the relational sentences described in Chapter 3. For example:

> In the middle of this picture is a cowboy.
>
> He has a big blue hat and a brown jacket.
>
> He wears a little moustache.

Following this the target sentence was presented.
The object in the centre was always topicalised in this way.
No attempt was made to equate the preamble sentences for length as it was feared this might encourage Ss to develop a rhythm.

The four kinds of target sentence for the above preamble set would be:

> An Indian is shooting the cowboy.
>
> The cowboy is shooting an Indian.
>
> An Indian is getting shot by the cowboy.
>
> The cowboy is getting shot by an Indian.

There were two slides for each pair so that all four sentences could be true or false. The "get" form of the passive was used because it was at one time my intention to run younger subjects, and this form is more frequent in their speech than the "be" form. (All preambles and the verbs used for each type are listed in Appendix D).

Note that pragmatic expectations were fully controlled by having two pairs of each type - e.g. one pair with the cowboy as topic and one pair with the Indian as topic - see Appendix.

4. Design and Procedure

The design is a 2^4 factorial one, the four factors being : voice (active vs. passive), truth (true vs. false), theme (whether the theme, which is also always the subject, is the object mentioned in the preamble or not). The theme factor is assigned for the no context condition by correspondence to the context condition. This factor is not, however, completely a pseudo-factor as the object mentioned in the preamble is always marked with the definite article and is in the centre of the picture, while the object mentioned only in the target sentence is marked with the indefinite article and is in the periphery of the picture.

The order of object pairs was kept constant for all subjects but sentence type order was randomised separately for each subject. Context and no context trials were presented separately, half the subjects receiving the one first, and half the subjects the other. All subjects had four practice trials all active, two context and two no context, one of each true.

The order of events was as follows:-

(1) E started the experiment having first checked that S's eye was adequately placed on the monitor, and in focus, and having given the subjects instructions as detailed below.

(2) E started the tape recorder. This played the verbal stimulus (either the one sentence or the preamble + sentence) at the end of which a pulse on track 2 triggered logic which stopped the tape and changed the slide (from a black blank to the stimulus).

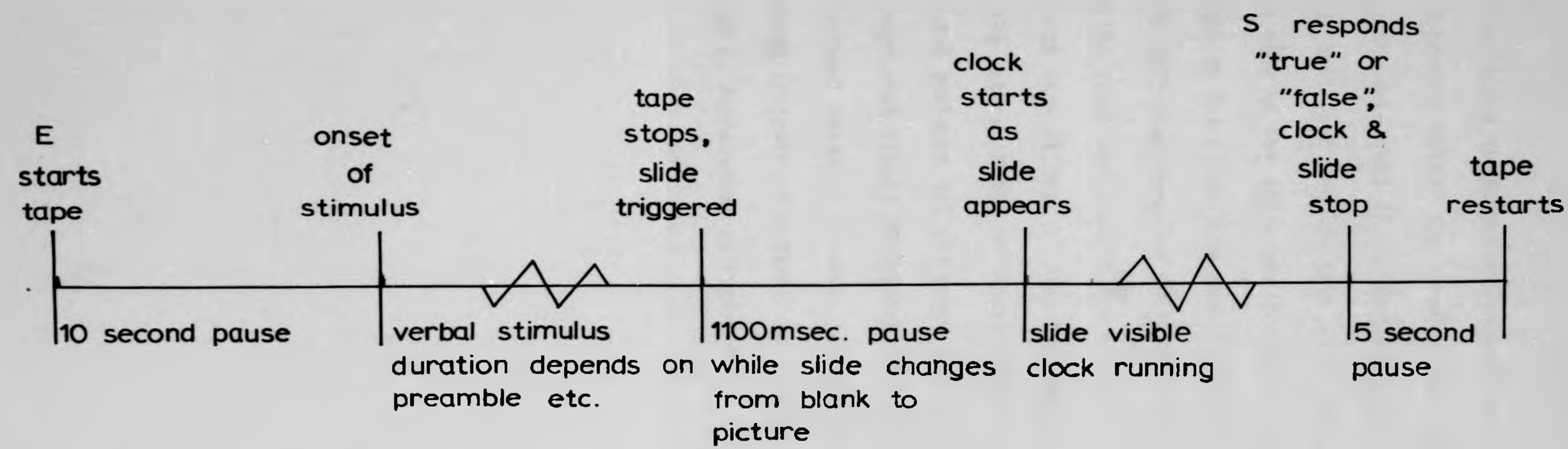


Figure 3 Time course of Experiment 6

(3) the light from the slide triggered a photocell which started a clock. This ran on until S pressed either the 'true' or 'false' keys on the control box beneath the chin rest ($\frac{1}{2}$ the Ss had 'true' on the left, $\frac{1}{2}$ on the right. Since Ss could not see the keys a red light flashed on inside the box on the same side as the false key whenever they pressed the false button, a green light on the other side whenever they pressed the true button. This was to help them remember which side was which). Pressing the button stopped the clock and changed the slide to a black blank in readiness for the next test slide. It also indicated to E which response was made. Due to the nature of the machinery there was a gap of 1100 msec. between the end of the sentence and the onset of the picture.
The tape recorder restarted itself automatically 5 secs. after the S had responded. An additional delay of 10 secs. was put on the tape between stimuli making a total inter-trial interval of 15 secs.

The time course is represented in Figure 3.

Instructions to Subjects : Experiments Six and Seven

Ss were told they would hear a series of descriptions on the tape recorder. After each description the slide described would appear in the box. They were to indicate whether the last sentence of the description was true or false using the keys in front of them. They were told half of the descriptions would be four sentences long. All sentences except the final one in each description would be true - the final one might be either true or false. They should not however ignore the others, but rather think about and try to imagine what they described as they would see that object in the picture. If the descriptions sounded childish that was because we were seeing how their performance compared with children.[5] The other half of the descriptions would only consist of the sentence which they had to verify.

They were told their eye movements would be recorded on the video and a reaction time measure would be taken from the key pressing so that it was important they should go fast, though more important that they should minimise errors. They should memorise the positions of the keys as they were not to move their head from the chinrest during the experiment.

They would have to do twenty trials : four practice and sixteen experimental. The first two would be one sentence trials, the next two four sentence ones. Similarly with the experimental trials : the first half (8) would be one sentence ones and the last half (8) would be four sentence ones.

(This last part was reversed for half the subjects.)

5 It was originally intended to extend this experiment to children but it proved both difficult to carry out and not sensitive enough.

TABLE 1 : Reaction Time Data : Experiment 6

(i) WITH PREAMBLE

	ACTIVE		PASSIVE	
THEME =	OLD	NEW	OLD	NEW
TRUE	1481 [a]	1606	1478	1783
FALSE	1665	1553	1696	1944

(ii) WITHOUT PREAMBLE

	ACTIVE		PASSIVE	
THEME = [b]	'OLD'	'NEW'	'OLD'	'NEW'
TRUE	1646	1647	1537	1514
FALSE	1706	1462	1770	1800

a : Units are milliseconds. Figures are overall means. N = 22

b : The 'Theme' factor is assigned by correspondence to condition

(i) It is effectively only a definiteness factor in condition

(ii) : theme = 'old' means the first noun is marked with the definite article, and the second noun with the indefinite article. In the theme = 'new' case this order is reversed.

Results

Three sets of results were derived from this experiment :

(1) the time from the onset of the picture (1100 msec. after the end of the sentence to be verified) to the subject responding.

(2) the number of fixations made on the picture in that period.

(3) the position of the second fixation. (The first fixation was always on the object in the centre of the picture).

1. Reaction Times (See Tables 1 - 4)

A five way analysis of variance was performed on the data. Factors were : theme (= previously mentioned object or not), voice (passive or active), truth (true or false), context (preamble or not) and subjects. This analysis yielded no significant F values other than that of subjects ($F_{21,21}$ = 8.05, $p < 0.001$) However there were slight trends towards true sentences being easier than false (1586 msec. vs. 1699 msec.; $F_{1,21}$ = 2.93, $p > 0.1$), and actives easier than passives (1595 msec. vs. 1690 msec.; $F_{1,21}$ = 3.07, $p > 0.1$) These effects, such as they are, are clearly the result of interactions. A truth value x voice interaction ($F_{1,21}$ = 3.44, $p \simeq 0.1$) seemed to show that only in the passive voice is there an effect of truth value : there is no difference between true and false actives, but true passives are 206 msec. faster on average than false passives. True passives did not differ from actives (1573 msec. and 1595 msec.) but false passives took much longer. A second interaction of theme with voice ($F_{1,21}$ = 3.53, $p \simeq 0.1$) showed that the position of the previously mentioned noun makes a difference of 140 msec. to passive times (they are lower when it is first in the sentence) but of only 57 msec. in the opposite direction for actives. Finally a third interaction of context and theme factors ($F_{1,21}$ = 3.04, $p > 0.1$) tended to show that the theme factor had little effect in the no context case (only 59 msec. - RTs being faster when it was the new noun) but a much larger effect in the opposite direction in the context case (1580 msec. vs. 1721 msec.) This last result is as one might expect if the theme factor was

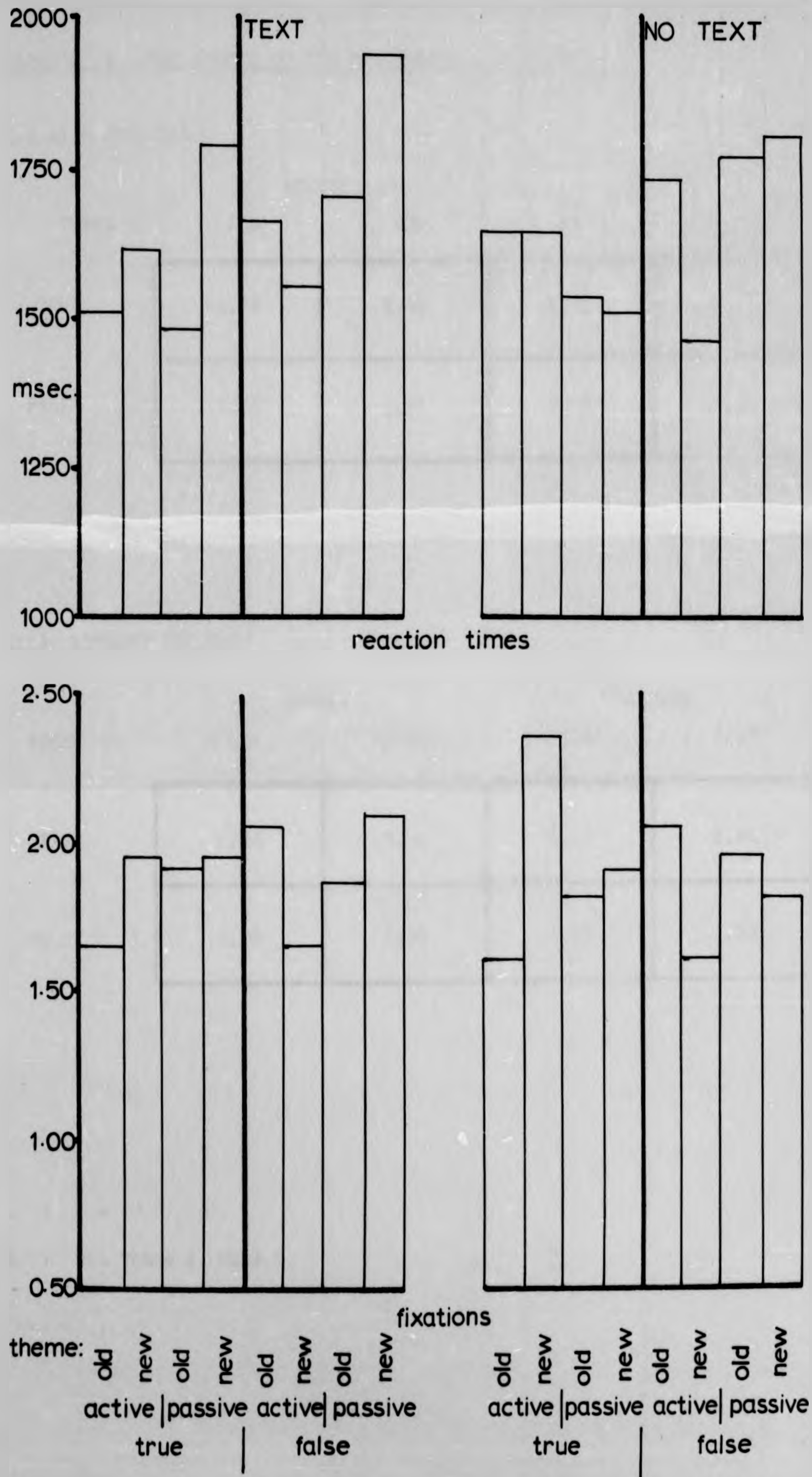


Figure 4 Results of Experiment 6

TABLE 2 : MEAN NUMBER OF EYE MOVEMENTS : EXPERIMENT 6

(i) WITH PREAMBLE

	ACTIVE		PASSIVE	
THEME =	OLD	NEW	OLD	NEW
TRUE	1.64[a]	1.95	1.91	1.95
FALSE	2.05	1.64	1.86	2.09

(ii) WITHOUT PREAMBLE

	ACTIVE		PASSIVE	
THEME =[b]	'OLD'	'NEW'	'OLD'	'NEW'
TRUE	1.59	2.36	1.82	1.91
FALSE	2.05	1.59	1.95	1.82

a : N = 22

b : See Table 1, Note b.

Table 3 : Experiment 6 : Analyses of Variance.

Effect	Reaction Time Data	Fixation Data
	F Values	
Theme	0.41	0.24
Voice	3.07	0.42
Truth	2.93	0.00
Context	0.03	0.00
Subjects	*** 8.05	*** 10.73
Theme x Voice	† 3.53	0.01
Theme x Truth	1.50	** 8.15
Theme x Context	3.04	0.00
Voice x Truth	† 3.44	0.24
Voice x Context	1.33	0.54
Truth x Context	0.07	0.35
Theme x Voice x Truth	1.01	* 7.44
Theme x Voice x Context	0.50	1.11
Theme x Truth x Context	0.09	0.90
Voice x Truth x Context	0.68	0.21
Theme x Voice x Truth x Context	0.05	0.00

All df are 1,21 except subjects which are 21,21

*$p < 0.05$

**$p < 0.01$

***$p < 0.001$

† $p < 0.1$

Table 4 : Experiment 6 : Means for major effects

A. Reaction Time Data (Figures are milliseconds)

1. Voice $F_{1,21} = 3.07$, not significant

Active	Passive
1595	1690

2. Truth $F_{1,21} = 2.93$, not significant

True	False
1586	1699

3. Voice x Theme $F_{1,21} = 3.53$, $p < 0.1$

	Theme = New	Theme = Old
Active	1566	1624
Passive	1760	1620

4. Context x Theme $F_{1,21} = 3.04$, not significant

	Theme = New	Theme = Old
With Preamble	1721	1580
Without Preamble	1606	1665

5. Truth x Voice $F_{1,21} = 3.44$, $p \leq 0.1$

	True	False
Active	1595	1596
Passive	1578	1802

B. Fixation Data (Figures are number of fixations)

1. Theme x Truth $F_{1,21} = 8.15$, $p < 0.01$

	Theme = New	Theme = Old
True	2.04	1.74
False	1.78	1.97

2. Theme x Voice x Truth $F_{1,21} = 7.44$, $p < 0.05$

	Active		Passive	
	Theme = New	Theme = Old	Theme = New	Theme = Old
True	2.15	1.62	1.92	1.86
False	1.62	2.05	1.95	1.90

genuinely the effect of the position of the previously mentioned noun rather than due to the asymmetry of the articles involved in the target sentence. If the latter were the case the no context case should be similar to the context case as the actual target sentences are identical.

2. Number of Eye Movements.[6]

A five way analysis of variance with the same factors as the reaction time data produced three effects : subjects ($F_{21,21} = 10.73$, $p < 0.001$), theme x truth ($F_{1,21} = 8.18$, $p < 0.01$) and theme x voice x truth ($F_{1,21} = 7.44$, $p < 0.025$). The theme x truth interaction shows that true sentences are easier when the previously mentioned noun is theme (1.73 vs. 2.04 fixations) and false ones when the new noun is theme (1.97 vs. 1.88 fixations). The higher order interaction shows that this is present only in the active voice (true sentences 1.61 vs. 2.15 fixations for theme = old and theme = new respectively; false sentences 2.05 vs. 1.61 respectively). With passives there are no effects (all roughly 1.9 fixations).

3. First two fixations

Subjects invariably look at the object in the centre of the picture first. The second fixation does not appear to be governed in any way by the sentence structure either. 2.8% of the time there is no second fixation - the first suffices. 8% of the time subjects fixate on the

6 The means presented in Table 7 are not directly comparable with those in Table 6. This is because means from the present data ignore the first fixation since this was always on the object in the middle of the picture. In Experiment 7 subjects often started with their eyes fixated in the middle of the picture, but often they did not. For that experiment if the first fixation was on the middle it was not counted - the first object was fixated. Accordingly the results for Experiment 7 ought to have means somewhat less than 1 fixation higher on average than those for Experiment 6. In fact they are much higher than that.

empty space. 89.2% of their second fixations are on the other object. If subjects were using the sentence to guide their scan of the picture they would always look at the empty space with false sentences. It is clear that they are not doing this : the figures for false sentences are 2.5%, 9.4% and 89.1% respectively; for true sentences they are 3%, 6.6% and 90.4%.

Discussion

Any conclusions drawn from the present experiment - particularly from the reaction time data, - must of necessity be very tentative as the effects are small and unreliable. This is probably due to the fact that comprehension and verification times are not separated and, more especially, that subjects are not highly practised and only receive a small number of trials (only one trial for each cell in the matrix). However, having said that, there do appear to be some grounds for a handful of limited conclusions.

Hall's experiment produced significant effect of both voice and truth in the same direction as those found here. It seems clear that in the present experiment these are largely due to high order interactions. In Hall's experiment subjects see the actor in position before each trial; this is similar in some ways to the use of a verbal preamble in the present experiment. In both cases attention is focused on one object, and subjects know this object will be in the centre of the display. But in Hall's experiment this object is always the actor in the sentence, whatever sentence voice, whereas in the present experiment its case role in the sentence to be verified is systematically varied by the voice and theme factors. The interaction of these two factors which appeared here, although failing to reach significance, suggests that when the preamble element is the patient passives are no harder than actives. Only with the actor being mentioned in the preamble does the voice effect appear : but this is just the situation in Hall's experiment, so that it is not surprising that she observes a main effect of voice. (Note : I am assuming here, for the reason given in the results section, that the voice x theme interaction is due to topicalisation and not simply definiteness asymmetry).

The slight truth value effect in the present experiment reaction time data was only evident in the passive voice. This result appears to be consistent with that of Gough (1966) who used only the no context case and observed the truth value effect only in the passive. However it is just the opposite of the effect Hall found with her eye movement data : she found no difference in number of eye movements to true and false passives, but a big difference between true and false actives. It is also not easily reconciled with the eye movement data from the present experiment which shows a relatively clear voice x theme x truth interaction. In fact Hall's data can be seen to follow from the eye movement results of the present experiment. In her experiment the theme is always the same as the object Ss have already seen in position in the active but always different from it in the passive. This is similar to having theme = previously mentioned item in the active and theme = new item in the passive. Now the present results show that neither theme nor truth has much effect on the number of fixations in the passive - hence Hall's finding of no effect of truth in the passive. But the present results show that less fixations occur when theme = previously mentioned item in the active if the sentence is true, but more if it is false. In Hall's experiment only the theme = previously mentioned item case occurs so that she finds a simple truth x voice interaction.

However one explains these data it seems clear that the explanation will have to be in two parts : an explanation of the reaction times and a separate explanation of the number of fixations. The two sets of results are quite different. On the one hand the RT data show effects of truth in the passive, but no effect in the active, and an effect of theme in the passive, but not in the active. On the other hand the fixation data show effects of truth and theme in interaction in the active, but no effects in the passive. In addition there is the slight influence of context on theme in the RT data.

The reaction time data - such as they are - are readily explained given the analysis of voice and theme presented in Chapter One and the Introduction to the present Chapter. The use of the passive is to enable one to have patient = theme without the need for marked theme and the main motive for doing this is a desire to continue to focus on the patient (together with a subsidiary desire to avoid marked information structure). The active is more neutral. These facts explain the direction of the context x theme result and the voice x theme result in a straightforward way. But why the voice x truth result? I suspect it is because the active, not having the presuppositions of the passive is more readily falsified. Of course this need not extend to non-reversible sentences, since these seem unlikely to need such thorough processing and comparison operations.

Turning now to the fixation data, the only explanation which occurs to me was suggested by doing the experiment and is largely phenomenological.. One gets the impression that it is easier to press the true button if the order of fixation and the order of mention of the two objects is the same, and easier to press the false button if they are different. Otherwise one tends to look again. This for some reason only occurs with actives - the knowledge that passives are "in the wrong order" anyway tends to lead one to drop that strategy with them. Now with actives fixation order is the same as order of mention with sentences with theme = old, and different with sentences with theme = new, so that true cases of the former will be easier than true cases of the latter but false cases of the former harder than false cases of the latter. This is what is found. This strategy presumably has no effect on RT because one can rescan very rapidly. Of course this is only a partial explanation but it is one which can be tested in a situation where subjects scan consistently left to right (or the reverse) rather than in the way they do in the present experiment. In fact in the next experiment they tend to adopt a consistent left to right strategy. Another way of avoiding this strategy is to present subjects with pictures in which more than the relevant two objects are depicted so that they cannot simply use peripheral visual information to locate the remaining object. This is done in the third eye movement experiment (Experiment 8) presented below.

Experiment 7

This experiment investigates the same factors as the previous one but differs in the nature of the materials. Instead of telling subjects in the preamble the position of one object it is now simply described as being in the picture. Instead of having one object in the middle and one on either the left or the right of the slide, the two objects are now at the sides only. The chief reason for doing this was to see if subjects would fixate first on actor or patient, theme or rheme.

Method

1. Subjects

22 first year undergraduates at Stirling University, aged approximately 17-19 years, 15 female, 7 male, fulfilling a requirement for the first semester psychology course.

2. Apparatus

As for Experiment 6.

3. Materials

As for Experiment 6 with the following exceptions:-

(1) Instead of the first sentence of the preamble saying "In the middle of this picture is a ....." it said only "In this picture is a .....". Otherwise the preambles and target sentences were the same.

(2) The slides were made up so that the objects were pictured one on either side of the picture with the middle space vacant. The size of eye movement necessary is increased therefore (see Fig.7). Again both objects always faced the right and were seen in profile. They were pictured near the top of the slides.

4. Design and Procedure

As for Experiment 6.

Results

The same three analyses were performed on the data from this experiment as on that from Experiment 6.

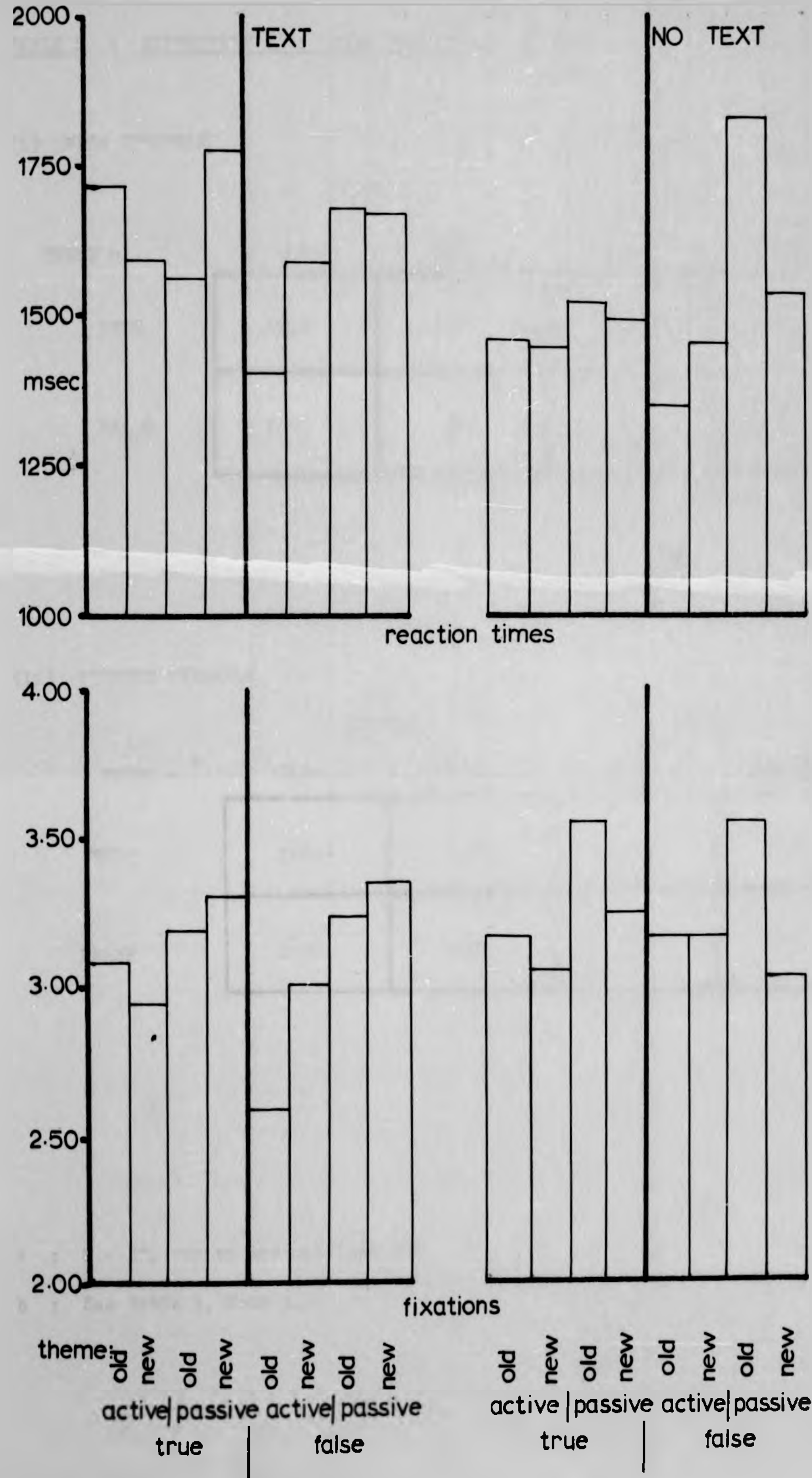


Figure 5 Results of Experiment 7

TABLE 5 : EXPERIMENT 7 : MEAN REACTION TIME DATA

(i) WITH PREAMBLE

	ACTIVE		PASSIVE	
THEME =	OLD	NEW	OLD	NEW
TRUE	1713[a]	1587	1563	1770
FALSE	1400	1587	1677	1661

(ii) WITHOUT PREAMBLE

	ACTIVE		PASSIVE	
THEME =[b]	'OLD'	'NEW'	'OLD'	'NEW'
TRUE	1465	1455	1523	1483
FALSE	1352	1459	1834	1540

a : N = 22. Units are milliseconds.

b : See Table 1, Note b.

(1) <u>Reaction Times</u>

Reaction times for this experiment are tabulated in Table 5. The overall mean here is slightly lower than Experiment 6 : 1566 msec. here, 1642 msec. in Experiment 6.

The analysis of variance yielded two significant results and one notable trend. The voice factor proved significant ($F_{1,21} = 6.27$, $p < 0.05$) with actives 128 msec. faster than passives (1502 msec. vs. 1630 msec.). Once again though there was an interaction with truth value ($F_{1,21} = 4.51$, $p < 0.05$) which showed that there was very little difference between true actives and true passives (1555 msec. and 1583 msec. respectively), but false actives were about 100 msec. faster than this (1449 msec.) and false passives about 100 msec. slower (1673 msec.). This result is similar to that found in the last experiment. However its interpretation is made quite complex by the presence of a theme x voice x truth interaction which only just fails to reach significance ($F_{1,21} = 3.85$, $p < 0.1$). The voice x truth effect appears to be confined to the theme = previously mentioned data. When the new noun is theme there does not appear to be any effect of truth, though the voice effect is enhanced (actives : 1521 msec. with true sentences, 1522 msec. with false; passives : 1626 msec. with true and 1800 with false). The subjects factor was, of course, significant ($F_{21,21} = 16.84$, $p < 0.001$).

(2) <u>Number of Eye Movements</u>

Number of fixations for this experiment are tabulated in Table 6. Even considering the reservations expressed in footnote 4 the number of fixations is clearly much higher here than in Experiment 6.

The analysis of variance produced a significant effect of voice ($F_{1,21} = 10.80$, $p < 0.01$) with actives leading to fewer fixations than passives (3.03 vs. 3.30). A context x theme interaction showed a non-significant trend ($F_{1,21} = 4.03$, $p < 0.1$) with the new noun as theme leading to more fixations in the context condition (3.17 for new = theme and 3.02 for old = theme), but less in the no context condition (3.12 vs. 3.36

TABLE 6 : EXPERIMENT 7 : MEAN NUMBER OF EYE MOVEMENTS

(i) WITH PREAMBLE

	ACTIVE		PASSIVE	
THEME =	OLD	NEW	OLD	NEW
TRUE	3.09 [a]	2.95	3.18	3.32
FALSE	2.59	3.04	3.23	3.36

(ii) WITHOUT PREAMBLE

	ACTIVE		PASSIVE	
THEME = [b]	'OLD'	'NEW'	'OLD'	'NEW'
TRUE	3.18	3.04	3.54	3.23
FALSE	3.18	3.18	3.54	3.04

a : N = 22

b : See Table 1, Note b

Table 7 : Experiment 7 : Analyses of Variance

	F Values	
Effect	Reaction Time Data	Fixation Data
Theme	0.00	0.66
Voice	* 6.27	** 10.80
Truth	0.01	0.22
Context	1.33	0.46
Subjects	*** 16.84	*** 10.76
Theme x Voice	0.54	0.54
Theme x Truth	0.01	0.19
Theme x Context	1.04	† 4.03
Voice x Truth	* 4.51	0.00
Voice x Context	0.45	0.47
Truth x Context	1.30	0.06
Theme x Voice x Truth	† 3.85	1.10
Theme x Voice x Context	1.22	1.36
Theme x Truth x Context	0.24	0.36
Voice x Truth x Context	0.15	0.54
Theme x Voice x Truth x Context	0.25	0.02

All df are 1,21 except subjects which are 21,21

* $p < 0.05$

** $p < 0.01$

*** $p < 0.001$

† $p < 0.1$

Table 8 : Experiment 7 : Means for major effects

A. Reaction Time Data (Figures are milliseconds)

1. Voice $F_{1,21} = 6.27$, $p < 0.05$

Active	Passive
1502	1630

2. Voice x Truth $F_{1,21} = 4.51$, $p \simeq 0.05$

	True	False
Active	1555	1449
Passive	1583	1678

3. Theme x Voice x Truth $F_{1,21} = 3.85$, $p < 0.1$

	Active		Passive	
	Theme = New	Theme = Old	Theme = New	Theme = Old
True	1521	1589	1626	1543
False	1523	1376	1600	1755

B. Fixation Data (Figures are number of fixations)

1. Voice $F_{1,21} = 10.80$, $p < 0.01$

Active	Passive
3.03	3.30

2. Theme x Context $F_{1,21} = 4.03$, $p < 0.1$

	Theme = New	Theme = Old
With Preamble	3.17	3.02
Without Preamble	3.12	3.36

respectively). The subjects factor was once again highly significant ($F_{21,21} = 10.76$, p 0.001).

(3) First Fixations

Subjects in both the context and no context conditions showed a highly consistent but uninteresting strategy : they scanned left to right. They did this on 84% of the trials in the no context condition and 77% of the trials in the context condition. Given the fact that the displays were oriented towards the right this means that they were scanning from the actual actor to the actual patient. They therefore saw the objects referred to in true actives and false passives in the same order as they were referred to in the sentence. The opposite was the case for false actives and true passives.

Since this simple strategy was so dominant there is insufficient data from the cases where this was not followed to observe any differences between actives and passives or true and false sentences.

Discussion

The reaction time analysis revealed a significant effect of voice with actives being reacted to faster than passives. Again though this appeared to be due to the fact that truth value and voice interact. As in the previous experiment true actives produced similar RTs to true passives. False passives again produced much longer RTs than either of these. The only difference is with false actives : these were reacted to faster than true actives in the present experiment, but marginally slower in the previous experiment. This is a small difference though. Of more importance is probably the presence of an almost significant three way interaction of these two factors with theme in the present experiment, though not in the previous one. This interaction is due to the fact that there is no effect of truth when theme = new item, or, to put it differently, that the truth x voice interaction is confined to cases where the previously mentioned item is theme. Passives are, as a result, consistently harder than actives when theme = new. If one equates the picture-focus procedure of the Olson and Filby experiments, with the preamble method used here, then they too found a three way interaction, though one with a rather different form. The form of the interaction for their three experiments and Experiments 6 and 7 here is indicated in the following table:

		Olson + Filby	Experiment 6	Experiment 7
Theme = old	Active	T < F	T < F	F < T
" " "	Passive	T < F	T < F	T < F
Theme = new	Active	F ≮ T	F < T	F ≮ T
" " "	Passive	F < T	T < F	F ≮ T

These results are clearly not consistent with one another but given the fact that the present results are unreliable, there seems little point in further discussion of this interaction until more data has been collected.

One rather disturbing aspect of the present data is the lack of any effects involving the context factor. The presence of context effects in

Experiment 6 but not here, despite the use of preambles in both, suggests the possibility that the locus of the context effect may not be in the use of a preamble at all, but rather may be in the use of pictures in which the "previously mentioned item" is always in the middle or (more generally) is in a position known to the subject before he sees the complete picture. This is the situation in both Experiment 5 and all of Olson and Filby's experiments, but not in Experiment 7. Hence the absence of any effects in the latter (Olson and Filby did not have the equivalent of a no context condition). To this argument I have two replies. Firstly, Chapters 2 and 3 amply demonstrate the influence of the preamble. Secondly it is possible that both preamble and knowledge of the position of the object may play a part. There do not, therefore, appear to be sufficient grounds for abandoning the assumption that it is the preamble which is the (or a) source of context effects, though this problem obviously requires further investigation.

Turning now to the fixation data, it seems fair to say that the hypothesis advanced in the Discussion of the previous experiment receives support from the present data. It was suggested there that the theme x voice x truth interaction (which was due to actives being easier when theme = old and true and theme = new and false than when theme = old and false and theme = new and true; with no effects in the passive) was due to a comparison effect. This was that when order of mention and order of fixation match it is easier to conclude "true", but when they mismatch easier to conclude that the sentence is false. Because of the different structure of the pictures in the present experiment the theme factor no longer has any effect on this process : with actives order of fixation (given the left-right scanning strategy) and order of mention match whenever the sentence is true and mismatch whenever it is false. The reverse is true of passives. Accordingly all that is observed is a simple voice effect not attributable to any interactions (as in the previous experiment). Actives consistently lead to fewer fixations than passives.

This is consistent with the hypothesis. The only other result in the fixation data is a non-significant trend towards the new = theme case leading to fewer fixations in the no context and more fixations in the context condition. I have no explanation for this effect, but it does provide some support for the assumption that it is the preamble which is the primary cause of context effects.

There appear to have been three major problems with both this experiment and the previous one:

(1) the failure to vary the side which is faced by the object in the picture has led to eye movement strategies which utilise this fact.

(2) the use of only the two objects mentioned in the sentence in the picture has meant that it is easy to use an "attracted scan" strategy, not bothering to compute a scan but rather allowing the presence of objects in the picture to attract one's eyes. This is efficient here because only relevant objects are present.

(3) there is a lot of "noise" in the data, preventing one from making more than highly tentative conclusions.

All these three problems are tackled in the next experiment in which the side faced is systematically varied, a third object is added to the pictures and two responses per subject are used to determine each plot instead of only one.

Experiment 8

Method

1. Subjects

16 Subjects, 9 female, 7 male, 12 were 1st. year undergraduates fulfilling a requirement for the first semester psychology course. 4 were postgraduates. Mean age of the subjects was approximately 19-20years.

2. Apparatus

As for Experiment 6 except that the slide projector was fitted with a shutter.

3. Materials

Sixteen tapes were prepared for each of the context and no context conditions, each of them of sixteen trials. All had different random orders of sentences. The pulse which activated the slide projector was placed on the second track of the tape and served to activate a shutter placed on the slide projector.

The slides differed from those of the previous two experiments in that there were now three objects in the picture : one on either side and one in the centre (see fig.3). This was done so as to encourage subjects to use the sentence to guide their scan of the picture. It also served to make sentences false in both a binary and non-binary fashion : e.g. a sentence such 'X ø Y' would not be false given

→
Z Y X

so that both 'Y ø X' and 'Z ø Y' are true.

All objects faced the same way, though the overall orientation of the picture was systematically varied.

4. Design and Procedure

This was as for Experiment 6 with the following exceptions. The design was now 2^5 factorial with the fifth factor being the side faced by the objects in the picture. This factor is collapsed for all analyses though, giving two responses per subject per plot. Each subject now had to view 32 experimental slides plus 8 practice slides. Since context and no context conditions were once again blocked this meant that they had 4 practice trials of the same sort as the condition to follow, before each condition. All practice slides were active, 2 true and two false, one of each facing left and one of each facing right. Since only 16 triples of objects were used each subject saw each triple twice : once in the context and once in the no context condition.

Half the subjects had the no context condition first and half the context condition. Each tape was used twice : once in the context first order and once in the context second order. No subject received the same random order in his context and no context trials.

There was a short break between conditions. Instructions were as per Experiment 6 with appropriate modifications for the number of trials and for the major procedure difference, namely the simultaneous onset of sentence and picture. The use of a shutter on the slide projector ensured that onset was simultaneous. The sequence of events was as indicated in the following diagram:

(P.T.O.)

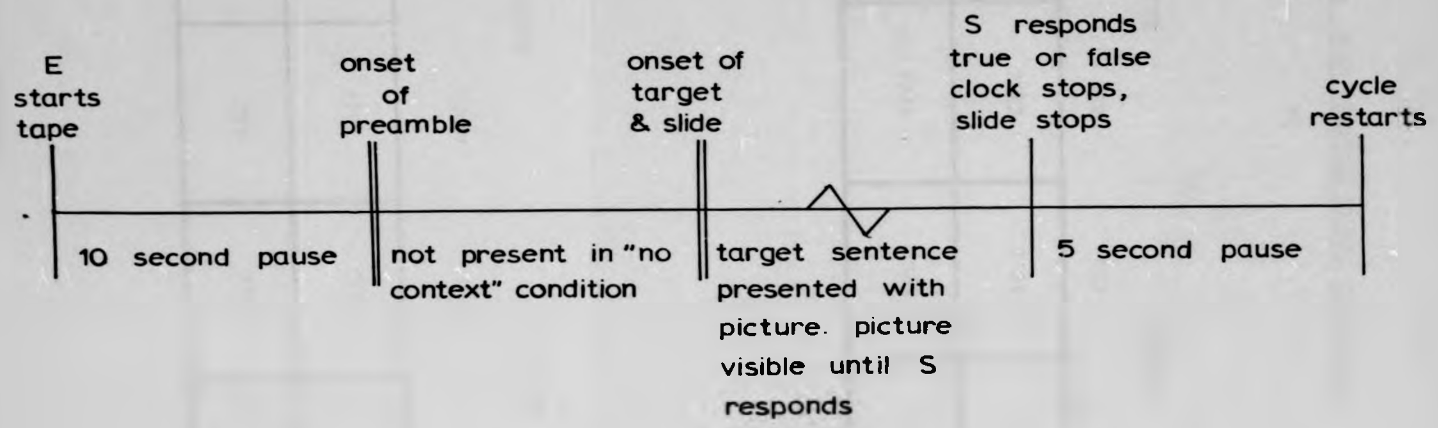


Figure 6 Time course of Experiment 8

TABLE 9 EXPERIMENT 8 : MEAN REACTION TIMES, UNADJUSTED

(i) WITH PREAMBLE

	ACTIVE		PASSIVE	
THEME =	OLD	NEW	OLD	NEW
TRUE	2449[a]	2413	2584	2741
FALSE	2517	2343	2883	2892

(ii) WITHOUT PREAMBLE

	ACTIVE		PASSIVE	
THEME =[b]	'OLD'	'NEW'	'OLD'	'NEW'
TRUE	2543	2754	2747	2694
FALSE	2666	2852	3021	2704

a. N = 16. Units are milliseconds.

b. See Table 1, Note b.

Results

Results were analysed by collapsing over the "faced side" factor and analysing the data in the same way as for Experiments 1 and 2, with a five factor ANOVA. Four analyses were performed:

(1) on RTs from the onset of the target sentence until the subject's response.

(2) on RTs from the offset of the verbal stimulus until the subject's response.

(3) on the number of fixations from the onset of the target sentence until the subject's response.

(4) on the first two fixations following the onset of the target sentence.

Data set (2) was derived in the following manner : each stimulus sentence was timed and this time then subtracted from the appropriate time in data set (1) to give the raw data for set (2).

(1) Total Reaction Times

Analysis of these data produced three effects significant at the $p < 0.05$ level : voice ($F_{1,15} = 8.18$) with actives taking less time than passives (2567 msec. vs. 2883 msec.), truth ($F_{1,15} = 4.73$) with true sentences taking less time than false (2615 msec. vs. 2734 msec.) and a three way interaction of context, voice and theme ($F_{1,15} = 4.53$). This last result appears to be due to the fact that, with context, actives are reacted to faster when the new item is theme, but passives when the old item is theme. The reverse is true when there is no context. This three way interaction gives rise to a context x voice effect which was not far from significance ($F_{1,15} = 2.99$, $p < 0.1$). There was no difference between active and passive in the no context data, but actives were nearly 350msec. faster than passives in the context data. The subjects factor was, of course, highly significant ($F_{15,15} = 28.07$, $p < 0.001$). There was a slight tendency for the context condition to lead to shorter times (2602 msec. vs. 2747 msec.) but this was not significant ($F_{1,15} = 2.57$, $p > 0.1$).

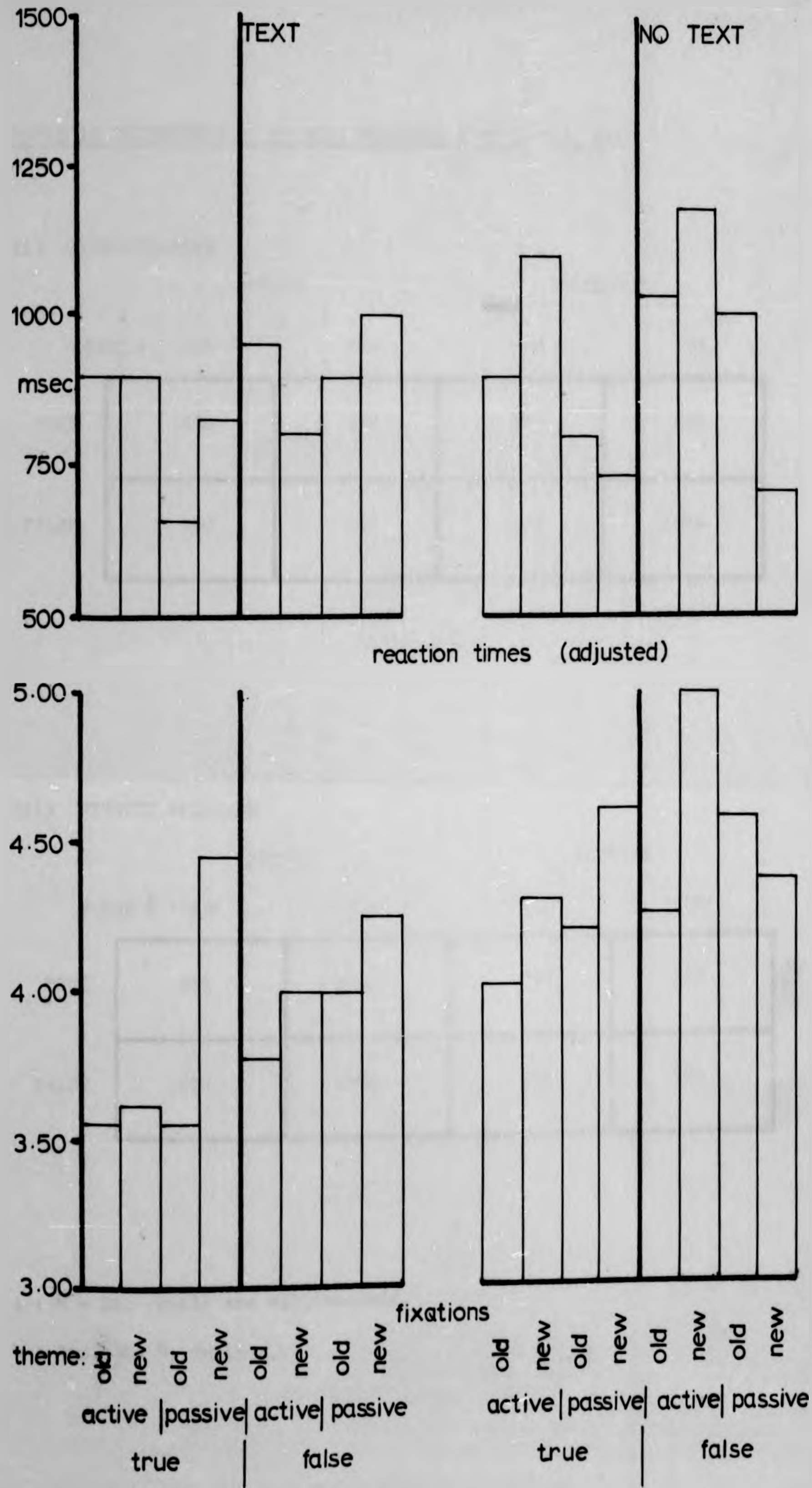


Figure 7 Results of Experiment 8

TABLE 10 EXPERIMENT 8 : MEAN REACTION TIMES, ADJUSTED

(i) WITH PREAMBLE

	ACTIVE		PASSIVE	
THEME =	OLD	NEW	OLD	NEW
TRUE	896[a]	890	663	832
FALSE	960	812	907	1014

(ii) WITHOUT PREAMBLE

	ACTIVE		PASSIVE	
THEME =[b]	'OLD'	'NEW'	'OLD'	'NEW'
TRUE	911	1099	792	727
FALSE	1037	1178	996	712

a : N = 16. Units are milliseconds.

b : See Note b, Table 1.

(2) Adjusted Reaction Times (Total Times minus speaking times).

The analysis of variance here revealed no significant effects whatsoever. The difference in speaking times of active and passive turned out to be about 360 msec. In addition there was a slight tendency for speaking times in the no context condition to be slightly longer : by about 90 msec. overall. This partly explains the tendency for the context total times to be shorter. It is not a complete explanation for obvious reasons : (1) because it is itself in need of explanation, (2) because it by no means follows that if speaking times had been identical overall so would total times - indeed it is quite possible that they would have been more different. In fact this 90 msec. difference seems likely to be due to an unconscious slowing down in the experimenter's speech to accomodate the listener since the no context sentences are richer in information than the context.

It is of course possible to apply this same argument to any adjustment of the times and this will be discussed below.

The only effect of any size at all in the adjusted times was a context x voice x theme interaction in the same direction as the main analysis ($F_{1,15} = 4.39$, $p < 0.1$).

(3) Number of Fixations

This analysis produced several significant effects : (1) when the theme was the previously mentioned item (= item marked with "the" = item in the middle of the picture) there were fewer fixations (3.90 vs. 4.32; $F_{1,15} = 10.28$, $p < 0.01$); (2) true sentences lead to fewer fixations than false sentences (4.04 vs. 4.28; $F_{1,15} = 4.93$, $p < 0.05$); (3) there were fewer fixations in the context than in the no context data (3.00 vs. 4.42; $F_{1,15} = 5.72$, $p < 0.05$); (4) there was a truth x voice x theme interaction ($F_{1,15} = 5.77$, $p < 0.05$). With theme = old (= noun marked with "the" = object in middle of the picture) false sentences produced more fixations than true and actives slightly more than passives. With theme = new (= noun marked with "a" = object in one of the peripheral slots) false

TABLE 11 EXPERIMENT 8 : MEAN NUMBER OF EYE MOVEMENTS

(i) WITH PREAMBLE

	ACTIVE		PASSIVE	
THEME =	OLD	NEW	OLD	NEW
TRUE	3.56 [a]	3.62	3.56	4.44
FALSE	3.78	4.00	4.00	4.25

(ii) WITHOUT PREAMBLE

	ACTIVE		PASSIVE	
THEME =[b]	'OLD'	'NEW'	'OLD'	'NEW'
TRUE	4.03	4.31	4.19	4.62
FALSE	4.28	5.00	4.59	4.37

a : N = 16. Units are milliseconds.

b : See Table 1, Note b.

Table 12 : Experiment 8 : Analyses of Variance

Effect	F values		
	Reaction Time Data		Fixation Data
	Unadjusted	Adjusted	
Theme	0.00	0.04	** 10.28
Voice	* 8.18	3.48	3.00
Truth	* 4.73	3.20	* 4.93
Context	2.57	0.41	* 6.72
Subjects	*** 28.07	*** 38.54	*** 18.74
Theme x Voice	1.34	0.52	0.00
Theme x Truth	1.20	0.89	0.46
Theme x Context	0.02	0.12	0.02
Voice x Truth	1.98	1.73	1.40
Voice x Context	† 3.99	2.66	0.79
Truth x Context	0.03	0.00	0.13
Theme x Voice x Truth	0.35	0.04	* 5.77
Theme x Voice x Context	* 4.53	† 4.39	2.40
Theme x Truth x Context	0.00	0.04	0.15
Voice x Truth x Context	0.58	0.84	0.41
Theme x Voice x Truth x Context	0.30	0.50	0.19

All df 1,15 except subjects which are 15,15

* $p < 0.05$

** $p < 0.01$

*** $p < 0.001$

† $p < 0.1$

Table 13 : Experiment 8 : Means for major effects

A. Unadjusted Reaction Times (Figures are milliseconds)

1. Voice $F_{1,15} = 8.18$, $p < 0.05$

Active	Passive
2567	2883

2. Truth $F_{1,15} = 4.73$, $p < 0.05$

True	False
2615	2734

3. Context x Voice $F_{1,15} = 3.99$, $p < 0.05$

	Active	Passive
With Preamble	2430	2775
Without Preamble	2704	2791

4. Context x Voice x Theme $F_{1,15} = 4.53$, $p = 0.05$

Active		Passive		
Theme = New	Theme = Old	Theme = New	Theme = Old	
2378	2483	2816	2733	With Preamble
2803	2604	2699	2884	Without Preamble

B. Adjusted Reaction Times (Figures are milliseconds)

1. Voice $F_{1,15} = 3.48$, not significant

Active	Passive
972	830

2. Truth $F_{1,15} = 3.20$, not significant

True	False
851	952

3. Context x Voice $F_{1,15} = 2.66$, not significant

	Active	Passive
With Preamble	889	854
Without Preamble	1056	806

Experiment 8 : Means for major effects (cont'd.)

4. Context x Voice x Theme $F_{1,15} = 4.39$, $p < 0.1$

	Active		Passive	
	Theme = New	Theme = Old	Theme = New	Theme = Old
With Preamble	851	928	923	785
Without Preamble	1138	974	719	894

C. Fixation Data (Figures are number of fixations)

1. Theme $F_{1,15} = 10.28$, $p < 0.01$

Theme = New	Theme = Old
3.99	4.32

2. Voice $F_{1,15} = 3.00$, not significant

Active	Passive
4.07	4.25

3. Truth $F_{1,15} = 4.93$, $p < 0.05$

True	False
4.04	4.28

4. Context $F_{1,15} = 6.72$, $p < 0.05$

With Preamble	Without Preamble
3.90	4.42

5. Theme x Voice x Truth $F_{1,15} = 5.77$, $p < 0.05$

	Active		Passive	
	Theme = New	Theme = Old	Theme = New	Theme = Old
True	3.96	3.79	4.53	3.87
False	4.50	4.03	4.31	4.29

sentences produced a lot more fixations than true with actives, though slightly less than true with passives.

(4) First Two Fixations

As in Experiment 6 there was a very strong tendency indeed to fixate first on the object in the centre of the picture. Of the total of 512 trials only on 4 occasions did a subject fixate first on any spot other than the centre.

The second fixation data is much more interesting than that for the other two experiments. Taking the context condition first. With theme = previously mentioned (and therefore the centre object - i.e. the one fixated first) subjects' second fixation is on the spot where the second object would be were the sentence true. There is no difference consequently between true and false sentences and little difference between actives and passives (the probability of this occuring is roughly 0.70 with actives and 0.65 with passives). When theme = new noun (i.e. the object not fixated first) subjects appear to behave as if they thought the sentence was a true active looking at the spot where the second object should be if it were a true active. The probability of this occurring is 0.70 with passives, 0.9 with true actives and 0.70 with false actives.

The no context data are more complex. When the first noun is marked with "the" (and therefore in the centre of the picture) subjects' second fixation is on the second object mentioned with passives (probability of this is roughly 0.75) but on the irrelevant object with actives (probability 0.60 of this). This only applies to true sentences : with false sentences there is a roughly even chance of them fixating either the second object referred to or the irrelevant object. In fact only the result for true actives really differs from random behaviour which perhaps suggests that subjects tend to assume a sentence beginning with "the" will be a passive, but conflicts between this and voice information derived from the sentence, and between the latter and the noun referred to in the second noun phrase lead to random behaviour.

When the first noun is marked with "a" (and is therefore in the periphery and not fixated first) subjects tend to look at the spot where the object should be with actives (probability of this is roughly 0.65) but at the spot where the object is with passives (probability = 0.70 for true passives and 0.75 for false). This tendency to fixate the theme with passives but to use a more calculating strategy with actives perhaps reflects the additional stress which goes with theme in the passive.

Discussion

The most outstanding result of the present experiment is the different pictures of the difficulty of the passive conveyed by the adjusted and unadjusted times. Any model of sentence comprehension which states that processing of the sentence cannot begin until the whole sentence has been input to the listener would have to predict that passives are simpler than the corresponding actives. The usual psychological interpretations of Chomskyan grammars make this assumption about processing, but at the same time insist on the greater transformational complexity of the passive. The present results appear to make this position totally untenable. However it is possible to suggest that the subject uses his knowledge of the constraints of the experimental situation to infer the structure of the later parts of the sentence given only the earlier parts - indeed this is assumed in describing the fixation results above. But this already presupposes that preliminary processing of the sentence can begin before the whole sentence is input - otherwise, to put the argument in its extreme form, the early stages of the sentence would not be regarded even as part of a meaningful sentence (as, for example, subject and verb of a transitive sentence). It is possible to maintain adherence to a Chomsky grammar which works on whole strings by postulating some sort of analysis by synthesis model which generates whole strings to match them to input and then formulates hypotheses about the sentence structure which some separate non-linguistic processor then compares with its knowledge of the situation, filtering out unlikely interpretations. This is something of a tall story, though, and a much simpler solution would be to devise a sentence comprehension model which worked in a left to right manner. Furthermore a left to right model would seem to be much more useful given the organisation of speech in time, since it does not place the load on memory of a whole string based approach. It would also be consistent with Clark's (1965) sentence generation data. Systemic grammar is much more sympathetic towards a left to right model than any other extant grammar (see Winograd, 1972).

One major difficulty does arise with the present data on relative difficulty of actives and passives. If the fundamental difficulty is simply one of the length of the input string, why did Miller and McKean find a voice effect even after they had deducted a measure of reading time from the overall times? Why also did Gough find an effect of voice on verification even with a delay between sentence presentation and the verification process? Since the present experiment appears to be the only one in which a verification task has been used with simultaneous presentation of sentence and picture it may be that the problem lies in the coding of the sentence in memory for the brief period before the picture is presented (or in the case of the picture first condition the coding of the picture until the sentence is presented). It is possible that in tasks where subjects have time to reflect on the nature of the sentence they have difficulty in remembering the extra information which the passive carries. In the present task no memory problem of this sort would occur. Of course we are left with the fact that in the first two of the present experiments when the previously mentioned noun is theme/subject there is no simple voice effect. But this is explicable on the same grounds : it is possible with a context to see the point of using the passive to make the old noun = theme, so that no extra information is conveyed. That the problem is not simply one of the amount of information conveyed by the passive but rather of remembering it for a short period appears to be confirmed by the fact that in the present experiment with the no context case passives are consistently easier - but the passive surely continues to carry more information - there is simply no memory component.

All this of course is subject to the criticism that the adjusted reaction times are not a legitimate measure. That position would be difficult to justify though, given the following consideration. In most verification tasks it is possible to make a correct response given the picture plus the sentence up to and including the main verb (this will of course include all operators - in particular 'not'). Now the passive always has one extra grammatical item in that section of the sentence, namely the auxiliary, so that someone knowing all task constraints would still take longer with passives - even if they were no harder. The present measure may paint too rosy a picture of passives because of the fact that they are also longer after the verb (because of the presence of "by") : a fact which might be thought to make the adjusted time too short by just that margin. However, "by" presumably does not take longer to say than 150 msec., and adding this to the passive times still makes them no longer than the active times. Also : unlike most verification experiments (including Experiments 5 and 7 in the present series) it is not possible in the current experiment to make a correct response until the whole sentence has been uttered. This is so because of the fact that sentences in the present experiment are always false (if false at all) in both a binary and non-binary fashion (see Materials).

This in itself does not make the adjusted times a good measure of difficulty but it does mean that in a situation like the present one with simultaneous presentation of sentence and picture they are bound to be at least as good a measure as the unadjusted times which inevitably make the passive appear harder. Further : as long as (1) there is a suspicion that the usual paradigms which present sentence and picture separately and in that order produce voice effects because of a memory component and (2) the picture-sentence order is subject to a picture coding effect (as shown by Olson and Filby), then the present measure seems at least as good as any other. Certainly the explanation presented in the last paragraph but one (in terms of the information conveyed, and to be remembered) seems to cover all the basic voice effects, including interactions with the position of any noun referring to an object which has been topicalised one way or another - be it by a preamble or by special construction of the picture.

This explanation does not cover two other effects observed in the reaction time data in the present experiment : namely the effect of truth value and the interaction of voice with theme in the no context case where, one is assuming, there is no topicalisation. This assumption might well be incorrect given the fact that people do use articles to mark topics (see Grieve, 1974, and Grieve and Wales, 1973), but it appears justified in the present experiment given the fact that the theme x voice interaction in the no context data shows the passive to be reacted to faster when theme = old - precisely the opposite of the result in the context condition. The simplest explanation is that subjects are using a comparison strategy here which involves decoding the sentence into ordered case roles. Given that subjects always fixate the centre object first, it follows that where theme = noun marked with 'the' (= noun referring to the object in the centre), in the active voice they are fixating the sentence actor, but in the passive voice the sentence patient. It appears easier to respond if one fixates the actor first. This is consistent with an account which states that subjects decode the sentence into an ordered format with the referring expressions in the same order as in an SAAD and then code the picture in a similar fashion. However it is also consistent with Hall's result with children that it is easier to fixate second on an object in front of the object fixated first, rather than behind it. If one expects the sentence to be true and fixates first on the actor then one's second fixation will be on the object which it faces. This is not true if one fixates first on the patient.

Hall's explanation is shown not to work since a simple prediction derived from it is readily disconfirmed. If it were true that adults, like children, have difficulty looking behind an object then one should find more cases of people looking in front of the centre object on the second fixation than behind it. In fact of the 254 second fixations in the no context data only 106 are to the object in front of the centre object.

Apparently one is left with the ordered case role explanation : subjects prefer (ceteris paribus) to code the sentence in some deep format which has elements ordered, perhaps Verb (Actor, Patient) and they find it easier to verify the sentence if they fixate the actor first in the picture. There is a difficulty here though : in Experiment 6, which uses displays identical to those here in respect of the order of scanning of the sentence case roles, there is no evidence in the no context data of a preamble element effect (Mean RT for Actor in the centre 1666 msec., for Patient 1604 msec.). Further : in Experiment 7, in which the properties of the display and subject's scanning strategies are different from those here, there is some evidence of a preamble element effect similar to that observed here (Mean RT for Actor in the preamble 1460 msec., for Patient 1565 msec.). These facts taken together tend to vitiate any simple explanation and lead one towards the rather pessimistic view of Glucksberg et al. that the process of comparing sentences against pictures is so dominated by the demands of any particular task that it is difficult, if not impossible to talk in terms of a single process.

This view is further reinforced by the presence of an overall main effect of truth value in the reaction times of the present experiment, compared with a tendency in both Experiments 6 and 7 towards a truth value x voice interaction.

Interpretation is not helped either by the complexity of the eye movement data. Certainly the data on the number of eye movements in the present experiment produced a significant main effect of truth and in that respect confirms the results of the reaction time data. However there are several difficulties here. Firstly this is the only effect evident in both the RT data and the fixation data. Secondly the fixation results here are quite different from those of either of the previous two experiments. Of the four significant results only one (the truth x voice x theme interaction) occurs in either of the other experiments, and then (Experiment 6) in a quite different form. Of particular interest is the presence of a very strong theme effect in the present experiment, but not in the previous ones. It seems that the reference to the object in the middle of the picture as S fixates it at the start of his scan greatly facilitates scanning. It is therefore all the more surprising that this effect did not occur in Experiment 6 which is so apparently similar. It seems likely that the difference lies in the fact that three objects are present in the pictures used here but not in the earlier experiment, although it is also possible that it lies in the use of simultaneous presentation here but not in the other experiment. I suspect it lies in the former : if in Experiment 6 one finds that the first object fixated is not the same object as that referred to by the first noun phrase one simply looks at the other one (in fact it is not even necessary to look). But in the present experiments one needs to look at at least one of the locations to see if the object being referred to is there. This would explain the theme effect without any need to invoke linguistic factors.

Any explanation of the three way interaction would have to be more complicated. In Experiment 6 this took the form of a more or less constant number of fixations for the passive, regardless of truth and theme, but a superiority with false actives when the new item was theme, but with true actives when the old item was theme. In the present experiment true actives produce less fixations than false ones, this effect being greater when the new item is theme. Passives, however, produce less fixations with the old item as theme (i.e. theme marked "the") when the sentence is true, but less when the new item is theme if the sentence is false. I have no explanation for this result or its difference from the earlier one. The lack of relationship between the fixation and RT data suggests that a scanning rather than a linguistic explanation would be appropriate.

The data on the first two fixations produced some quite enlightening results. As already noted subjects almost invariably fixate on the object in the centre of the picture first ($p > 0.99$). The second fixation varies with the language used and, to a lesser extent, with truth. Taking the context data first it is apparent that when the theme is the object fixated first subjects tend to fixate second on where the other object would be were the sentence true. This is true regardless of either voice or truth, with a probability overall of occurring almost seven times in ten. When theme = new (i.e. is marked with "a") subjects appear to behave as if they thought the sentence was a true active. That is they look (on the whole) where the new object would be if the sentence were a true active. The probability of this is roughly 0.9 if it _is_ a true active and 0.7 if it is a passive or a false active. Note that with false actives and true passives subjects look away from the actual object mentioned first - or, in other words, behind the object fixated first. This not only goes against the tendency to look in the direction which an animal or person is facing but it also means they are not using peripheral information to direct their scan. - or rather not using it very much : the increased probability of fixating the theme in a true active shows it is used a little . It would appear that the tendency to assume that a sentence beginning with a noun marked with the indefinite article is active, when there is an object which has been topicalised, is very strong indeed. This effect would not show up, of course, in any paradigm other than simultaneous presentation of sentence and picture because of the fact that subjects would have voice information available to them before beginning to scan in the sentence-first case.

Subjects behave much less consistently in the no context case. When the first noun is marked with "the" they show a slight tendency to behave as though the sentence were a true passive looking at the point where the second object would be were that the case. However this only exceeds chance level with actual true passives ($p = 0.75$), being 0.5 with false sentences and 0.60 with true actives. The fact that both voice and truth clearly have some kind of effect here reflects the fact that subjects are less dependent on the sentence structure to direct their scan. As noted in the results section when the first noun is marked with the indefinite article subjects tend to look at where the noun referred to by the theme is with passives (roughly $p = 0.75$) but at where it ought to be with actives ($p = 0.65$). Since subjects are scanning while the sentence is continuing it is hard to see how they can know whether a sentence is going to be active or passive in time to alter their scanning in this way. It is possible though that while they are fixating on the centre object (the object mentioned second in the sentence) the sentence reaches the verb and so enables them to project a scan. It is impossible to gain a precise idea of the time relationships from the experimental data. It is apparent that they compute and carry out the scan with actives as if they were assuming it to be true whereas with passives there is more of a tendency to find the object referred to by the theme. This too is in line with the idea that passives serve to [illegible] the subject and indicate the speaker's point of departure : the listener here appears to need to start from the same point of view as the speaker by finding the theme. This does not seem to be the case with actives. On the other hand one should not overstress this point given the different behaviour when the first noun is marked with "the".

Experiments 6,7 and 8 : General discussion.

The RT data reveal a much more consistent pattern of results across the three experiments with the context than with the no context condition. This is especially so with passives. In all three experiments RTs to true passives are lower when the theme refers to the previously mentioned object (i.e. is marked with "the" and is in the centre of the picture in Experiments 1 and 3) : 305 msec. in Experiment 6, 207 msec. in Experiment 7 and 169 msec. in Experiment 8 (adjusted times). This is an unweighted mean of 227 msec. False passives do not produce such consistent results, the figures being 248 msec.,-16 msec. and 107 msec. respectively - an overall mean of 124 msec. This figure is considerably lower than that for true sentences but it should be viewed in the light of the fact that false no context passives show a considerable decrement [illegible] when the theme = noun marked with "the" (compared to the case where the theme is marked "a"), while true passives show little effect of theme (see below).

There is much less pattern in the data from active sentences in the context condition. There is no overall effect of theme in true sentences : in Experiment 6 responses are faster to the case where the first nominal is marked with "the" by 125 msec., in Experiment 7 they are slower by 126 msec., and in Experiment 8 slower by 5 msec. (adjusted times). The picture is similar with false sentences : in the sixth experiment responses are slower when the first nominal is definitely marked by 112 msec., in Experiment 7 faster by 187 msec., and in Experiment 8 slower by 148 msec. (overall mean 24 msec. slower when the first nominal is definitely marked - i.e. when the actor is the old item). These overall means of 2 msec. and 24 msec. argue for the neutrality of the active in context. There appears to be a voice x truth value interaction present in two of the three sets of data (Experiments 6 and 8) due to actives being about as difficult whether the sentence is true or false, but passives being easier if it is true. All three experiments show little difference overall between actives and passives if the sentence is true but substantially shorter times in the active if it is false.

The no content data are more confusing. There is very little evidence of the usual voice effect (actives consistently easier). The voice x truth value effect really only occurs in the no content data in Experiment 1. False sentences seem to be reacted to faster if passive with the first noun marked with "a" or active with the first nominal marked with "the". This is quite different from the data for true sentences.

Any assessment of the linguistic parameters involved is inevitably clouded in these data by the different scanning strategies subjects used in the different experiments. There appears to be more variability across experiments in the no content data RTs than in the context data RTs. This may be due to the greater susceptibility of this condition to strategy effects. If one makes the assumption that the scanning and comparison processes are essentially independent of context then a clearer picture begins to emerge. Of course I have thrown some doubt on that assumption in the course of discussing the three experiments, but let us adopt it for the sake of argument. Given this assumption it makes sense to say that one can simply subtract the no context times from the context times to arrive at a figure for the facilitating or inhibiting effect of context on any particular sentence type. One can use this method given the assumption that one of the stages (interpretation, scanning, comparison) is simply speeded up. Otherwise one would be subtracting incommensurables. The justification of the method will be largely empirical anyway : if it leads to a simpler picture it is worth accepting - otherwise not. It does appear to lead to a simpler picture.

Taking first passives : the theme effect is more or less constant for all three experiments for both truth values. The mean is [illegible] msec. with a range of [illegible] msec. to 301 msec. facilitation of theme = old over theme = new. This is very impressive given the variability of the three sets of experimental materials and procedures. The results for actives are not quite so simple. In both Experiment [illegible] and Experiment [illegible] there does not seem to be any interaction between truth value and theme. Experiment 6 shows an advantage of theme = old (124 msec. with true sentences, 132 msec. with false) and Experiment 8 of theme = new (194 msec. with true and 289 msec. with false). Experiment 7 shows a more complex pattern : actives are only facilitated by having theme = new if they are true (by 116 msec.), the reverse being true if they are false (by 8[illegible] msec.). All this averages out at a net advantage of roughly 50 msec. when the theme refers to the object mentioned in the preamble (i.e. marked with "the") but with facilitation ranging from 289 msec. to -132 msec. this is not a very helpful figure. Nevertheless comparing it with the mean figure for the passive there does seem to be evidence that the passive, unlike the active, is not neutral with respect to topicalisation. This point is emphasised if one compares the overall facilitation of the context over the no context condition. With actives with old = theme (i.e. first nominal marked with "the") facilitation of context over no context ranges from 165 msec. to -248 msec. with a mean of 0 msec. Actives with new = theme have a mean facilitation figure of 44 msec. and a range of 360 msec. to -132 msec. The corresponding figures for passives are 78 msec. (with a range from 157 msec. to -40 msec.) and -204 msec. (with a range from -105 msec. to 302 msec.). The very similar overall times for actives in the context and no context conditions emphasises the relatively "context free" nature of the active, while the considerable context effects with the passive emphasise the context sensitive nature of this choice. The direction of the effect with the passive brings out the role of the passive as an option selected in order to provide cohesion with prior discourse by making the topic of the prior discourse thematic in the sentence. The use of the subtractive method clearly rules out explanations based on either definiteness asymmetry or fixation order.

The eye movement data treated by this subtractive method provide more support for this position. Taking passives first, there is only one case of the theme = new examples giving rise to more facilitation than the to be any interaction between truth value and theme. Experiment 6 shows an advantage of theme = old (124 msec. with true sentences, 132 msec. with false) and Experiment 8 of theme = new (194 msec. with true and 289 msec. with false). Experiment 7 shows a more complex pattern : actives are only facilitated by having theme = new if they are true (by 116 msec.), the reverse being true if they are false (by 80 msec.). All this averages out at a net advantage of roughly 50 msec. when the theme refers to the object mentioned in the preamble (i.e. marked with "the") but with facilitation ranging from 289 msec. to -132 msec. this is not a very helpful figure. Nevertheless comparing it with the mean figure for the passive there does seem to be evidence that the passive, unlike the active, is not neutral with respect to topicalisation. This point is emphasised if one compares the overall facilitation of the context over the no context condition. With actives with old = theme (i.e. first nominal marked with "the") facilitation of context over no context ranges from 165 msec. to -248 msec. with a mean of 0 msec. Actives with new = theme have a mean facilitation figure of 44 msec. and a range of 366 msec. to -132 msec. The corresponding figures for passives are 78 msec. (with a range from 157 msec. to -40 msec.) and -204 msec. (with a range from -105 msec. to 302 msec.). The very similar overall times for actives in the context and no context conditions emphasises the relatively "context free" nature of the active, while the considerable context effects with the passive emphasise the context sensitive nature of this choice. The direction of the effect with the passive brings out the role of the passive as an option selected in order to provide cohesion with prior discourse by making the topic of the prior discourse thematic in the sentence. The use of the subtractive method clearly rules out explanations based on either definiteness asymmetry or fixation order.

The eye movement data treated by this subtractive method provide more support for this position. Taking passives first, there is only one case of the theme = new examples giving rise to more facilitation than the theme = old (figures for Experiment 6,7 and 8 respectively are -0.05, 0.45, and 0.45 for true sentences and 0.36, 0.[illegible]3 and 0.35 for false sentences). Once again these results are quite impressive in their uniformity. Again the actives are quite different (figures in the same order are -0.46, 0, -0.02 and 0.05, 0.45 and -0.50). There are exceptions but these results as a whole clearly demonstrate the superiority of the passive with theme = old over the passive with theme = new, and reaffirm the lesser importance of choice of theme in the active.

In using the subtractive method one is getting rid of anything common to both the context and no context conditions. Hopefully we are cutting out the effects of specialised strategies since we are here interested in the tasks for the light they throw on linguistic processing rather than for their own sake. However it may be that common syntactic processing is being cut out too. This seems possible - even probable. However one has to consider what might be involved here. Since context apparently facilitates some processing (as well as, it would appear, inhibiting other processing) there is clearly some of the reaction time in the no context case which is being used for non-basic (i.e. non-common) syntactic processing. So it would seem that, <u>if there is any basic syntactic processing</u>[6], there is no way we can observe it since it is not isolable. Unless that is we can use a large number of tasks and find what is common to all of them. This would not be easy, needless to say, and there would be no point in going through with it unless there was some independent reason for believing in such processing. Chomskyan grammar does not impute meaning to syntactic choices, so that the subtractive method used here should not produce variation between two sentences whose meaning is the same at a deep level, being related by a transformation. But it is clear that

6 The complete absence of a correlation between mean comprehension times for the 32 sentences of Experiment 1 (r = -0.05) suggests that there may not be any common syntactic processing.

after the subtractive method has been applied there are great differences between actives and passives. This constitutes a good reason for rejecting Chomsky-style grammars as a model of how we process sentences.

Systemic theory, because of its greater emphasis on the function of different grammatical choices, is in a far better position to cope with this kind of contextual variation. Nobody has so far attempted to produce a performance model from systemic theory[7] so that it is difficult to see exactly how it would be applied to the present problem. But it seems clear that the deepest level of coding will, in the short term anyway, contain emphasis and topicalisation data and that the verification process will not be independent of that.

7 The nearest to a performance model is Winograd (1972) and this uses only a fragment of SG - no facilities for handling the "textual" and "interpersonal" components are incorporated in his system.

Chapter 6 : An Experiment Involving Answering Wh- Questions

Introduction

This experiment extends the range of phenomena covered to include transitive questions. This has the advantage of providing further coverage of thematic options within the transitive clause type, including voice, while at the same time investigating in a rather preliminary fashion some of the relationships between mood and thematic choice. It may be as well to state here at the outset that the move towards questions was not motivated in any way by a desire to investigate a different "speech act" (Searle, 1969). Indeed it is not clear to me that the questions of the present experiment do differ from the statements of earlier experiments in regard to their speech act status. The whole notion of the illocutionary force of an utterance is dependent upon the complex system of social roles and conventions of relevance to the ongoing situation. In some cases the felicitous rendition of a speech act is dependent on the syntax and semantics of the utterance produced. This is certainly not always the case. Though we do not have anything like a thorough analysis of speech acts, it is clear that the surface structure of most sentences is related in only a very indirect way to their speech act role. The complex of social roles and demands of the present experiment and those which preceded it are sufficiently similar to suggest that the speech acts on the experimenter's part are the same in all the experiments, despite the syntactic variation.

The primary reason for switching to questions was the variation in both the range of choices available to a speaker and in the way the various choices are expounded in questions as compared with indicatives.

This point will be discussed in depth below. There were a number of secondary reasons for choosing questions. Firstly subjects cannot just use a set of simple matching strategies which would enable them to respond correctly given only a part of the sentence. Secondly although it is possible in theory to work out the answer before the question has been fully asked, if one quickly divines the task structure, subjects do not interrupt a question to answer : they wait until it is complete. (I should add that divining the possibilities to the extent of being able to predict how certain of the sentences must finish is extremely difficult.) This is not of great importance but it does make the paradigm neater. Of greater importance is the fact that subjects have to produce a linguistic response : this makes the paradigm rather more natural than the verification task. This is even more true in that one has to make an oral response. These are all minor reasons but they add up to a wholly different attitude to the task than the one encountered in verification tasks. In doing the verification task one feels that it is merely a question of "getting the hang of it" - it's simply a trick to be learnt. This is not at all the feeling of the question task, it feels stranger and more taxing and one is not aware of "getting the hang of it".

<u>Linguistic Analysis of a subset of Wh- Questions</u>

The primary interest in looking at questions is, I repeat, the different range of choices available in the question. This can be brought out by a consideration of the parallels with the corresponding declaratives. The experiment is restricted to consideration of simple transitive Wh-questions and so the discussion which follows will be restricted to these with very little consideration of linguistic problems not immediately related to the sentences used in the experiment.

There are two primary Wh- question structures, exemplified by

1. Who hit Fred?

2. Who did John hit?

and their passive counterparts

1a. Who was hit by John?

2a. Who was Fred hit by?

It may be that in fact the correspondence should be 1 : 2a and 2 : 1a since the questioned element in both 1 and 2a is the actor and in 2 and 1a it is the patient. However the surface structure Wh-Vb-N is common to both 1 and 1a and the structure Wh-aux-N-Vb is common to both 2 and 2a. Since claims about surface structure involve, on the face of it, less preconceptions than those about deep structure I will assume the latter to be the correct correspondence, at least for the moment.

As a first approximation one might suppose that the indicative sentences corresponding to the above interrogatives are (i) to (iv) respectively:

(i) John hit Fred

(ii) Fred, John hit

(i)a Fred was hit by John

(ii)a John, Fred was hit by.

This seems at first to be right because (ii) and (ii)a certainly don't seem satisfactory as answers to 1 and 1a, whereas (i) and (i)a seem perfectly alright as answers to 1 and 1a. However (i) and (i)a seem also to be alright as answers to 2 and 2a. Now the (ii)s differ from the (i)s in that the former expound a syntactic choice which Halliday refers to as "marked theme" (Notes, 2, 218ff). The temptation is to suggest that this is what also distinguishes the 1s from the 2s. This ignores the definition of theme, though. According to Halliday theme is a function of mood. In particular the unmarked theme of an indicative is the subject of the sentence, the subject being that element in concord with the verb (i.e. actor in an active, patient in a passive sentence). In an interrogative the unmarked theme is the modal or auxiliary verb in a polar interrogative and the Wh- item in a Wh- interrogative. This follows directly from Halliday's definition of the theme as the subject's point of departure for the sentence : in a question that is obviously the request for information, at least in the usual case. A marked theme is one where a decision as to sentence initial position is made which does not accord with the decision which would be made on the basis of mood alone. Marked themes are of two sorts : intrinsic case roles, and adjuncts. Adjuncts are much the most common and fall into four types:

(a) conjunctions e.g. Although John hit Fred --------

This is intrinsic to the sentence structure unlike (b).

(b) discourse adjunct e.g. "However", "Despite that", "But".

These serve to relate the sentence to what has gone before.

(c) modal adjunct e.g. "perhaps", "probably" etc.

(d) complement e.g. "yesterday" etc.

It seems that these can all occur in a single clause generating a complex theme e.g. Meanwhile, back at the ranch, perhaps because they were

Meanwhile,	back at the ranch,	perhaps	because
Disc.	compl.	modal	conj.

feeling ill, ---------

(Halliday is not explicit on this point). The other type of marked theme is much less common and is exemplified by (ii) above for the case of an indicative. For the Wh- interrogative an example of a marked theme would be

3. John hit who?

Here the non-Wh- item in the sentence is theme : a marked case for the interrogative.

To return again to the examples above : given this definition of theme it is clear that the 1s do not differ from the 2s on this dimension. Both have Wh- as thematic and are therefore unmarked. This is in agreement with the fact that the (i)s seem reasonable answers to both the 1s and 2s. The difficulty here seems to arise from a failure to represent the intonation pattern. If we distinguish

(v) John hit Fred

from (vi) John hit Fred

Where underlining denotes additional stress then we can see that 1 can only be answered by (v) and 2 by (vi). Similarly if we introduce

(vii) Fred was hit by John

and (viii) Fred was hit by John

we have the correspondence 1a : (vii) and 2a : (viii).

Theme here is unmarked and intonation may be marked (as in (v) and (vii)) or unmarked (as in (vi) and (viii)). But if a marked theme is chosen this necessarily takes up an intonation contour of its own (see Halliday p.218-222) so excluding any possibility of this element being old information (i.e. the part which would be 'given' in the corresponding question). So (ii) and (ii)a have to correspond to 2 and 2a because the marked theme cannot be old information, which is what it is given as in 1 and 1a. The two syntactic structures in the interrogative correspond to a difference of information structure in the indicative : a difference which may be expounded in either intonation alone ((v) vs (vi)) or intonation and syntax ((vi) vs (ii)). The latter seems to be a much more emphatic option and it may seem surprising that the question structure does not appear to convey this. In fact it may do so in one of two ways : either by additional stress of the Wh- item, or by marked syntax. The former may still only correspond to highly emphasised version of (v) to (viii), but the latter seems to clearly correspond to the indicative marked theme. An example would be 3 above : 'John hit who?'

The need for two unmarked structures for the question seems to derive from the very limited variation possible in the question intonation in English compared with the almost endless variety in the indicative (Halliday 1967 b). This is of course due partially to the focus on new information - i.e. the Wh- item. A full list of the suggested indicative/ interrogative correspondences is given in Table 1.

There are several problems with this linguistic analysis as it stands when one starts to bring in aspects of the situation in which the various questions might be asked. Note firstly that all of the questions in Table 2 presuppose more shared knowledge than the simple "What happened?" question. Secondly the passive structures presuppose an agreement over a topic prior to the utterance of the question - something which is not true of the actives. This follows directly from Halliday's analysis of the

Table 1 : Suggested list of correspondences between indicative and Wh-interrogative for the set considered.[a]

Who did John hit?	:	John hit Fred
Who hit Fred?	:	John hit Fred
Who was hit by John?	:	Fred was hit by John.
Who was Fred hit by?	:	Fred was hit by John.
John hit who? / Who did John hit?	:	Fred, John hit/John hit Fred.
Who hit Fred?	:	John hit Fred.
Who was hit by John?	:	Fred was hit by John
Fred was hit by who?/	:	John, Fred was hit by.
Who was Fred hit by?		/Fred was hit by John

a. underlining indicates stress differing from the unmarked (sentence final) form, or not deducible from the syntactic pattern, or additional to that deducible from the syntactic pattern.

function of the passive. But given this, the question type "Who was ∅ by x?" is very odd since it questions the identity of the patient by means of a structure which should only be selected if the patient is old information. This does not, of course, apply to the question type "Who was x being ∅ by?" nor to either of the active types since these have less specific entry conditions.

<u>Predictions derived from the linguistic analysis</u>

Translating this linguistic analysis into predictions as to subjects' reaction time to the various question types, we find the following. First actives should be easier than passives, though passives with the patient questioned should be more difficult than passives with the agent questioned. This should apply regardless of whether there is a context or not. However matters will not be quite this simple where there is a preamble topicalising one element. Most importantly, the passive should be relatively easier when the topicalised element is patient in the question : again this is an application of Halliday's notion of the function of the passive. It works with the other application of that explanation (cited above) to produce a rather counterintuitive set of predictions for sentence types A - D (Questions here are assumed to be based on the display Fred John Jack where "←" denotes the direction faced and all three people are running; John is previously mentioned - the other two are not) :

A. Who is being chased by Jack?

B. Who is being chased by John?

C. Who is Fred being chased by?

D. Who is John being chased by?

Firstly one can predict from Halliday's analysis that C and D should be easier than A and B. Secondly A should be easier than B and D easier than C.

This analysis of the nature of these Wh- questions makes no predictions as to the relative difficulty of the various active questions. They should all produce rather similar reaction times both with and without a topicalised element.

The experiment reported below once again measures eye movements during the scanning of the picture. Again though reaction times will be used as the definitive measure of processing difficulty. As the previous experiments showed eye movement data may help in interpretation of RTs but their exact relationship to processing difficulty remains, as yet, very obscure indeed.

Method.

1. Subjects 14 first year undergraduate psychology students fulfilling a course requirement. 5 males and 9 females. Average age approximately 19.

2. Apparatus

The same viewing box, projection and video equipment was used as in experiments 6, 7 and 8. Additional equipment was as follows : as before a Revox A77 taperecorder was used to present the materials to the subject, a Revox A700 tape deck recorded the whole procedure both the materials presented and the subject's response, which was spoken into a microphone placed within the viewing box 2 - 4 cm. from the subject's mouth. It was from this recording that all time measures were obtained.

3. Design, Materials and Procedure

The design is basically four factor, within subjects, the four factors being 1. whether there is a preamble or not 2. whether the object mentioned in the preamble is mentioned in the question 3. whether the question is active or passive 4. whether the noun mentioned in the question is early or late in the sentence. In the no context case factor 2 is assigned by correspondence to the context condition.

If x and y denote the object mentioned in the preamble and one of the other objects in the picture, respectively, and Ø denotes one of the four verbs used (viz. 'chase', 'follow', 'shoot', 'watch'), then the eight sentence types are as follows (with factors 2, 3 and 4 cycling in that order)[1] :

What is Ø x?

What is Ø y?

What is being Ø by x?

What is being Ø by y?

1. Note that all nominals in the present experiment are definitely marked as the subject can see the referent at the same time as he hears the nominal.

What is x Ø?

What is y Ø?

What is x being Ø by?

What is y being Ø by?

As in Experiment 8 the direction faced by the objects in the slides was systematically varied so that one response was obtained from each subject to each sentence type with the picture oriented to the left and one with it oriented towards the right. This amounted to 16 responses per subject in each of the context and no context conditions. Eight different random orders of the 16 sentence type/faced side combinations were generated and a tape made for each of these for both the context and no context conditions. Each subject received a different context and no context random order. Seven subjects received the no context condition first, seven the context condition. No random order was used more than twice for either the 'context' or 'no context' cases. Slides used were identical to those used in Experiment 8 the object described in the preamble always being in the middle of the picture. Practice trials comprised four trials before each run - with context if the run was with context, without if without. The same four trials were used throughout - sentence types 1 and 3, one of each with the slide facing left and one of each with it facing right.

When the experiment commenced subjects were shown the apparatus and what it did and the chin rest and seat height were adjusted so that the subject was comfortable and a good view of the right eye was obtained on the monitor. They were told the experiment was in two halves and the instructions for the first half were given them. If they were to receive the context condition first they were told that the slide would onset and simultaneous with it there would be a description lasting three sentences describing one of the objects in the picture. At the end of the description there would be a question. They were to answer this as quickly and briefly as possible - preferably in one word. They were told that it was not important that the name they gave the object was absolutely precise so long as it could not be confused with one of the other objects, that is, so long as the experimenter could tell which object they meant. The necessity for speed was stressed.

Subjects in the no context condition were given these instructions amended appropriately, the slide now onsetting from the onset of the question.

The slide onset was operated from a voice key fed from the recorder with the stimulus materials on it. The voice key opened a shutter on the slide projector. When the subject responded the experimenter closed the shutter by means of a key which also advanced the projector one slide. The tape recorder with the stimulus materials continued to play throughout. Trials were spaced at 10 second intervals so that if the subject did not respond in this time E closed the shutter ready for the next trial. This in fact only happened once altogether.

After the first block (i.e. either the context or no context condition) there was a break of about three minutes while E altered the orientation of the slides ready for the next block. The same slides were used in the same order in both conditions.

Results

All subjects were asked which half of the experiment they found easier. All seven subjects who received the context condition first found the no context condition easier. Three of those with the no context condition first found the context condition easier and four found the no context condition easier. Thus there seems to be a straightforward order effect, with the second half easier, but also an overall tendency to think the no-context condition easier. This is interesting for two reasons : (1) it would seem to show that the shorter exposure of the slide in the no context condition was not felt as a hindrance and (2) the reaction times were significantly slower for the no context condition, in apparent contradiction of subject's expressed opinion of the difficulty of the two cases (see below).

Several analyses were performed on the data. It was felt useful to have an analysis of the lengths of the different sentences and so an analysis of variance was performed on the figures for their durations with the four experimental factors as fixed and the 8 orders as random factors. In other words the four factors (and interactions) were tested against the order x factor interaction as error.

In addition tests were performed on the overall RT's from the onset of the question to the onset of the answer (onset-onset times) and from the offset of the question to the onset of the answer (offset-onset times). Analysis was also carried out on the number of fixations from the onset of the question to the onset of the answer. Separate analysis is performed on the context and no context data, as well as the main analyses with all the data in a single anova.

Rather than go through each analysis separately I will go through them factor by factor, considering all the separate analyses at once. Separate tables for each analysis are attached. This method of presentation has the advantage that one can consider the effects of duration of the stimulus material on other measures very easily.

	NOUN BEFORE MAIN VERB		NOUN AFTER MAIN VERB	
	ACTIVE	PASSIVE	ACTIVE	PASSIVE
NOUN PREVIOUSLY MENTIONED	923	1261	940	1253
NOUN NOT PREVIOUSLY MENTIONED	1017	1400	1051	1353

"CONTEXT"

	NOUN BEFORE MAIN VERB		NOUN AFTER MAIN VERB	
	ACTIVE	PASSIVE	ACTIVE	PASSIVE
NOUN PREVIOUSLY MENTIONED	976	1259	1078	1403
NOUN NOT PREVIOUSLY MENTIONED	943	1367	1043	1450

"NO CONTEXT" N = 8

TABLE 2 MEAN STIMULUS DURATIONS (msec.)[1]

1. Note that the figures from this table plus the offset-onset times do not add up to the onset-onset times because of the fact that materials for the 14 subjects were randomly selected from this set.

	NOUN BEFORE MAIN VERB		NOUN AFTER MAIN VERB	
	ACTIVE	PASSIVE	ACTIVE	PASSIVE
NOUN PREVIOUSLY MENTIONED	2656	2685	2421	3645
NOUN NOT PREVIOUSLY MENTIONED	2563	2919	3058	3309

"CONTEXT"

	NOUN BEFORE MAIN VERB		NOUN AFTER MAIN VERB	
	ACTIVE	PASSIVE	ACTIVE	PASSIVE
NOUN PREVIOUSLY MENTIONED	3334	3139	3199	3922
NOUN NOT PREVIOUSLY MENTIONED	2930	3084	3392	4109

"NO CONTEXT" N = 14

TABLE 3 MEAN ONSET-ONSET TIMES (msec.)

	NOUN BEFORE MAIN VERB		NOUN AFTER MAIN VERB	
	ACTIVE	PASSIVE	ACTIVE	PASSIVE
NOUN PREVIOUSLY MENTIONED	1742	1427	1483	2387
NOUN NOT PREVIOUSLY MENTIONED	1536	1522	2003	1971

"CONTEXT"

	NOUN BEFORE MAIN VERB		NOUN AFTER MAIN VERB	
	ACTIVE	PASSIVE	ACTIVE	PASSIVE
NOUN PREVIOUSLY MENTIONED	2362	1894	2130	2610
NOUN NOT PREVIOUSLY MENTIONED	2029	1726	2357	2708

"NO CONTEXT" N = 14

TABLE 4 MEAN OFFSET-ONSET TIMES (msec.)

	NOUN BEFORE MAIN VERB		NOUN AFTER MAIN VERB	
	ACTIVE	PASSIVE	ACTIVE	PASSIVE
NOUN PREVIOUSLY MENTIONED	3.36	4.29	3.43	4.68
NOUN NOT PREVIOUSLY MENTIONED	4.46	5.39	5.18	5.50

"CONTEXT"

	NOUN BEFORE MAIN VERB		NOUN AFTER MAIN VERB	
	ACTIVE	PASSIVE	ACTIVE	PASSIVE
NOUN PREVIOUSLY MENTIONED	5.54	5.54	5.11	5.64
NOUN NOT PREVIOUSLY MENTIONED	5.14	5.57	5.29	6.43

"NO CONTEXT" N = 14

TABLE 5 MEAN NUMBER OF FIXATIONS

	NOUN BEFORE MAIN VERB		NOUN AFTER MAIN VERB	
	ACTIVE	PASSIVE	ACTIVE	PASSIVE
NOUN PREVIOUSLY MENTIONED	3.36	4.29	3.43	4.68
NOUN NOT PREVIOUSLY MENTIONED	4.46	5.39	5.18	5.50

"CONTEXT"

	NOUN BEFORE MAIN VERB		NOUN AFTER MAIN VERB	
	ACTIVE	PASSIVE	ACTIVE	PASSIVE
NOUN PREVIOUSLY MENTIONED	5.54	5.54	5.11	5.64
NOUN NOT PREVIOUSLY MENTIONED	5.14	5.57	5.29	6.43

"NO CONTEXT" N = 14

TABLE 5 MEAN NUMBER OF FIXATIONS

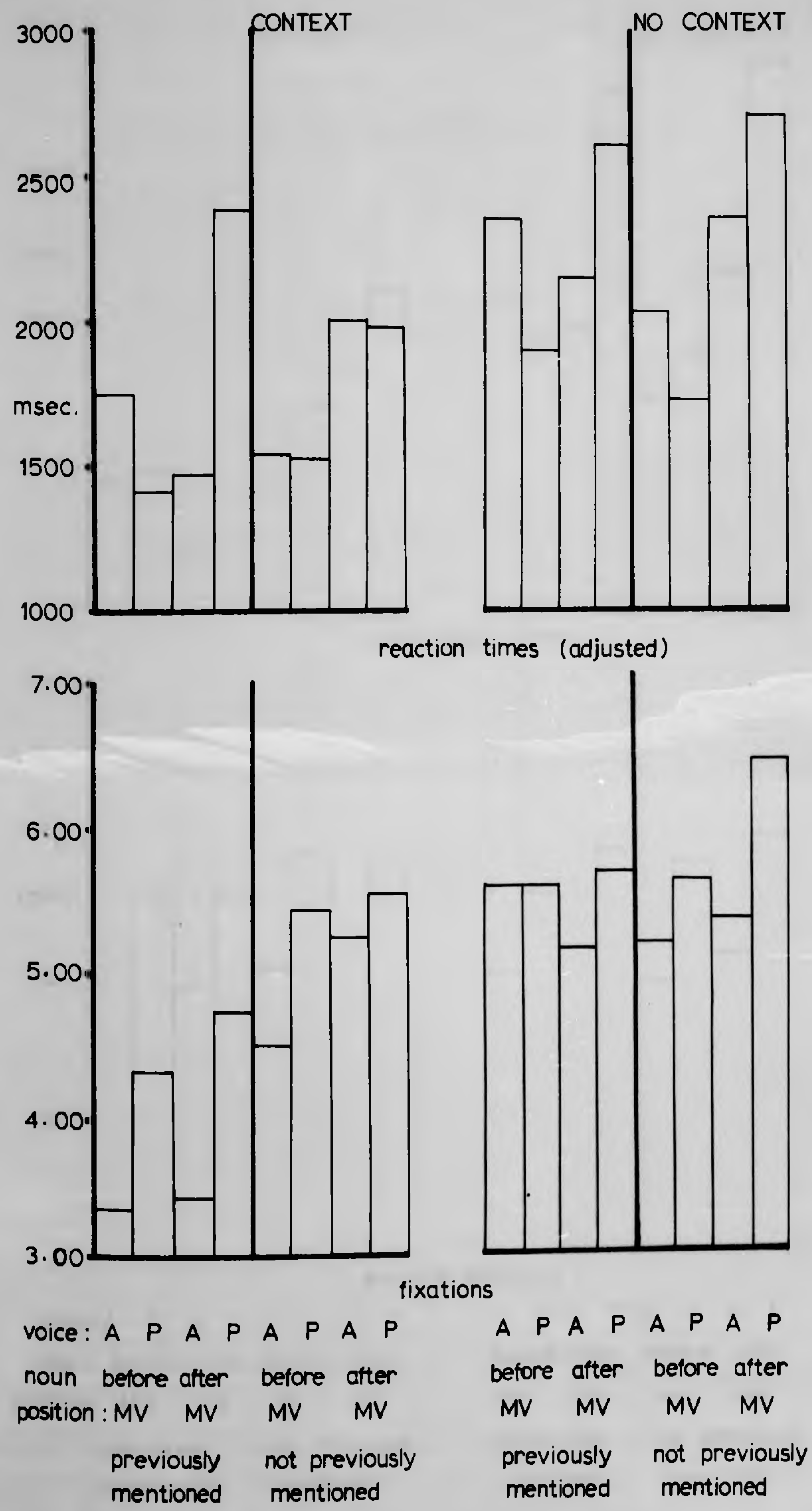


Figure 1 Experiment 9: principal results (accompanies Tables 4 & 5)

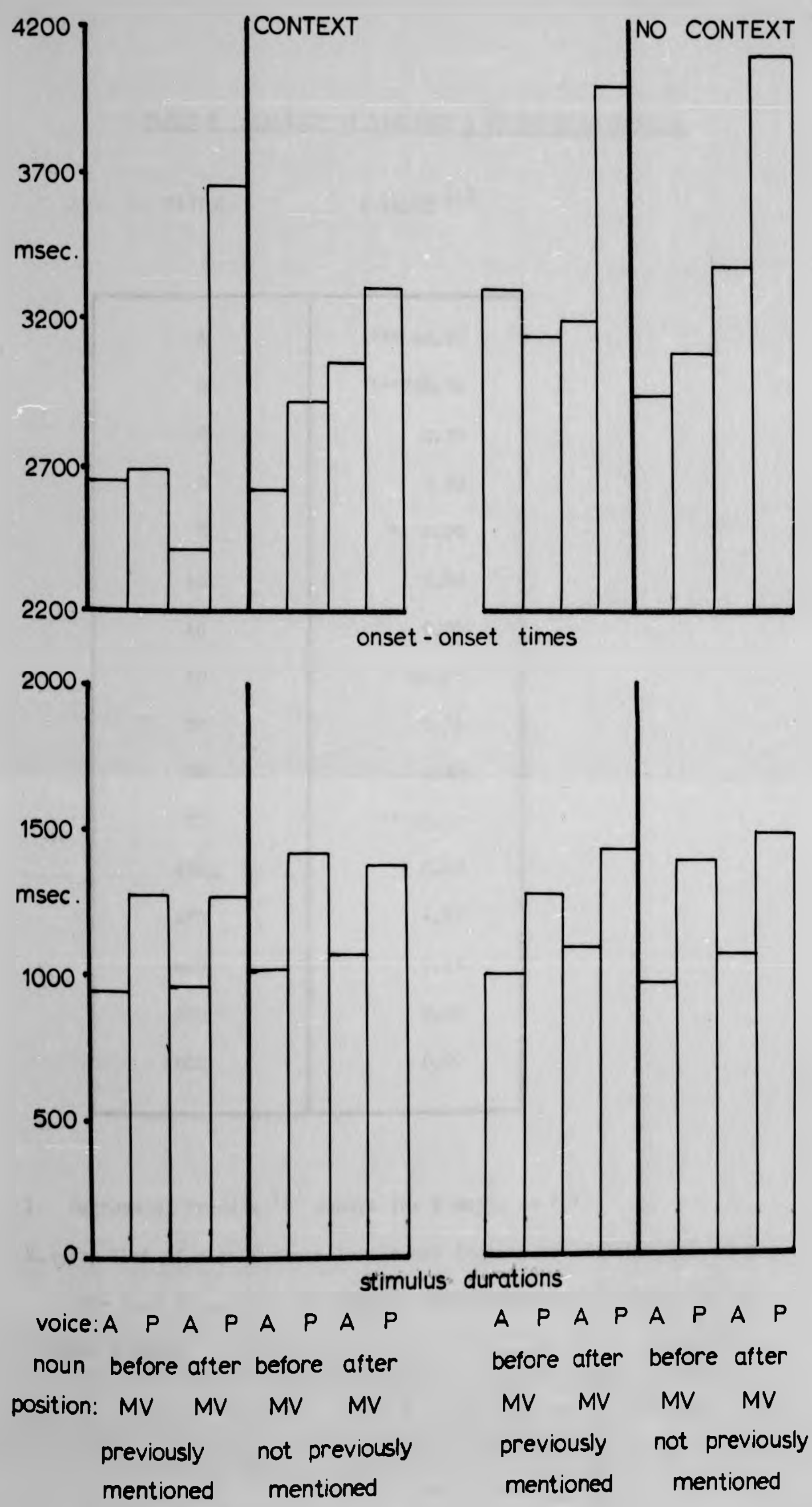


Figure 2 Experiment 9: subsidiary results (accompanies Tables 2 & 3)

TABLE 6 ANALYSIS OF VARIANCE : STIMULUS MATERIALS.

FACTOR	F VALUE [1,2]
A	*** 40.50
B	***738.38
C	3.73
D	3.23
E	* 4.39
AB	3.60
AC	0.35
AD	*11.20
BC	0.18
BD	0.46
CD	** 13.93
ABC	0.79
ABD	4.17
ACD	0.24
BCD	0.83
ABCD	0.00

1. Degrees of Freedom ^{1,7} except for E which is 7,7.
2. For list of significance levels and factors see onset-onset anova. Note that E is here not subjects but quadruples of object triples and a verb.

Table 6A Summary of Significant Effects : Stimulus Durations.[a]

1. Previous Mention $F_{1,7} = 40.50$, $p < 0.001$

Previously mentioned	Not Previously Mentioned
1136	1203

2. Voice $F_{1,7} = 738.38$, $p < 0.001$

Active	Passive
996	1343

3. Previous Mention x Context $F_{1,7} = 11.20$, $p < 0.05$

	Previously Mentioned	Not Previously Mentioned
Context	1094	1205
No Context	1179	1201

4. Syntactic type x Context $F_{1,7} = 13.93$, $p < 0.01$

	Noun Position	
	Before Main Verb	After Main Verb
Context	1150	1149
No Context	1136	1243

a. Figures are milliseconds.

TABLE 7 ANALYSIS OF VARIANCE : ONSET-ONSET TIMES

FACTOR [1]	ALL DATA [2]	CONTEXT ONLY	NO CONTEXT ONLY
A	0.25	0.52	0.03
B	*** 20.08 [3]	*** 23.19	* 6.14
C	** 14.93	* 6.42	** 14.10
D	* 4.89	-	-
E	*** 17.69	*** 5.26	*** 9.60
AB	0.05	0.48	0.24
AC	1.82	0.48	2.34
AD	0.37	-	-
BC	** 11.30	3.85	*** 16.68
BD	0.49	-	-
CD	0.41	-	-
ABC	2.94	4.54	0.60
ABD	0.60	-	-
BCD	0.35	-	-
ABCD	2.15	-	-

1. Factors are as follows:-
   - A. Whether the noun in the sentence is referred to in the preamble.
   - B. Voice : Active or Passive.
   - C. Syntactic Type : Wh- (aux)-Vb-N or Wh-aux-N-Vb.
   - D. Context : Preamble or No Preamble.
   - E. Subjects.
2. Figures are F Values. All degrees of freedom 1,13 except E which is 13,13.
3. Significance levels are denoted as follows:
   - * $p < 0.05$
   - ** $p < 0.01$
   - *** $p < 0.001$

Table 7A Summary of Significant Effects : Onset-Onset times. [a,b]

1. Voice $F_{1,13} = 20.08$, $p < 0.001$

Active	Passive
2944	3361

2. Syntactic Type $F_{1,13} = 14.93$, $p < 0.01$

Noun Position	
Before Main Verb	After Main Verb
2914	3391

3. Context $F_{1,13} = 4.89$, $p < 0.05$

Context	No Context
2907	3397

4. Voice x Syntactic Type $F_{1,13} = 11.30$, $p < 0.01$

	Noun Position	
	Before Main Verb	After Main Verb
Active	2871	3017
Passive	2957	3765

a. Figures are milliseconds.

b. The separate analyses of context and no context data are not given here. Where of interest these are given in the accompanying text.

TABLE 8 ANALYSIS OF VARIANCE : OFFSET-ONSET TIMES.[1,2]

FACTOR	ALL DATA	CONTEXT ONLY	NO CONTEXT ONLY
A	0.04	0.00	0.07
B	0.63	2.35	0.01
C	** 10.83	* 5.63	* 7.42
D	4.18	-	-
E	*** 19.15	*** 6.80	*** 9.87
AB	0.48	0.78	0.00
AC	1.68	0.14	2.71
AD	0.02	-	-
BC	** 13.79	4.52	*** 17.14
BD	0.58	-	-
CD	0.05	-	-
ABC	3.04	* 5.67	0.34
ABD	0.43	-	-
BCD	0.46	-	-
ABCD	3.18	-	-

1. Figures are F Values, degrees of freedom 1,13 except for E which is 13,13.

2. For list of factors and significance levels see previous Table Notes 1 and 2.

Table 8A Summary of Significant Effects : Offset-Onset Times.[a,b]

1. Syntactic Type $F_{1,13} = 10.83$, $p < 0.01$

Noun Position	
Before Main Verb	After Main Verb
1779	2206

2. Voice x Syntactic Type $F_{1,13} = 13.79$, $p < 0.01$

	Noun Position	
	Before Main Verb	After Main Verb
Active	1917	1993
Passive	1642	2419

a. Figures are milliseconds.

b. The separate analyses of context and no context data are not given here. Where of interest these are given in the accompanying text.

TABLE 9 ANALYSIS OF VARIANCE : FIXATIONS [1,2]

FACTOR	ALL DATA	CONTEXT ONLY	NO CONTEXT ONLY
A	** 10.89	*** 15.42	0.25
B	*** 20.64	*** 32.70	2.60
C	1.86	2.60	0.53
D	** 10.88	-	-
E	** 6.04	2.57	*** 5.67
AB	0.01	0.68	1.53
AC	1.42	0.13	2.88
AD	* 5.41	-	-
BC	0.88	0.12	2.16
BD	0.66	-	-
CD	0.42	-	-
ABC	0.23	0.71	0.05
ABD	1.60	-	-
ACD	0.82	-	-
BCD	1.37	-	-
ABCD	0.99	-	-

1. Figures are F Values, degrees of freedom 1,13 except for E which is 13,13.

2. For list of significance levels and factors see onset-onset anova Notes 2 and 3.

Table 9A Summary of Significant Effects : Number of Fixations. [a]

1. Previous Mention $F_{1,13} = 10.89$, $p < 0.01$

Previously Mentioned	Not Previously Mentioned
4.68	5.37

2. Voice $F_{1,13} = 20.64$, $p < 0.001$

Active	Passive
4.69	5.38

3. Context $F_{1,13} = 10.88$, $p < 0.01$

Context	No Context
4.53	5.52

4. Context x Previous Mention $F_{1,13} = 5.41$, $p < 0.05$

	Context	No Context
Previously Mentioned	3.94	5.43
Not Previously Mentioned.	5.13	5.61

a. Separate analyses of context and no context data are not given here - where of interest these are given in the accompanying text.

1. Context

Several analyses produced a significant difference between the context and no context cases. The onset-onset times showed a significant effect ($F_{1,13} = 4.89$, $p < 0.05$) with the context sentences overall 490 msec. quicker than the no context (2907 msec. vs. 3397). This effect also occurred with the offset-onset times though in that case it just failed to reach significance (1759 msec. vs. 2227 msec.; $F_{1,13} = 4.18$, $p < 0.1$). There was in addition a strong effect in the fixation data ($F_{1,13} = 10.88$, $p < 0.01$) with the context sentences averaging 0.99 fixations less than the no context (4.53 vs. 5.52 fixations). There was a slight but non-significant tendency for the stimulus materials to be briefer in the context condition ($F_{1,7} = 3.23$, n.s.). However measured in terms of number of milliseconds this effect is very small : only 40 msec. in over 1100 (1149 msec. vs. 1189 msec.), and it is certainly insufficient to account for the significant effect in the onset-onset data which is over twelve times as great (measured in milliseconds).

2. Previous mention

The analysis of materials showed a significant difference between sentences in which the object mentioned in the preamble is referred to and those in which a new object is referred to ($F_{1,7} = 40.50$, $p < 0.001$), but again in terms of time the effect is quite small vis 1136 msec. for the former to 1203 msec. for the latter. This is confounded by a Previous mention x Context interaction ($F_{1,7} = 11.20$, $p < 0.05$) which shows that the effect is primarily in the context data (1094 msec. vs. 1205 msec. with context; 1179 msec. vs. 1201 msec. without context). However none of the other reaction time data show either effect with all relevant F values extremely close to zero. The data on number of fixations do show the result, though. The previous mention main effect is highly significant ($F_{1,13} = 10.89$, $p < 0.01$) as is the context x previous mention interaction ($F_{1,13} = 5.41$, $p < 0.05$). The figures for the interaction are context, previously mentioned 3.94, context not previously mentioned 5.13, no context previously mentioned 5.43, no context not previously mentioned 5.61. These results are parallel to those for the analysis of stimulus durations.

3. Voice

The analysis of stimulus durations produced a very highly significant effect of voice ($F_{1,7}$ = 738.38, $p < 0.001$) with a mean difference of 347 msec. between actives and passives (Actives 996 msec., Passives 1343 msec.). This effect is also evident in the onset-onset times for both context and no context data. On the overall analysis there is a highly significant $F_{1,13}$ value of 20.08, $p < 0.001$. Although there is no trace of an interaction with context in the overall analysis ($F < 1$) it is apparent that the voice effect is larger with the context data (2674 msec. vs. 3139 msec. for actives and passives respectively. Corresponding figures for the no context data are 3214 msec. and 3583 msec.). This is reflected in an F value of 23.19 ($p < 0.001$) for the context data, but one of only 6.14 ($p < 0.05$) for the no context data. The answering or offset-onset times show no significant effects with an overall F value less than one. Again though the context data is suggestive of an effect ($F_{1,13}$ = 2.35, n.s.), whereas the no context data is not ($F_{1,13}$ = 0.01, n.s.).

The fixation data show a strong voice effect ($F_{1,13}$ = 20.64, $p < 0.001$), however this is again evident chiefly in the context data ($F_{1,13}$ = 23.70, $p < 0.001$) with the no context data failing to produce a significant effect ($F_{1,13}$ = 2.60, n.s.). However the context x voice interaction produced an F value of less than one. This despite the fact that the context data show a large difference between actives and passives (4.11 and 4.96 fixations respectively) while the no context data show a smaller difference even though the mean scores are much higher (5.27 vs. 5.79 fixations for actives and passives respectively).

This factor shows a close relationship between the stimulus duration, onset-onset times and number of fixations, with the offset-onset times correlated with them, but only weakly. However the evidence from the previous mention and context factors above shows that although number of fixations and stimulus duration are closely related the relation of these two to onset-onset times is quite weak.

Interpretation of this evidence on the voice effect is made more difficult by the presence of a voice x syntactic type interaction which will be considered below, after discussing the final main effect.

4. Syntactic Type

The position of the noun in the sentence shows effects in several aspects of the data. The materials analysis shows that the sentences were spoken faster by the experimenter when the noun occurs before the main verb - but only in the no context condition (Context x Syntactic Type Interaction : $F_{1,7} = 13.93$, $p < 0.01$). The effect is fairly substantial : 1243 msec. as against 1136 msec.; there is no evidence at all of this effect in the context data (1149 msec. vs. 1150 msec.), and the main effect of Syntactic type is accordingly non-significant ($F_{1,7} = 3.73$, n.s.).

The onset-onset times show a rather different pattern : the context x syntactic type interaction is non significant ($F < 1$) but there is a highly significant main effect ($F_{1,13} = 14.93$, $p < 0.01$). This is present in both context and no context data when analysed separately ($F_{1,13} = 6.42$, $p < 0.05$ and $F_{1,13} = 14.10$, $p < 0.01$ respectively), though again the no context data show the effect rather more (3108 msec. vs. 2706 msec. for context, 3675 msec. vs. 3122 msec. for no context).

The offset-onset times show much the same pattern : an overall significant F value ($F_{1,13} = 10.83$, $p < 0.01$) with a similar overall difference in time between the two (2205 msec. vs. 1780 msec.). Again the interaction with context is non-significant ($F < 1$) though the no context data show the effect a little more ($F_{1,13} = 5.63$, $p < 0.05$ for the context data and $F_{1,13} = 7.42$, $p < 0.05$ for the no context data with respective means 1960 msec. vs. 1557 msec. and 2451 msec. vs. 2003 msec.).

The fixation data show no significant effects (overall $F_{1,13} = 1.86$, n.s. Context alone $F_{1,13} = 2.60$, n.s. No Context alone $F_{1,13} = 0.53$, n.s.).

It would appear that the effects here are not due to scanning : even the longer duration of some of the sentences does not appear to lead to more fixations. Although there is the tendency in the stimulus materials for a context x syntactic type interaction, this has not produced such an effect in either the onset-onset latencies or the answering latencies. The 100 msec. or so effect in the no context data appears to have carried through (as one might expect) to the total times, but has not affected the answering times. There is an additional effect of about 400 msec. evident in both the context and no context data : this represents the difficulty subjects have with the noun later in the sentence.

5. Voice x Syntactic type

Both the interpretation of the syntactic type effect and the voice effect are affected by the presence of a syntactic type x voice interaction. This is not present in the materials ($F<1$). It comes out in the onset-onset analysis as well as the offset-onset analysis. In the onset-onset analysis it is highly significant ($F_{1,13} = 11.30$, $p<0.01$) with actives 748 msec. quicker than passives in the noun-second case but only 86 msec. faster in the noun first case. This effect is present in both the context and no context data but is very much stronger in the latter ($F_{1,13} = 3.85$, $p<0.1$ and $F_{1,13} = 16.68$, $p<0.001$ respectively). With context actives are 738 msec. faster than passives with the noun second but only 193 msec. faster with the noun first. The corresponding figures for the no context cases are 759 msec. and -21 msec.

The answering (i.e. offset-onset) times show a similar pattern. There is an overall significant voice x syntactic type interaction ($F_{1,13} = 13.79$, $p<0.01$) with actives 425 msec. faster than passives with the noun second but 275 msec. slower with the noun first. Again the effect is much stronger with the no context data. With the context data actives are 389 msec. faster than passives with the noun late in the sentence (1743 msec. vs. 2087 msec.) and 165 msec. slower when the noun is early in the sentence (1639 msec. vs. 1474 msec.). The corresponding figures for the no context data are 416 msec. (2243 msec. vs. 2659 msec.) and - 390 msec. (2195 msec. vs. 1805 msec.) ($F_{1,13}$ values are 4.52, $p < 0.1$ and 17.14, $p < 0.001$ respectively). In both the onset-onset and the offset-onset data the context x voice x syntactic type interaction fails to reach significance ($F < 1$ in both cases).

In neither the overall analysis of the fixation data nor in the separate analysis of context and no context data does this effect appear.

6. Voice x Previous mention x Syntactic type

This effect is rather complex but it is undoubtedly the most interesting result. There is no evidence of it in the analysis of stimulus materials ($F < 1$) or in the fixation data ($F < 1$ for overall analysis as well as separate context and no context analyses).

It appears only in the context data of both the onset-onset analysis and the offset-onset analysis, in the former just failing to reach significance however ($F_{1,13} = 4.54$, $p < 0.1$ and $F_{1,13} = 5.67$, $p < 0.05$ respectively). There is no trace of it in the no context data (both F's < 1) and it comes out neither as a 3-way interaction nor as a 4-way interaction with context in either of the analyses of both context and no context data together. For the onset-onset times F values for the overall voice x previous mention x syntactic type effect, and the 4-way interaction with context are 2.94 and 2.15 respectively. For the offset-onset times corresponding figures are 3.04 and 3.18. None of these reaches significance. However the means for the offset-onset times of the context data show what at first sight appear to be dramatic effects. When the noun is at the end of the sentence actives are 520 msec. faster if it refers to the object previously mentioned (1483 msec. vs. 2003 msec.) whereas passives are 416 msec. faster if it does <u>not</u> refer to the object previously mentioned. When the noun is early in the sentence (before the main verb) actives are 196 msec. faster if the noun does <u>not</u> refer to the object previously mentioned (1742 msec. vs. 1546 msec.) whereas passives are 95 msec. faster if it is previously mentioned (1427 msec. vs. 1522 msec.). The noun-second

	NOUN AFTER AUXILIARY		NOUN AFTER MAIN VERB	
	ACTIVE	PASSIVE	ACTIVE	PASSIVE
NOUN PREVIOUSLY MENTIONED	5	3	1	8
NOUN NOT PREVIOUSLY MENTIONED	2	2	2	17

"CONTEXT"

	NOUN AFTER AUXILIARY		NOUN AFTER MAIN VERB	
	ACTIVE	PASSIVE	ACTIVE	PASSIVE
NOUN PREVIOUSLY MENTIONED	4	2	4	7
NOUN NOT PREVIOUSLY MENTIONED	3	3	3	8

"NO CONTEXT"

N = 14 (2 responses per subject per cell).

TABLE 10 TOTAL NUMBER OF ERRORS.

results are particularly dramatic.

Errors

Errors (see Table 10) tended to be at an acceptably low level with one notable exception. This was sentences of the form "Who is being Ø by a?". These were difficult both in the context and no context conditions but especially so in the former where the nominal refers to the object previously mentioned. Here errors were over 60% of the total. Unlike the other three cells for this question type the context, nominal = previously mentioned cell does not have noticeably long RTs. This RT figure should therefore be treated with caution.

Discussion

The present experiment produced a noticeable difference in both speed of performance and number of fixations between the context and no context conditions, with the context condition giving rise to faster reaction times and fewer fixations. This contrasts with the earlier eye movement experiments, all three of which showed no benefit of context on RT, and two of which showed no benefit of context on number of fixations either. Presumably the prime reason for this is the exposure of the picture throughout the preamble in the present experiment, giving subjects considerably more time to study the picture in the context condition than in the no context condition. In addition, as pointed out in the introduction to the present Chapter, these questions carry rather more presuppositions than the indicatives used in the earlier experiments and this may make the no context condition relatively harder. However, the main interest of the present experiment lies in the patterning of responses in the two conditions, rather than any overall differences in the two sets of data. It is to these that we now turn. Once again I will take the adjusted (offset-onset) times as definitive and largely ignore the unadjusted times. A justification of this position is given in the previous chapter.

The analysis of all the offset-onset data produced two significant effects : a tendency for questions to be responded to faster if the noun precedes the main verb, rather than following it; and a tendency for this effect to be very much larger in the passive than in the active. In fact the size of the effect is very small in the active (only 76 msec.) so that it is doubtful whether there is a real difference at all. This result provides strong support for the prediction made in the introduction based upon an interpretation of Halliday's account of the role of the passive. There it was stated that sentences of the type "Who is being Ø by x?" are anomalous because they query the identity of the patient while at the same time having passive voice, which is a means of thematising the patient as old (shared) information. This anomaly is inherent in the linguistic options selected and should be affected relatively little by context. This appears to be the case.

At the same time though it was predicted that previous mention would have some effect on these times and although the four way interaction with context failed to reach significance, separate analysis of the context and no context data does provide support for this position. There was a significant interaction between the syntactic type, voice and previous mention factors in the context data. As predicted passives are easier when the patient is mentioned in the question if this is the previously mentioned item, and easier if the actor is mentioned in the question if this is not the previously mentioned item. Again questions with the patient mentioned in the question are very much easier than those with the actor mentioned in the question (a result which is supported by the very high error rates with the latter). This is all as predicted from the interpretation of Halliday's account of the role of the passive (see the introduction to this chapter). The fact that the no context sentences show little effect of previous mention (and what there is in the opposite direction to the context sentences) provides further support for this position.

The analysis presented in the Introduction is rather less successful with the active questions. It predicted no differences between the various actives. However there do appear to be some differences. As noted already there is very little difference between active questions in which the noun precedes the main verb and those in which it follows it in terms of overall average. The previous mention factor appears to have a different effect in the one case than in the other though. With sentences of the form "Who is Ø a?", RTs are briefer if a refers to the previously mentioned item (i.e. the object in the middle of the picture), whereas with sentences like "Who is a Ø?", RTs are briefer if a refers to the item not previously mentioned (i.e. the object behind the object in the centre of the picture).

What this amounts to is that RTs are shorter if (1) the object referred to in the preamble/in the centre of the picture is the patient, rather than the actor; (2) RTs are shorter if the second relevant object is behind rather than in front of the object in the centre. Note that this second result is just the reverse of the effect Hall found with her children's fixation data, viz. that they found it easier to look to the front than to the rear. The first result is very similar to the familiar Huttenlocher/Clark result that sentences are easier when the new object is actor/the old object is the reference point. It is worth emphasising that the results from the passives are very different. In the no context case although passives are easier when the questioned element is the actor (this is the difference between noun before main verb and noun after main verb sentences), this has no relation to previous mention/position in the picture. The passive with context case has already been discussed and does not fit such a simple model as that it is easier to answer the question when the questioned element is the actor, or the object fixated after the centre object. Furthermore this simple model would seem to predict that actives should be easier when the noun follows the main verb (i.e. the questioned element is actor) but the data show, if anything, the opposite result.

To summarise this discussion so far :

(1) the explanation in terms of the function of the passive fully explains the passive results, but fails to explain apparent differences between actives.

(2) the two derivations from the Clark/Huttenlocher effect (viz. that sentences are easier when the new item is actor/old item is patient, and that sentences are easier when the first object fixated/preamble object is patient - and hence the second (new?) object actor) is only partly supported by the active data, but not supported at all by the pattern of results in the passive.

Essentially the same picture emerges from the analysis of onset-onset times as from these answering times. Indeed the only major difference between the two analyses concerns the voice effect : passives take significantly longer than actives if all the time from the onset of the question to the onset of the response is used as the measure, but there is a non-significant trend in the opposite direction if only the answering times are used. The fixation data give a quite different picture from either of these analyses. The voice effect is significant in the same direction as the onset-onset times. This is not surprising as one would expect some correlation between number of fixations and RT and such a big difference in sentence length as there is between active and passive should clearly have some effect on number of fixations. But the remaining fixation data show quite clearly that the positive correlation between fixations and RT is quite a weak one. There is a substantial effect of whether the object referred to explicitly in the question is the previously mentioned object or not, with approximately 13% fewer fixations if it is. This effect is much larger with context than without, though (23% fewer with the former but only 3% fewer with the latter). The influence of the context factor here would seem to support the view that the presence of a preamble is the primary topicalising device here, the absence of even a slight effect with the no context data perhaps indicating that there is no visual prominence associated with central position in the picture. However one should beware of associating this effect directly with linguistic topicalisation given the absence of this effect in the RT data. What is perhaps most noticeable about these data is the remarkably large number of fixations which occur in the context data when the explicit reference in the question (i.e. not the reference of the Wh- item) is not to the object in the preamble/centre of the picture. It is as though they had not had any time to look at the picture prior to the onset of the question - the number of fixations being quite close to the no context data. This is very strong evidence of the importance of the sentence being matched in terms of topic to the situation. Note, though, that this effect seems to be independent of the case role of the topic in the sentence (contrary to what Huttenlocher's hypothesis would suggest). It seems a little odd that an effect of this size should not be evident at all in the RT data, even given the obviously weak relationship between number of fixations and RT. Partly this is because despite the size of the differences in means the interaction is only significant at the five per cent level, indicating very large variances. Partly it may also be due to an ability of people to rapidly restructure a scan on the basis of reference made in an accompanying sentence (this must surely happen all the time in interactions between adults and young children while playing). This hypothesis is readily testable by using the paradigm of the experiment reported in Chapter Two, only presenting the picture after the subject has indicated that he has understood the sentence, and using the comprehension time as the measure.

One quite interesting aspect of the present results is the uneven distribution of errors. 54% of the errors were on sentences of the type "Who is being ∅ by a?", 62.5% of the context errors and 44% of the no context errors being on sentences of this type. What is more there seems to be a clear influence of previous mention in the context case - more errors occurring if the nominal refers to the object not previously mentioned than if it refers to the previously mentioned object. Subjects seem overwhemingly to treat this sentence as "Who is ∅ a?" in the case where a ≠ previously mentioned item, but are slightly less consistent when a = previously mentioned item. There were 17 errors in the 28 responses to the context case where a ≠ previously mentioned item, and of these 11 (64%) were correct answers to "Who is ∅ a?", 2 (12%) produced no answer and 4 (24%) produced some other incorrect answer. Equivalent figures for the no context case are 8 errors, 7 wrong answers of the first sort and one other wrong answer. I have no explanation for this pattern of results, but will simply make three comments : (1) this large number of errors suggests one should treat the

RT data with caution[2]; (2) the errors occur on the sentence type predicted to be most difficult, so this supports the explanation put forward of the importance of the functional role of the passive; (3) if subjects are processing the sentence as ""Who is Ø a?" they would appear to be correctly processing surface order information but largely ignoring voice information, which again brings out the importance of order information in processing a sentence.

Finally a note on the analysis of stimulus durations. There are a number of interesting points here. Firstly of course there is the substantial difference between actives and passives in terms of length of the utterance. The difference here (347 msec.) is rather smaller than that found with indicatives in the previous experiment (458 msec.), however it is again very substantial and is strong evidence that great care should be taken in comparing active and passive in terms of some unqualified notion of absolute processing difficulty. An equally interesting result is the finding that the experimenter spoke the sentences faster when the explicit reference in the question was to the previously mentioned item - but really only in the context case. The only explanation for this effect seems to be that the speaker makes allowances for the hearer's knowledge in uttering the sentence and says it more slowly if it is known to consist entirely of new information for the hearer. It is possible that the old information part of the sentence alone is shortened, rather than an overall slowing down taking place. This is an interesting possibility and would repay further study. Of

2. With regard to this point however there is no consistent difference in RT between correct and incorrect responses for this sentence type viz.:-

Context : Previous mention : Correct : 1634 Incorrect : 3140

Context : No Previous mention : Correct : 2027 Incorrect : 1915

No Context : Previous mention : Correct : 2838 Incorrect : 2382

No Context : No Previous mention : Correct : 2407 Incorrect : 3009.

course it is essentially a part of the vexed question of the nature of stress - duration being widely considered to be a major component of the complex notion of stress. On that basis one would expect additional duration only in the stressed part - i.e. only in the new information part - of the clause. The third effect in the stimulus material analysis - that sentences with the noun preceding the main verb are spoken faster by the experimenter, but only in the no context condition - is very odd and I can think of no explanation for it.

Chapter 7 : Conclusions.

In this chapter I attempt to do two things. Firstly I give a small number of substantive conclusions drawn from the whole series of experiments. In doing this I try to cut through the mass of detail surrounding the experiments and produce some fairly general statements. Partly as an antidote to excessive generality, and partly because a series of experiments with so many diverse tasks demands a general discussion of the relationship between the tasks, in a second section I discuss the tasks and measures used and attempt to compare them. This thesis is already much too long so I have attempted to keep this final chapter to a bare minimum.

Substantive Conclusions.

Of the nine experiments reported in this thesis seven include some attempt to manipulate sentence comprehension or production through having subjects read a chunk of text, or produce a chunk of text, prior to the target sentence. There can be no doubt that this manipulation - or family of manipulations - has a profound effect on the way the sentence is processed, or on the structure of the sentence which people produce. There is a natural inclination to assimilate this result to those of Olson and Filby (1972) and Wright (1969) which show that sentence processing is affected by the presence of some sort of previously stored code with which one has to compare the coding of the sentence. But in none of the present experiments does the picture precede the sentence so it cannot be (as in Olson and Filby's experiments) that the sentence is being compared to some previously encoded picture. Nor is it being compared to some other sentence (as in Wright's experiment) since the sentences of the preamble are not really comparable to the target sentence in this way. If anything is being compared with anything it is the actual sentence produced with the set of alternative structures which could have been used to convey truth-conditionally equivalent information,

Chapter 7 : Conclusions.

In this chapter I attempt to do two things. Firstly I give a small number of substantive conclusions drawn from the whole series of experiments. In doing this I try to cut through the mass of detail surrounding the experiments and produce some fairly general statements. Partly as an antidote to excessive generality, and partly because a series of experiments with so many diverse tasks demands a general discussion of the relationship between the tasks, in a second section I discuss the tasks and measures used and attempt to compare them. This thesis is already much too long so I have attempted to keep this final chapter to a bare minimum.

Substantive Conclusions.

Of the nine experiments reported in this thesis seven include some attempt to manipulate sentence comprehension or production through having subjects read a chunk of text, or produce a chunk of text, prior to the target sentence. There can be no doubt that this manipulation - or family of manipulations - has a profound effect on the way the sentence is processed, or on the structure of the sentence which people produce. There is a natural inclination to assimilate this result to those of Olson and Filby (1972) and Wright (1969) which show that sentence processing is affected by the presence of some sort of previously stored code with which one has to compare the coding of the sentence. But in none of the present experiments does the picture precede the sentence so it cannot be (as in Olson and Filby's experiments) that the sentence is being compared to some previously encoded picture. Nor is it being compared to some other sentence (as in Wright's experiment) since the sentences of the preamble are not really comparable to the target sentence in this way. If anything is being compared with anything it is the actual sentence produced with the set of alternative structures which could have been used to convey truth-conditionally equivalent information, in the light of the topic of the prior discourse. It is possible that this is also what people do in the Olson and Filby experiment - that the major influence on the reaction times which they observed was not the processes involved in comparing the two codes to produce a truth judgement, but rather the comparison of what was said with what might have been said given the way they focussed their attention. Of course stating it this way makes the process sound far too abstract and mechanical. I am not suggesting that people compare the sentence with a list of alternative formulations - merely that in processing the sentence they are aware (at some level) that it could have been put differently, that meaningful choices are involved.

There are many objections to this view. For the moment I will just deal with one straightforward one. It goes like this : in Experiments 1, 2, 3, 6, 8 and 9 (i.e. all but one of the relevant experiments) people will encode the picture with a "voice" because the object in the centre of the picture will be the natural topic of the picture coding. Since in all but one of these experiments no measure is taken until after the picture has been presented it is still possible for the effect to be substantially due to the comparison process which Olson and Filby believe is involved. The reply to this is simple : (1) nevertheless in most of these experiments there is a strong influence of the text factor which should not occur if it were simply a case of the structure of the picture determining the effects; (2) there are clear effects in the comprehension data of Experiment 1 - effects which could not be due to the picture coding since this measure is taken prior to the presentation of the picture. There may be effects of picture coding in addition to the effects of the preamble, but the fact remains that many effects are due to the preamble. If we are to assimilate the present results to Olson and Filby's it will have to be on the basis that similar processes of interpretation occur in the two experiments because of the presence of a topic - not on the basis that the two sets of experiments both involve comparing different codes.

Many of the experiments reported here show the importance of the position of the topic in the sentence in determining both production frequencies and reaction times in the comprehension tasks. This factor interacts with several others, however, so that the picture is quite a complex one. Huttenlocher's description of the major effect in her placement tasks as due to a correspondence/non-correspondence between the perceived actor (or new item) and the logical subject of the sentence provides an extremely good first approximation to the results observed here.[1] Reaction times tend to be faster to the copular sentences when the new item is grammatical subject (i.e. first NP with unmarked syntax and second NP with marked syntax) and to transitive sentences when the new item is actor (i.e. surface subject with actives and object with passives.) (Production frequencies are inversely related to reaction times) This characterisation is only an approximation though since the effect is very much larger with marked syntax. That is to say that the difference in RT between the new item = grammatical subject/actor and the new item = grammatical object/patient is very much greater with passives and copular sentences in which the locative phrase comes first than with actives and copular sentences in which the locative phrase comes later. Hence the Clark and Huttenlocher characterisations are in need of modification. The explanation which has been repeatedly put forward here rests upon an account of the distribution of new and old information in the clause and the function of certain grammatical constructions in realising unusual configurations of new and old information. The passive is seen as a way of maintaining a focus on an object which happens to be the recipient of a given action. It maintains the usual given/new structure - namely given information early in the clause and new information later - but at the

1 As repeatedly noted this is equivalent to Clark's characterisation of the effect as due to the patient in transitive sentences and the locative phrase in copular sentences being the natural reference point.

expense of a marked voice option. Psychologically this option is no harder to process than the "equivalent" active - so long as it can be seen to be motivated. But if its selection does not make sense then people have difficulty in understanding it. For example if there is no previously-focussed-upon object then the function of maintaining focus upon that object is simply irrelevant. This account of the passive has the additional advantage that the so-called shortened passive can be seen as less complex than the full passive which is merely the shortened form plus an extra piece of new information. This analysis of the full passive as more highly marked than the shortened form is in accord with the measured frequency of the two forms - the shortened form being more frequent.

The copular sentence with locative phrase preceding the copular is related to the passive in that it too serves to produce an unmarked configuration of new and old information. As the oral description study shows quite clearly it is more complex than this though. In the passive sentence-initial position (theme), grammatical (modal) subject, and given information are all realised in the same surface item. In the marked copular construction only theme and given information are associated, and the modal subject role is associated with new information. This has the consequence that the next clause is likely to be about the new information of the present clause since for some reason - as yet unexplained - the subsequent clause seems more likely to be about the grammatical subject of the present clause than about its grammatical object. Hence the marked copular construction essentially performs a topic switching role in the overall structure of a text. This may in part explain why there remains some residual difficulty in understanding this construction in the context of Experiment 1 - even when the topic is in the locative phrase. Subjects know that the relational sentence is the final sentence of the text and hence that it cannot be performing its topic-switching function. Hence they can never completely justify its use to themselves. One can easily check this conclusion by embedding sentences of this sort in the middle of longer pieces of text and measuring comprehension times to each sentence in the manner of Experiments 1, 4 and 5.

Given the explanation of Huttenlocher's effect in terms of new and old information and the demonstration that it is not independent of syntax, where this is also governed by information structure, it comes as no surprise to discover that the effect is partly dependent upon another information-structural device, namely pronominalisation. It was clear in examining the three term series problem experiment (Experiment 5) that the advantage of having the new item first in the second premise is increased if the mode of crossreference to the first premise is by means of a pronoun. The new item position effect was here explained as being due to the subject's knowledge that a third item would be mentioned and his need to "solve for" this item, combined with the extra prominence which derives from sentence-initial placement. When the object which he is most interested in is referred to by the nominal in the most prominent position then the sentence is easier to process than if it is referred to by the other nominal. At the same time though there is a difficulty arising from the fact that the normal order of information in the clause is old then new. Where pronouns are used to cross refer this conflict is rapidly resolved, but where they are not the subject may have to check with what he already has stored to confirm that the information structure of the clause is new then old.

Although Experiment 4 seemed to show a simple speeding up of comprehension where pronouns are used to crossrefer rather than proper names or other definite noun phrases, several of the experiments demonstrate that the situation is not as simple as that. Firstly there is the fact just referred to that the advantage of having the new item as grammatical subject is affected by the use of pronouns. Secondly there is the fact, demonstrated also in Experiment 5, that the position of the coreferential nominal in the preceding sentence has a clear effect on the speed of comprehension of the sentence including the pronoun. That is : when the pronoun is coreferential with the grammatical subject of the previous sentence, the sentence including the pronoun is understood more rapidly than when the pronoun is coreferential with the object of the previous sentence. This effect is clearly related to the fact, noted in Experiment 3, that a subsequent sentence is more likely to be about the grammatical subject of the present sentence than about the grammatical object (though, as we have already seen, this is affected by syntactic type). A third factor affecting pronoun use was noted in Experiment 3 : namely the tendency for pronouns to be coreferential with a nominal in the immediately preceding sentence. If anaphora extends over a larger interval then nominals marked with "the" tend to be used instead of pronouns. Of course there must be many other constraints on pronoun use : for example the number of nominals in the preceding sentence and whether a pronoun can be used to pick one of them out uniquely on the basis of gender. These phenomena were not investigated in the present series of experiments, however.

The use of the definite article is obviously not unrelated to pronoun use. Although in all the comprehension experiments, except Experiments 4 and 5, definitely marked noun phrases were used where pronouns would clearly have been appropriate, the oral description study showed remarkably well that pronouns are used if subjects are not constrained to use the definite article. It is hard to assess in detail quite what effect this restriction had on reaction times, but it seems clear that it must have had some effect and that this is unlikely to have been entirely independent of syntactic type. More studies are needed to investigate this. A very noticeable feature of Experiment 1 was the lack of any simple effects of definiteness marking on subjects' reaction times. This is in sharp contrast to the undoubtedly very strong influence of this factor on subjects' responses in the production experiments. This result seems likely to have been an artifact of the design of Experiment 1.

Nominals were as likely to be marked "incorrectly" as "correctly" and this must surely have encouraged people to ignore definiteness marking as far as possible. This phenomenon is in need of further investigation - especially in view of the abundant evidence for the importance of this factor in production experiments.

One of the dominant interests of the first half of this thesis is the nature of lexical marking. On the whole the evidence supports the interpretation outlined in Chapter 1 based upon the good reason principle. Thus Experiment 1 shows the same effect as observed by Clark in the no context case (viz : the unmarked term "in front of" is reacted to faster than the marked term "behind"), but no difference is observed in the context case. Related to this, subjects reacted faster to sentences with "behind" when the two nominals were marked differently, whereas with "in front of" RTs tended to be faster when both nominals were marked the same. This was interpreted as due to a tendency to pick the marked term for topicalisation reasons - i.e. when the two nominals are different in importance.[2] However it is presumably not a very strong tendency since there was no real trace of this trend in the production experiments. Of more interest are the data from Experiment 5. These support the position that (1) the selection of a marked term without any obvious reason will lead to longer RTs - probably because of the fact that this choice is interpreted as being due to a desire to convey absolute information (this gives rise to main effects of first premise marking in both the second premise and question times : we said it is unlikely that there could be a genuine topicalisation principle for picking a marked term at the start of a discourse); (2) the selection of a marked term in the presence of a possible topicalisation justification will not lead to longer RTs (hence the absence of a main effect of second premise marking and the tendency

2 Note that there is another explanation based on preferred directions for building displays which cannot be discounted - see Experiment 3.

for first premise marked, second unmarked to be especially difficult).

Both these results support the "good reason" interpretation of marking. As a general principle this is given further support by the results for copular sentences with marked and unmarked syntax, and those for transitive sentences in active and passive voices. These have already been referred to. Unfortunately this principle may present difficulties when it comes to formulating lexical entries. At the present time it seems reasonable to think in terms of lexically marked/unmarked pairs as really being triples composed of a neutral superordinate and two non-neutral subordinates, one of which happens to be homonymous with the superordinate. But in the first place it seems strange to apply this analysis to prepositions and in the second place we are here arguing for neutralisation of the subordinates in the presence of topicalisation choices. This second fact clearly presents problems for a feature theory in which a set of features are always attached to a word since several different levels of description would need to be involved in coping with this neutralisation.

The analysis of lexical marking which one adopts is intimately related to one's approach towards a more general issue : that of canonical form representations. A fixed-feature analysis is obviously more in the spirit of canonical form theories than a view which sees the meaning of a word as very highly environmentally conditioned. It is in the nature of evidence bearing on such a fundamental kind of precept as Clark's view of the importance of canonical form that it should be indirect. This is certainly true of the evidence presented in the current series of experiments. But taken as a whole it does tend to undermine Clark's precept. The presence of several effects involving the text factor, in the verification times of Experiment 1 is very important in this respect. Evidence demonstrating that surface form has an influence on reaction times after even a long delay is not terribly damaging to canonical form theories since they can always accept two-trace (surface as well as deep) models. But evidence that factors which are in one sense deep, but are not represented in the canonical form - namely topicalisation features - can influence verification times, even after a delay, is surely quite serious for canonical form theories. In addition it seems to contradict Clark's principle of the primacy of functional relations which explicitly denies the importance of topicalisation factors. Also the influence of pronouns in solving the three term series problem can surely not be accounted for by any standard canonical form theory - especially in view of the fact that the pronoun/name factor seems to interact with other factors. Of course one could still accept that canonical forms are important at some level and appeal to superficial decoding strategies (of the sort discussed by Bever, 1970) to explain the data. But what then is the point of talking of canonical form? There can be little doubt that in some tasks with only a few parameters, highly practised subjects develop strategies based on canonical forms. But the evidence suggests that in much language processing they are either not used at all or else so deeply embedded in a complex of other processes as to make analysis in terms of them at best unhelpful and at worst highly misleading.

Methodological Conclusions.

The experiments reported in this thesis use quite a variety of different methods and four different measures. These measures are: (1) error rates. Although almost all the experiments (the exceptions being Experiments 2, 3 and 4) produce error data, I do not go into this in very much detail at any point, although in some cases I have presented an analysis of the errors. The reasons for this are twofold. Firstly I have concentrated to a large extent on keeping the number of trials any one subject has to undertake to a minimum. This is primarily in order to avoid the development of special strategies (in so far as this is possible). The result of this policy is that I have insufficient data to perform adequate analyses on the error data. The second reason is that, so far as I am able to judge on the basis of the error data I have, the number of errors varies directly with reaction time. Given this state of affairs, it seems unnecessary to discuss both measures since for the most part they will give the same results, and any differences will only tend to add to the complexity of an already very complex picture.

(2) fixations. This measure is used in four experiments only. It is only a crude measure as used here - being basically only a measure of the number of fixations involved in scanning the picture. The relationship between number of fixations and reaction time to the sentence is, at best, extremely unclear. It is obvious that structural properties of the picture must affect number of fixations (see for example the much larger number in Experiment 8 where three objects were depicted compared to Experiments 6 and 7 where only two were pictured). The amount of uncertainty is also important (for example in the text conditions of Experiments 6 and 8 subjects know where one object will be and this accordingly reduces the number of fixations). But how the structure of the sentence is related to the number of fixations is an almost total mystery. Hall's (1975) experiments tended to show a close direct relationship between number of fixations and sentence difficulty as measured by RT, but this is not found in the present experiments. It may only have been true of her experiments because of the fact that her subjects always knew the position of the sentence actor in the picture. In other words she may not have sampled enough of the possible combinations of previous knowledge/picture structure/sentence structure/truth and so have observed an artifactual concurrence of RT and number of fixations measures (see Introduction to Chapter 5). It seems likely that studies in which picture and sentence are presented simultaneously, and the object on which S is fixating at any moment is compared with a moment by moment analysis of the sentence, will give a much fuller picture. This could be done in both an explicit verification task (where S has to say whether the sentence is true or false) and in a kind of comprehension task where, for example, S is told he has to remember the text/picture and the picture/text is there to assist him. The sentence could be embedded in long pieces of

prose or not, etc. This kind of study, though technically very difficult, should well be within the capabilities of many psychology laboratories, and may well be very revealing. The gross number of eye movements measure clearly is not, though.

(3) production frequencies. This measure provided remarkably clean data in both of the studies in which it was used. This is especially true of Experiment 2 in which subjects were severely constrained as to what they could write in order to produce a fairly broad range of responses. However this method means that subjects may be approaching the sentence generation problem in an extremely artificial way. This may not be a bad thing if what we are interested in is the traditional basis of grammatical theory : namely intuitions about sentencehood and grammaticality. But the relationship of these to natural language processing is partly what we are trying to investigate and we cannot assume that this kind of highly artificial exercise is an example of natural sentence construction. The much less artificial situation of Experiment 3 cannot be criticised on this count. The open-ended nature of the task revealed a number of facts which would not have turned up otherwise. As a general method though it has the disadvantage that one has very little control over the situation. One consequence of this is that one only observes a rather restricted range of the possible utterances and so information on the relative difficulty of the more improbable cases is, to say the least, sparse. Of course there is the additional fact that the relationship between comprehension difficulty and production frequency is problematic, though this is partly overcome by the fact that the results of Experiments 1 - 3 fit quite well with one another.

(4) reaction times. This is by far the most important measure used in this series of experiments - all but two experiments using at least one reaction time measure. The measures taken differed considerably, though. Comprehension times were taken in Experiments 1, 4 and 5, verification times in Experiment 1, comprehension + Verification times in Experiments 6,

7 and 8 and question-answering times in Experiments 5 and 9. What is more there are important features of experimental design which make all of these times unique - with the exception of those for Experiments 6 and 7 which are very similar to one another. Furthermore there are aspects of the question-answering times of Experiment 9 which make them similar to the comprehension + verification times of Experiment 8. Looking over all the experiments it seems true to say that the comprehension + verification times produced the least clear results, while the comprehension-only times produced the most clear (though question-answering results are also reasonably clear). While the relationship between these comprehension times and the other sorts of RT measures is unclear it does seem unlikely that the comprehension measure is unrelated to these others. This for two reasons : (1) all the tasks must involve some sort of comprehension or relatively deep processing - otherwise they have not done what they were intended to do (see Chapter 1); (2) the "pure" comprehension times must be "contaminated" by some of the elements contributing to the other times since the comprehension times only measure part of a task which also includes in one case verification and in another question-answering. Indeed it may be a mistake to suppose that there could be any such thing as "pure comprehension" or something which is a sub-task of every task involving language processing. But given the results derived from the various methods used here, it seems likely that investigation of language processing using much less rigorous methods of checking for understanding than verification and question answering - for example the kind of sentence by sentence presentation with instructions to "press when you understand", used in Experiments 1, 4 and 5 as well as by Clark in his recent work (e.g. Clark and Haviland, 1976) - may be more fruitful in the long run. If we are to continue using verification tasks the method used in Experiment 1 seems to be the most satisfactory. Even if the separate measures are not really measures of comprehension and verification, we are introducing additional sampling into the process and may as a result succeed in getting

a fuller picture of what is occurring. The use of a fixation measure is of course another step in the same direction. However the bald number of fixations measure is too gross, as we have already noted, and the most promising method here would seem to be one in which sentences are presented orally and simultaneous with the picture so that fixations can be directly correlated with sentence structure. Successive presentation of sentence and picture, though in general more desirable from the point of view of reaction time measures, seems likely to encourage heavily strategy-laden scanning strategies, especially in cases where there is a delay between offset of the sentence and onset of the picture so S can formulate a scanning strategy (and in an interval not being sampled, at that!). Of course a major difficulty with simultaneous presentation of sentence and picture is that one is faced with the decision as to what reaction time measure to take. Should we use the total time from the onset of the sentence to the onset of the response or is the time from the offset of the sentence to the onset of the response a more legitimate measure? As I have indicated the answer to this question depends partly on the general theory one wishes to apply. A theory in which sentence decoding is seen as involving the use of operators working on whole strings is best tested by the use of sentence offset - response onset times. But for other models it is not clear what is the appropriate measure. However it is clear that the use of onset-onset times will inevitably lead one to the conclusion that the passive is harder to process that the active, because of its greater length. On the other hand offset-onset times may underestimate its difficulty. There is no straightforward solution to this dilemma - partly because it leads on to questions about the meaningfulness of comparing different sentence structures as to absolute difficulty - questions which cut deep into the heart of one's theory of language. These are important issues, and ones which deserve considerably more debate than they usually receive.

Most of the experiments reported in this thesis have used a multifactorial design. In some cases this has lead to results so complex that it is almost impossible to conceptualise them. A natural response to this is to question the use of such complex designs : why waste one's time producing uninterpretable data? The simple answer is : because the phenomena one is studying are so complex one needs complex designs. The very fact that such complicated effects are observed is a justification of the design. Even if five way interactions are uninterpretable, two, three and four way interactions are not and by using simpler designs which only throw up these lower order interactions one is, in a way, burying one's head in the sand. In any case the experiment may only produce simple effects in which case one has learnt quite a lot more than would have been discovered in a simpler design producing the same effects. Of course there is another side to this argument : it may be that by using complex designs one is manufacturing complex results because of the strategies which subjects develop for dealing with the set of possibilities with which they are presented. I have suggested that the failure to observe any simple effects of definiteness marking in Experiment 1 is due to a strategy developed to cope with the fact that the design ensures that nominals are as likely to be marked correctly as incorrectly. This kind of thing is undoubtedly a problem, but on the whole I feel that a research strategy which starts from complex multifactorial designs and then goes on to use simpler designs to examine possible artifacts, is likely to be more efficient than one which works the other way around.

One major decision in applying complex designs is which factors to make within-subjects ones and which between-subjects. One important consideration in the present series was the desire to keep subjects as naive and unpractised as possible (consistent with efficient use of subjects, equipment etc.) and this necessitated between-subjects factors in several experiments. However, wherever possible, within-subjects factors were used. This may have been a mistake. Experiments 1 and 5 in which the major factors (context/no context and pronoun/name respectively) were between-subjects gave clearer results where those factors were involved than Experiments 6 - 9 which used entirely within-subjects designs. This may be because the approach developed in one condition is carried over to the other condition and so minimises any difference between the conditions. On the other hand Experiments 1 and 5 used the "press when you understand" technique with written materials, while Experiments 6 - 9 used oral presentation at a rate controlled entirely by the experimenter. It is possible that oral presentation is less subject to strategy effects than written presentation and in that way minimises differences between the conditions. These questions need further examination.

Finally, one more comment on oral presentation. A major problem with using spoken material is the control of intonation contours. This is particularly true where one is interested in the way subjects process new and old information in the sentence since intonation is the major means of expounding these options - of which there are a very large number (Halliday, 1967a,b). There is no doubt that insufficient control was exercised over this parameter but it is difficult to see how this can be remedied. It is possible that the experimenter's efforts to avoid very expressive intonation contours which would alone be sufficient to indicate new and old information in the sentence, was a contributory factor in producing reaction times to the questions of Experiment 9 which showed a pattern similar to what one might expect from statements. Of course the decision to use a reasonably flat tone is itself not really a neutral one. But then there is no really neutral case.

This last point is yet another example of a problem which has recurred at intervals throughout this thesis. If one rejects ideas of canonical forms, kernel structures and so on, everything becomes relative and impossible to assess without considering several other parameters. It is like trying to cross a landscape where the reference points constantly move.

It is not an easy task but I hope I have shown that we can, assisted by some of the concepts of systemic grammar, at least make an attempt at it.

It is not an easy task but I hope I have shown that we can, assisted by some of the concepts of systemic grammar, at least make an attempt at it.

It is not an easy task but I hope I have shown that we can, assisted by some of the concepts of systemic grammar, at least make an attempt at it.

References.

Anderson, J. (1974). The Grammar of Case : Towards a Localistic Theory. London : Cambridge University Press.

Anderson, J.R. (1974). "Verbatim and propositional representation of sentences in immediate and long term memory". Journal of Verbal Learning and Verbal Behaviour, 13, 149 - 162.

______ and Bower, G.H. (1973). Human Associative Memory. Washington, D.C. : V.H. Winston.

Anisfeld, M., and Klenbort, I. (1973). "On the functions of structural paraphrase : the view from the passive voice." Psychological Bulletin, 79, 117 - 126.

Austin, J.L. (1962). How to Do Things with Words. Edited by J.O. Urmson. Oxford : University Press.

Barclay, J.R. (1973). "The role of comprehension in remembering sentences". Cognitive Psychology, 4, 229 - 254.

Begg, I., and Paivio, A. (1969) "Concreteness and imagery in sentence meaning." Journal of Verbal Learning and Verbal Behaviour, 8, 821 - 827.

**

Bierwisch, M. (1969). "On certain problems of semantic representation". Foundations of Language, 5, 153 - 184.

______ (1970). "Semantics". In J. Lyons (Ed.) New Horizons in Linguistics, 166 - 184. Harmondsworth : Penguin Books.

Bierwisch, M. (1971) "On classifying semantic features". In D.D. Steinberg and L.A. Jakobovits (Eds.) Semantics. 410 - 435. London : Holt, Rinehart and Winston.

Bolinger, D. (1972). That's That. The Hague : Mouton.

Bransford, J.D., and Franks, J.J. (1973). "The abstraction of linguistic ideas : a review". Cognition, 1, 211 - 249.

Brewer, W.F., and Lichtenstein, E.H. (1974). "Memory for marked semantic features versus memory for meaning". Journal of Verbal Learning and Verbal Behaviour, 13, 172 - 180.

Bruner, J.S. (1973). "The ontogenesis of speech acts". Unpublished manuscript, Oxford University.

__________ (1975a). "From communication to language - a psychological perspective". Cognition, 3, 255 - 288.

__________ (1975b) "Pragmatic deep structures in language acquisition". Paper given at British Psychological Society Conference Language in Social Context at Stirling University, January, 1975.

Brunswik, Egon. (1950). The Conceptual Framework of Psychology International Encyclopaedia of Unified Science. Volume 1, Number 10. Chicago : University of Chicago Press.

Campbell, R.N., and Wales, R.J. (1969) "Comparative structures in English". Journal of Linguistics, 5, 215 - 251.

Carpenter, P.A., and Just, M.A. (1972). "Semantic control of eye movements in picture scanning during sentence-picture verification". Perception and Psychophysics, 12, 61 - 64.

__________ (1975). "Sentence comprehension : a psycholinguistic processing model of verification". Psychological Review, 82, 45 - 73.

Chafe, W.L. (1970). Meaning and the Structure of Language. Chicago : University of Chicago Press.

Chase, W.G., and Clark, H.H. (1971). "Semantics in the perception of verticality". British Journal of Psychology, 62, 311 - 326.

__________ (1972). "Mental operations in the comparison of sentences and pictures". In L. Gregg (Ed.), Cognition in Learning and Memory. New York : Wiley.

Chomsky, A.N. (1957). Syntactic Structures. The Hague : Mouton.

__________ (1964). Current Issues in Linguistic Theory. The Hague : Mouton.

__________ (1965). Aspects of the Theory of Syntax. Cambridge, Mass. : M.I.T. Press.

__________ (1971). "Deep structure, surface structure and semantic interpretation". In D.D. Steinberg and L.A. Jakobovits (Eds.), Semantics, 183 - 216. London : Cambridge University Press.

Chomsky, and Halle, M. (1968). The Sound Pattern of English. Cambridge, Mass. : M.I.T. Press.

Clark, E.V. (1973). "What's in a word? On the child's acquisition of semantics in his first language". In T.E. Moore (Ed.), Cognitive Development and the Acquisition of Language. London : Academic Press.

Clark, H.H. (1965). "Some structural properties of simple active and passive sentences". Journal of Verbal Learning and Verbal Behaviour, 4, 365 - 370.

________ (1969a). "Linguistic processes in deductive reasoning". Psychological Review, 76, 387 - 404.

________ (1969b). "The influence of language in solving three term series problems". Journal of Experimental Psychology, 82, 205 - 215.

________ (1970). "Word associations and linguistic theory". In J. Lyons (Ed.), New Horizons in Linguistics, 271 - 286. Harmondsworth : Penguin Books.

________ (1971). "More on 'Adjectives, Comparatives and Syllogisms' : A reply to Huttenlocher and Higgins". Psychological Review, 78, 505 - 514.

________ (1972). "Difficulties people have in answering the question 'Where is it?'". Journal of Verbal Learning and Verbal Behaviour, 11, 265 - 277.

Clark, H.H. (1973a). "Space, time, semantics and the child". In T.E. Moore (Ed.), Cognitive Development and the Acquisition of Language. New York : Academic Press.

________ (1973b). "Comprehension and the Given/New Contract". Paper given at conference "Die Rolle der Grammatik in der interdisciplinaren linguistischen Forschung", Universität Bielefeld, December 1973.

________ (1973c). "The Language-as-Fixed-Effect Fallacy : a critique of language statistics in psychological research". Journal of Verbal Learning and Verbal Behaviour, 12, 335 - 359.

________ (1974). "Semantics and Comprehension". In T.A. Sebeok (Ed.), Current Trends in Linguistics Volume 12 : Linguistics and Adjacent Arts and Sciences, 1291 - 1428. The Hague : Mouton.

________ (1976). Paper given to Stirling University Psychology Department, May 1976.

________, and Begun (1968). "The use of syntax in understanding sentences". British Journal of Psychology, 59, 219 - 229.

________ and Begun (1971). "The semantics of sentence subjects". Language and Speech, 14, 34 - 46.

________ and Clark, E.V. (1968). "Semantic distinctions and memory for complex sentences". Quarterly Journal of Experimental Psychology, 20, 129 - 138.

Clark, H.H. and Chase, W.G. (1972). "On the process of comparing sentences against pictures". Cognitive Psychology, 3, 472 - 517.

_________ and Haviland, S. (1976). "Comprehension and the Given-New Contract". In R. Freedle (Ed.), Discourse Production and Comprehension. Hillside, N.J. : Lawrence Erlbaum, In press.

_________ Carpenter, P.A., and Just, M.A. (1974) "On the meeting of semantics and perception". In W.G.Chase (Ed.), Visual Information Processing. London : Academic Press.

Coleman, E.B. (1965). "Learning prose written in four grammatical transformations". Journal of Applied Psychology, 49, 332 - 341.

Collins, A.M., and Quillian, M.R. (1969). "Retrieval time from semantic memory". Journal of Verbal Learning and Verbal Behaviour, 8, 240 - 247.

Danes, F. (1964). "A three-level approach to syntax". Travaux Linguistiques de Prague, 1, 225 - 240.

Davidson, D. (1970). "Semantics for natural languages". In Linguaggi nella societa e nella tecnica. Milan : Olivetti.

Derwing, B. (1973). Transformational Grammar as a Theory of Language Acquisition. London : Cambridge University Press.

De Soto, C.B., London, M., and Handel, S. (1965). "Social reasoning and spatial paralogic". Journal of Personality and Social Psychology, 2, 513 - 521.

Dijk, T.A. van (1973). Aspects of Text Grammars. The Hague : Mouton.

Donaldson, M. (1974). "On judging truth values". Unpublished manuscript, Edinburgh University.

Donnellan, K. (1971). "Reference and definite descriptions". In D.D. Steinberg and L.A. Jakobovits (Eds.), Semantics. London : Cambridge University Press.

Fillenbaum, S. (1971). "Psycholinguistics". Annual Review of Psychology, 22, 251 - 308.

________ (1973). Syntactic Factors in Memory. The Hague : Mouton.

Fillmore, C.J. (1968). "The Case for Case". In E. Bach and R.T. Harms (Eds.), Universals in Linguistic Theory. London : Holt, Rinehart and Winston.

________ and Langendoen, D.T. (Eds.) (1971). Studies in Linguistic Semantics. London : Holt, Rinehart and Winston.

Fredericksen, C.H. (1975a). "Acquisition of semantic information from discourse : effects of repeated exposures". Journal of Verbal Learning and Verbal Behaviour, 14, 158 - 169.

_______________ (1975b). "Representing logical and semantic structure of knowledge acquired from discourse". Cognitive Psychology, 7, 371 - 458.

Garrod, S., and Trabasso, T. (1973). "A dual-memory information-processing interpretation of sentence comprehension". Journal of Verbal Learning and Verbal Behaviour, 12, 155 - 167.

Geach, P.T. (1973). Logic Matters. Oxford : Blackwell.

Glucksberg, S., Krauss, R.M., and Weisberg, R. (1966). "Referential communication in nursery school children : method and some preliminary findings". Journal of Experimental Child Psychology, 3, 333 - 342.

_________, Trabasso, T., and Wald, J. (1973). "Linguistic structure and mental operations". Cognitive Psychology, 5, 358 - 370.

Goodman, L.A. (1969). "On partitioning X^2 and detecting partial association in three-way contingency tables". Journal of the Royal Statistical Society Series B, 31, 486 - 498.

___________ (1970). "The multivariate analysis of qualitative data : interactions among multiple classifications". Journal of the American Statistical Association, 65, 226 - 256.

Goodman, L.A. (1971). "Partitioning of χ^2, analysis of marginal contingency tables, and estimation of expected frequencies in multidimensional contingency tables". Journal of the American Statistical Association, 66, 339 - 344.

Gough, P.B. (1965). "Grammatical transformations and speed of understanding". Journal of Verbal Learning and Verbal Behaviour, 4, 107 - 111.

__________ (1966). "The verification of sentences : the effects of delay of evidence and sentence length". Journal of Verbal Learning and Verbal Behaviour, 5, 492 - 496.

Greenberg, J.H. (1966). Language Universals. The Hague : Mouton.

Greene, J. (1970). "The semantic function of negatives and passives". British Journal of Psychology, 61, 17 - 22.

__________ (1972). Psycholinguistics. Harmondsworth - Penguin Books.

Grieve, R. (1974). "Definiteness in discourse". Language and Speech, 16, 365 - 372.

__________ and Wales, R.J. (1973). "Passives and topicalisation". British Journal of Psychology, 64, 173 - 182.

Grice, P. (1969). "Utterer's meaning and intentions". The Philosophical Review, 78, 147 - 177.

Grimes, J.E. (1975). The Thread of Discourse. The Hague : Mouton.

Goodman, L.A. (1971). "Partitioning of χ^2, analysis of marginal contingency tables, and estimation of expected frequencies in multidimensional contingency tables". Journal of the American Statistical Association, 66, 339 - 344.

Gough, P.B. (1965). "Grammatical transformations and speed of understanding". Journal of Verbal Learning and Verbal Behaviour, 4, 107 - 111.

__________ (1966). "The verification of sentences : the effects of delay of evidence and sentence length". Journal of Verbal Learning and Verbal Behaviour, 5, 492 - 496.

Greenberg, J.H. (1966). Language Universals. The Hague : Mouton.

Greene, J. (1970). "The semantic function of negatives and passives". British Journal of Psychology, 61, 17 - 22.

__________ (1972). Psycholinguistics. Harmondsworth - Penguin Books.

Grieve, R. (1974). "Definiteness in discourse". Language and Speech, 16, 365 - 372.

__________ and Wales, R.J. (1973). "Passives and topicalisation". British Journal of Psychology, 64, 173 - 182.

Grice, P. (1969). "Utterer's meaning and intentions". The Philosophical Review, 78, 147 - 177.

Grimes, J.E. (1975). The Thread of Discourse. The Hague : Mouton.

Haberman, S.J. (1972). "Log linear fit for contingency tables". Applied Statistics, 21, 218 - 225.

Hall, L.C. (1975). Linguistic and Perceptual Constraints on Scanning Strategies : some Developmental Studies. Unpublished Ph.D thesis : University of Edinburgh.

Halliday, M.A.K. (1961). "Categories of the theory of grammar". Word, 17, 241 - 292.

__________ (1963). "The tones of English". Archivum Linguisticum, 15, 1 - 28.

__________ (1966a). "Some notes on 'deep' grammar". Journal of Linguistics, 2, 55 - 67.

__________ (1966b). "Lexis as a linguistic level". In C.E.Bazell, J.C. Catford, M.A.K. Halliday and R.H.Robins (Eds.), In Memory of J.R. Firth, 148 - 162. London : Longman.

__________ (1967a - 1968). "Notes on transitivity and theme in English". Parts One - Three. Journal of Linguistics, 3, 37 - 81 and 199 - 244 (1967), 4, 179 - 215.

__________ (1967b). "Some aspects of the thematic organisation of the English clause". Unpublished manuscript.

__________ (1970). "Language structure and language function". In J. Lyons (Ed.), New Horizons in Linguistics, 140 - 165. Harmondsworth : Penguin Books.

Halliday, M.A.K. (1971). "Relevant models of language". E.L.T. Documents, 4, 2 - 9.

______________ (1972). Towards a Sociological Semantics. Universita di Urbino : Centro Internazionale di Semiotica e di Linguistica.

______________ (1973). Explorations in the Functions of Language. London : Edward Arnold.

______________ (1974). Language and Social Man. London : Longman.

______________ (1975). Learning How to Mean - Explorations in the Development of Language. London : Edward Arnold.

______________ (undated). "Options and functions in the English clause". Unpublished manuscript.

______________ and Hasan, R. (1976). Cohesion in English. London : Longman.

______________, McIntosh, A., and Strevens, P. (1964). The Linguistic Sciences and Language Teaching. London : Longman.

Hamilton, H.W., and Deese, J. (1971). "Does linguistic marking have a psychological correlate?". Journal of Verbal Learning and Verbal Behaviour, 10, 707 - 714.

Harris, R.J. (1973). "Answering questions containing marked and unmarked adjectives and adverbs". Journal of Experimental Psychology, 97, 399 - 401.

Hintikka, K.J.J. (1969). Models for Modalities. Dordrecht : Reidel.

____________ (1973). Logic, Language - Games and Information. London : Oxford University Press.

Hornby, P.A. (1971). "The role of topic-comment in the recall of cleft and pseudocleft sentences". In Papers from the Seventh Regional Meeting of the Chicago Linguistic Society, 445 - 453. Chicago : Chicago Linguistic Society.

____________ (1972). "The psychological subject and predicate". Cognitive Psychology, 3, 632 - 642.

____________ (1973). "Intonation and syntactic structure in the development of presupposition". Paper presented at the Biennial Meeting of the Society for Research in Child Development, Philadelphia.

____________ (1974). "Surface structure and presupposition". Journal of Verbal Learning and Verbal Behaviour, 13, 530 - 538.

____________, Hass, W.A., and Feldman, C.F. (1970). "A developmental analysis of the psychological subject and predicate of the sentence". Language and Speech, 13, 182 - 193.

Hudson, R. (1971). English Complex Sentences. London : North Holland Publishing Company.

Hunter, I.M.L. (1957). "The solving of three-term series problems". British Journal of Psychology, 48, 286 - 298.

Huttenlocher, J. (1968). "Constructing spatial images : a strategy in reasoning". Psychological Review, 75, 550 - 560.

_______________, Eisenberg, K., and Strauss, S. (1968) "Comprehension : relation between perceived actor and logical subject". Journal of Verbal Learning and Verbal Behaviour, 7, 527 - 530.

_______________ and Higgins, E.T. (1971). "Adjectives, comparatives, and syllogisms". Psychological Review, 78, 487 - 504.

_______________ and Strauss, S. (1968). "Comprehension and a statement's relation to the situation it describes". Journal of Verbal Learning and Verbal Behaviour, 7, 300 - 304.

_______________ and Weiner, S.A. (1971). "Comprehension of instructions in varying contexts". Cognitive Psychology, 2, 369 - 385.

Jackendoff, R.S. (1972). Semantic Interpretation in Generative Grammar. Cambridge, Mass. : M.I.T. Press.

James, C.T., Thompson, J.G., and Baldwin, J.M. (1973). "The reconstructive process in sentence memory". Journal of Verbal Learning and Verbal Behaviour, 12, 51 - 63.

Johnson-Laird, P.N. (1968a). "The interpretation of the passive voice". Quarterly Journal of Experimental Psychology, 20, 69 - 73.

_______________ (1968b). "The choice of the passive voice in a communication task". British Journal of Psychology, 59, 7 - 15.

Johnson-Laird, P.N. (1969). "On understanding logically complex sentences". Quarterly Journal of Experimental Psychology, 21, 1 - 13.

__________ (1972). "The three-term series problem". Cognition, 1, 57 - 82.

__________ (1974). "Experimental Psycholinguistics". Annual Review of Psychology, 25, 135 - 160.

__________ and Tridgell, J.M. (1972). "When negation is easier than affirmation". Quarterly Journal of Experimental Psychology, 24, 87 - 91.

Just, M.A. (1974). "Comprehending quantified sentences : the relation between sentence-picture and semantic memory verification". Cognitive Psychology, 6, 216 - 236.

__________ and Carpenter, P.A. (1975). "The semantics of locative information in pictures and mental images". British Journal of Psychology, 66, 427 - 441.

Katz, J.J. and Postal, P. (1964). An Integrated Theory of Linguistic Descriptions. Cambridge, Mass. : M.I.T. Press.

Kempson, R. (1975). Presupposition and the Delimitation of Semantics. London : Cambridge University Press.

Kennedy, A. (1973). "Associative encoding of words in passive sentences". British Journal of Psychology, 64, 169 - 172.

Kintsch, W. (1972). "Notes on the structure of semantic memory". In E. Tulving and W. Donaldson (Eds.), Organisation of Memory. London : Academic Press.

Krauss, R.M. and Glucksberg, S. (1969). The development of communication : competence as a function of age". Child Development, 40, 255 - 266.

__________, Vivekananthan, P.S., and Weinheimer, S. (1968). ""Inner speech" and "external speech" : characteristics and communication effectiveness of socially and non-socially encoded messages". Journal of Personality and Social Psychology, 9, 295 - 300.

Lakoff, G. (1972). "Linguistics and natural logic". In G. Harman and D. Davidson (Eds.), Semantics of Natural Language. Dordrecht : Reidel.

Langendoen, D.T. (1968). The London School of Linguistics. Cambridge, Mass. : M.I.T. Press.

Lyons, J. (1968). An Introduction to Theoretical Linguistics. London : Cambridge University Press.

Marks, L.E. and Miller, G.A. (1964). "The role of semantic and syntactic constraints in the memorization of English sentences". Journal of Verbal Learning and Verbal Behaviour, 3, 1 - 5.

Mathesius, V. (1975). A Functional Analysis of Present Day English on a General Linguistic Basis. Edited by J. Vachek. The Hague : Mouton.

Menyuk, P. (1968). "The role of distinctive features in children's acquisition of phonology". Journal of Speech and Hearing Research, 11, 138 - 146.

_________ (1971). The Acquisition and Development of Language. Englewood Cliffs : Prentice Hall.

Miller, G.A., and McKean, K.E. (1964). "A chronometric study of some relations between sentences". Quarterly Journal of Experimental Psychology, 16, 297 - 308.

Morton, J. (1966). "Discussion of "On Hearing Sentences"." In. J. Lyons and R.J. Wales (Eds.), Psycholinguistics Papers, 19 - 23. Edinburgh : Edinburgh University Press.

Nelson, K. (1974). "Concept, Word and Sentence : inter-relations in acquisition and development". Psychological Review, 81, 267 - 285.

Nilsen, D.L.F. (1972). Towards a Semantic Specification of Deep Case. The Hague : Mouton.

Olson, D.R., and Filby, N. (1972). "On the comprehension of active and passive sentences". Cognitive Psychology, 3, 361 - 381.

Osgood, C.E., Suci, G.J., and Tannenbaum, P.H. (1957). The Meaning of Meaning. Urbana : University of Illinois Press.

Quirk, R., Greenbaum, S., Leech, G., and Svartvik, J. (1972). A Grammar of Contemporary English. London : Longman.

Rommetveit, R. (1968). Words, Meanings and Messages. London : Academic Press.

___________ (1975). On Message Structure. A Framework for the Study of Language and Communication. London : Wiley.

Ross, J.R. (1972). "Act". In D.Davidson and G. Harman (Eds.), Semantics of Natural Language, 70 - 126. Dordrecht : Reidel.

Rubin, A. (1972). "Grammar for the people : flowcharts of SHRDLU's grammar". M.I.T. Artificial Intelligence Memo No. 282.

Russell, B. (1904). "On Denoting". Mind, new series, 14, 479 - 493.

Ryle, G. (1949). The Concept of Mind. London : Hutchinson.

Sachs, J.S. (1967). "Recognition memory for syntactic and semantic aspects of connected discourse". Perception and Psychophsics, 2, 437 - 442.

Savin, H.B., and Perchonock, E. (1965). "Grammatical structure and the immediate recall of English sentences". Journal of Verbal Learning and Verbal Behaviour, 4, 348 - 353.

Schank, R. (1972). "Conceptual dependency : a theory of natural language understanding". Cognitive Psychology, 3, 552 - 631.

Searle, J.R. (1969) Speech Acts. Cambridge : Cambridge University Press.

Seymour, P.H.K. (1969). "Response latencies in judgements of spatial location". British Journal of Psychology, 60, 31 - 39.

Seymour, P.H.K. (1973). "Judgements of verticality and response availability". <u>Bulletin of the Psychonomic Society</u>, **1**, 196 - 198.

______________ (1974a). "Asymmetries in judgements of verticality". <u>Journal of Experimental Psychology</u>, **102**, 447 - 455.

______________ (1974b). "Pictorial coding of verbal descriptions". <u>Quarterly Journal of Experimental Psychology</u>, **26**, 39 - 51.

Shaffer, J.P. (1973). "Defining and testing hypotheses in multidimensional contingency tables". <u>Psychological Bulletin</u>, **79**,127 - 141.

Shanaham, P.J. (1976). Unpublished Ph.D. thesis. University of Stirling.

Slobin, D.I. (1966). "Grammatical transformations and sentence comprehension in childhood and adulthood". <u>Journal of Verbal Learning and Verbal Behaviour</u>, **5**, 219 - 227.

___________ (1968). "Recall of full and truncated passive sentences in connected discourse". <u>Journal of Verbal Learning and Verbal Behaviour</u>, **7**, 876 - 888.

Smith, K.H., and McMahon, L.E. (1970). "Understanding order information in sentences : some recent work at Bell Laboratories". In G. Flores d'Arcais and W.J.M. Levelt (Eds.), <u>Advances in Psycholinguistics</u>. Amsterdam : North Holland.

Strawson, P.F. (1959). <u>Individuals</u>. London : Methuen.

Strawson, P.F. (1964). "Identifying reference and truth values". Reprinted in Logico-linguistic Papers. London : Methuen, 1971.

Svartvik, J. (1966). On Voice in the English Verb. The Hague : Mouton.

Tannenbaum, P.H., and Williams, F. (1968). "Prompted word replacement in active and passive sentences". Language and Speech, 11, 220 - 229.

Thorndike, E.L., and Lorge, I. (1944). The Teacher's Word Book of 30,000 Words. New York : Teachers College, Columbia University.

***

Trabasso, T., Rollins, M. and Shaughnessy, E. (1971). "Storage and verification stages in processing concepts". Cognitive Psychology, 2, 239 - 289.

Turner, E.A., and Rommetweit, R. (1967a). "The acquisition of sentence voice and reversibility". Child Development, 38, 64 - 80.

__________ and Rommetveit, R. (1967b). "Experimental manipulation of the production of active and passive voice in children". Language and Speech, 10, 169 - 180.

__________ and Rommetveit, R. (1968). "Focus of attention in recall of active and passive sentences". Journal of Verbal Learning and Verbal Behaviour, 7, 543 - 548.

Vachek, J. (1966). The Linguistic School of Prague. Bloomington, Indiana : Indiana University Press.

Wason, P.C. (1965). "The contexts of plausible denial". Journal of Verbal Learning and Verbal Behaviour, 4, 7 - 11.

Wason, P.C. (1971). "In real life negatives are false". Communication and Cognition, 4, 239 - 253.

__________ and Johnson-Laird, P.N. (1970). "A conflict between selecting and evaluating information in an inferential task". British Journal of Psychology, 61, 509 - 515.

Watt, W.C. (1970). "On two hypotheses concerning psycholinguistics". In J.R. Hayes (Ed.), Cognition and the Development of Language. New York : Wiley.

Winograd, T. (1972). Understanding Natural Language. Edinburgh : Edinburgh University Press.

Wright, P. (1969). "Transformations and the understanding of sentences". Language and Speech, 12, 156 - 166.

__________ (1972). "Some observations on how people answer questions about sentences". Journal of Verbal Learning and Verbal Behaviour, 11, 188 - 195.

ADDENDA

**Bever, T.G. (1970). "The influence of speech performance on linguistic structures". In G.B. Flores d'Arcais and W.J.M. Levelt (Eds.) Advances in Psycholinguistics. Amsterdam: North Holland.

***Trabasso, T. (1972). "Mental operations in language comprehension". In J.B. Carroll and R.O. Freedle (Eds.) Language Comprehension and the Acquisition of Knowledge. Washington: Winston.

Appendix A

Preamble sentences for Experiment 2. Those for Experiment 1 are quite similar in style, though the pictures used were not the same. In the interests of economy, I have omitted those.

1. Here is a picture of a giraffe.
   It is quite a small specimen which does not yet have horns, although its markings are already quite distinct.

2. Here is a picture of a pelican.
   His neck is raised and his wings are outstretched.
   He is looking towards the camera.

3. Here is a picture of a horse.
   It is a small black foal with white 'socks' around its ankles.
   It has a short, bushy tail.

4. Here is a picture of a camel.
   It is a rather shaggy example of the two humped variety.
   It has a blue brace and bit around its nose.

5. Here is a picture of a wild boar.
   It is a ferocious looking beast with little white tusks and long dark hair.

6. Here is a picture of a tiger.
   It is sitting down but it is leaning forward with its ears laid flat.
   Its mouth is open.

7. Here is a picture of a walrus.
   It is a dark steely grey colour.
   It is pretty big and fat with long white tusks.

Appendix A

Preamble sentences for Experiment 2. Those for Experiment 1 are quite similar in style, though the pictures used were not the same. In the interests of economy, I have omitted those.

1. Here is a picture of a giraffe.
   It is quite a small specimen which does not yet have horns, although its markings are already quite distinct.

2. Here is a picture of a pelican.
   His neck is raised and his wings are outstretched.
   He is looking towards the camera.

3. Here is a picture of a horse.
   It is a small black foal with white 'socks' around its ankles.
   It has a short, bushy tail.

4. Here is a picture of a camel.
   It is a rather shaggy example of the two humped variety.
   It has a blue brace and bit around its nose.

5. Here is a picture of a wild boar.
   It is a ferocious looking beast with little white tusks and long dark hair.

6. Here is a picture of a tiger.
   It is sitting down but it is leaning forward with its ears laid flat.
   Its mouth is open.

7. Here is a picture of a walrus.
   It is a dark steely grey colour.
   It is pretty big and fat with long white tusks.

8. Here is a picture of a duck.

It is brown with a pink breast.

It is surrounded by four yellow ducklings.

9. Here is a picture of a lion.

It is a large adult male with a huge mane.

Its mouth is open wide in a rather aggresive manner.

10. Here is a picture of a penguin.

Its neck is stretched forward and its wings are spread wide.

It looks about to dive.

11. Here is a picture of a kettle.

It is finished in polished stainless steel.

It is a modern electric model - although its flex is not attached.

12. Here is a picture of an antelope.

He is striped but only on his legs and hindquarters.

His neck is outstretched and his tongue is stuck out.

13. Here is a picture of a shepherd.

He is wearing a grey hat and large yellow apron.

He is holding a lamb in one hand and a crook in the other.

14. Here is a picture of a dog.

It is a collie.

Its head, shoulders and tail are white - the rest of it is light brown.

15. Here is a picture of a pig.

In fact it is a rather uncommon saddleback sow.

It looks as though it is eating.

16. Here is a picture of a rhinoceros.

In fact it is a small baby rhino.

Its head is lowered and one of its paws is raised as though it were about to charge.

17. Here is a picture of an antelope.

    It is an unusual species with a long shaggy coat and long slender curved horns.

18. Here is a picture of a horse.

    It looks like a Clydesdale.

    It is a large brown stallion with black mane and tail.

19. Here is a picture of a vulture.

    It has large black wings and a long bald neck.

    It is perched on a dead tree trunk.

20. Here is a picture of a monkey.

    In fact it looks like a rhesus.

    It is standing on its hind legs with its arms reaching out.

21. Here is a picture of a chimpanzee.

    It is sitting quite still with one arm resting on its knees and the other one casually reaching out.

22. Here is a picture of an aeroplane.

    It is a small jet airliner carrying B.O.A.C. markings.

    Its undercarriage is down.

23. Here is a picture of a car.

    It is a small green racing car with the number '32' on its side.

    Its boot is open.

24. Here is a picture of a goat.

    It is a fully matured nanny goat.

    She is wearing a blue collar with a little bell on it.

25. Here is a picture of a turkey.

    It is an adult male with the familiar red neck and head.

    Its feathers are black but the tail feathers have white tips.

26. Here is a picture of a farmer.

He wears a brown cap and grey jacket and is smoking a pipe.

He carries a shotgun.

27. Here is a picture of a deer.

It is a full grown stag with a fine set of antlers.

It has very heavy fur on its chest but much lighter fur over the rest of its body.

28. Here is a picture of an anteater.

It has the characteristic bushy tail.

It also has unusual white markings on its back and a white nose.

29. Here is a picture of a gorilla.

It is a large black male.

He is standing on his hind legs with his arms raised high.

30. Here is a picture of a rabbit.

It is a small albino.

Although it is crouched down its ears are pricked up.

31. Here is a picture of a tractor.

It has a farmer in the driving seat.

It is a new Ford 5000 with a safety cab.

32. Here is a picture of a cow.

It is a beautiful golden Jersey cow.

It has finely curved greyish horns and is standing perfectly still.

33. Here is a picture of a woman.

She is wearing a beige smock and purple trousers and carrying a handbag.

She has long brown hair.

34. Here is a picture of a hippopotamus.

It is a very dark baby hippo.

It is sitting down and appears to be eating.

35. Here is a picture of a bull.

It is a huge black and white Frisian with a ring through its nose.

It's starting to run.

36. Here is a picture of a sandal.

It has a thick cork platform sole.

The upper is made of blue and yellow leather.

Appendix B.

Stimulus materials for Experiment 4 : only the versions with "the" + noun are listed.

1. Here is a gorilla called Fred.

   The gorilla is standing on his hind legs.

2. Here is a tiger called Theo.

   The tiger is sitting down.

3. Here is a lion called Leo.

   The lion is lying down.

4. Here is a girl called Anna.

   The girl is dancing alone.

5. Here is a boy called Jimmy.

   The boy is singing noisily.

6. Here is a horse called Ed.

   The horse is galloping along.

7. Here is a cow called Daisy.

   The cow seems to be pregnant.

8. Here is a dog called Bonzo.

   The dog is barking loudly.

9. Here is a cat called Ming.

   The cat is purring quietly.

10. Here is a hamster called Hammy.

    The hamster is having a snooze.

11. Here is a woman called Louise.

    The woman looks very elegant.

12. Here is a man called Will.

    The man is talking incessantly.

13. Here is a cock called Dave.

    The cock is strutting about.

14. Here is a sow called Bessie.

    The sow is very dirty.

15. Here is an elephant called Majid.

    The elephant is asleep.

16. Here is a badger called Ben.

    The badger is sniffing about.

17. Here is a goat called Billy.

    The goat is jumping a brook.

18. Here is a rabbit called Snowy.

    The rabbit is crouching down.

19. Here is a bear called Bruin.

    The bear lives in Edinburgh zoo.

20. Here is a bull called Pete.

    The bull has a broken horn.

21. Here is a mouse called Jerry.

    The mouse lives in Murray Hall.

22. Here is a camel called Mohammad.

    The camel has a damaged foreleg.

23. Here is a coot called Malcolm.

The coot lives by the loch.

24. Here is a swan called Nat.

The swan is flying low.

Practice Problems

1. Fred is taller than Jane.
   Jane is taller than Jack.
   Who is tallest?

2. John is bigger than [illegible].
   Bill is bigger than John.
   Who is smallest?

3. Pete's dog is wilder than [illegible]'s.
   His is tamer than [illegible]'s.
   Whose is wildest?

4. Will is noisier than [illegible].
   Sam is quieter than [illegible].
   Who is noisiest?

5. Cliff is younger than Janet.
   She is younger than [illegible].
   Who is youngest?

6. Ernie is worse than [illegible].
   Charlie is worse than him.
   Who is best?

7. James' office is cooler than [illegible]'s.
   James' is warmer than Frank's.
   Whose is coolest?

8. Herb's tree is lower than Ellen's.
   Tom's is higher than Ellen's.
   Whose is highest?

Practice Problems

1. Fred is taller than Jane.
   Jane is taller than Jack.
   Who is tallest?

2. John is bigger than Mary.
   Bill is bigger than John.
   Who is smallest?

3. Pete's dog is wilder than Ann's.
   His is tamer than Dave's.
   Whose is wildest?

4. Will is noisier than Ruth.
   Sam is quieter than her.
   Who is noisiest?

5. Cliff is younger than Janet.
   She is younger than Robin.
   Who is youngest?

6. Ernie is worse than Sue.
   Charlie is worse than him.
   Who is best?

7. James' office is cooler than Liz's.
   James' is warmer than Frank's.
   Whose is coolest?

8. Herb's tree is lower than Ellen's.
   Tom's is higher than Ellen's.
   Whose is highest?

Problem set 1.

1. Neville's garden is longer than Helen's.
   Peter's is longer than Neville's.
   Whose is longest?

2. Roger's house is farther than Margaret's.
   Bob's is farther than Roger's.
   Whose is nearest?

3. George is happier than Fiona.
   Raymond is happier than George.
   Who is happiest?

4. Ian's car is faster than Eve's.
   Harold's is faster than Ian's.
   Whose is slowest?

5. Mike is tidier than Karen.
   Karen is tidier than Joe.
   Who is tidiest?

6. Simon's driveway is wider than Rachel's.
   Rachel's is wider than Henry's.
   Whose is narrowest?

7. Tony is fatter than Elizabeth.
   Elizabeth is fatter than Gerald.
   Who is fattest?

8. Philip is brighter than Elspeth.
   Elspeth is brighter than Andrew.
   Who is dullest?

9. Harry's hair is lighter than Julie's.

Christopher's is darker than Julie's.

Whose is lightest?

10. Matt's garage is cleaner than Kathy's.

Fergus' is dirtier than Kathy's.

Whose is dirtiest?

11. Karl's house is hotter than Morag's.

Julian's is colder than Morag's.

Whose is hottest?

12. Max's wine is smoother than Linda's.

Nigel's is rougher than Linda's.

Whose is roughest?

13. Steve's knife is sharper than Laura's.

Steve's is blunter than Mark's.

Whose is sharpest?

14. Adam's tent is wetter than Iris'.

Adam's is drier than Hugh's.

Whose is driest?

15. Jonathan's bread is fresher than Diana's.

Jonathan's is staler than Calum's.

Whose is freshest?

16. Bert is wiser than Veronica.

Bert is stupider than Ronald.

Who is stupidest?

17. Brian's coat is looser than Monica's.

Clive's is tighter than Monica's.

Whose is tightest?

18. Bruce's cider is sourer than Christine's.

Duncan's is sweeter than Christine's.

Whose is sourest?

19. Keith's cereal is soggier than Theresa's.

Arnold's is crisper than Theresa's.

Whose is crispest?

20. Richard's drink is cloudier than Sally's.

Martin's is clearer than Sally's.

Whose is cloudiest?

21. Sean is poorer than Rona.

Sean is richer than Terence.

Who is richest?

22. Stewart is cruder than Vivien.

Stewart is subtler than Brendan.

Who is crudest?

23. Timothy's pool is shallower than Janet's.

Timothy's is deeper than Cecil's.

Whose is deepest?

24. Donald's case is lighter than Sarah's.

Donald's is heavier than Patrick's.

Whose is lightest?

25. Edward is sicker than Ursula.

Roy is sicker than Edward.

Who is healthiest?

26. Wilf's bed is softer than Zoe's.

Lennie's is softer than Wilf's.

Whose is softest?

27. Rodney's hat is cheaper than Sheila's.

Ivan's is cheaper than Rodney's.

Whose is dearest?

28. Nick's beer is weaker than Wendy's.

Geoffrey's is weaker than Nick's.

Whose is weakest?

29. Barry's tunes are softer than Joy's.

Joy's are softer than Daniel's.

Whose is loudest?

30. Neil's stamps are rarer than June's.

June's are rarer than Dennis'.

Whose are rarest?

31. Reg's book is easier than Kirsty's.

Kirsty's is easier than Dominic's.

Whose is easiest?

32. Norman's bow is slacker than Lydia's.

Lydia's is slacker than Derek's.

Whose is tautest?

Problem Set 2.

1. Brian's coat is tighter than Monica's.

Clive's is tighter than Brian's.

Whose is loosest?

2. Bruce's cider is sweeter than Pamela's.

Duncan's is sweeter than Bruce's.

Whose is sweetest?

3. Keith's cereal is crisper than Theresa's.

Arnold's is crisper than Keith's.

Whose is soggiest?

4. Richard's drink is clearer than Sally's.

   Martin's is clearer than Richard's.

   Whose is clearest?

5. Sean is richer than Rona.

   Rona is richer than Terence.

   Who is poorest?

6. Stewart is subtler than Vivien.

   Vivien is subtler than Brendan.

   Who is subtlest?

7. Timothy's pool is deeper than Sally's.

   Sally's is deeper than Cecil's.

   Whose is shallowest?

8. Donald's case is heavier than Sarah's.

   Sarah's is heavier than Patrick's.

   Whose is heaviest?

9. Edward is healthier than Ursula.

   Roy is sicker than Ursula.

   Who is sickest?

10. Wilf's bed is harder than Zoe's.

    Lennie's is softer than Zoe's.

    Whose is hardest?

11. Rodney's hat is dearer than Sheila's.

    Ivan's is cheaper than Sheila's.

    Whose is cheapest?

12. Nick's beer is stronger than Wendy's.

Geoffrey's is weaker than Wendy's.

Whose is strongest?

13. Barry's tunes are louder than Joy's.

Barry's are softer than Daniel's.

Whose is softest?

14. Neil's stamps are commoner than June's.

Neil's are rarer than Denise'.

Whose are commonest?

15. Reg's book is harder than Kirsty's.

Reg's is easier than Dominic's.

Whose is easiest?

16. Norman's bow is tauter than Lydia's.

Norman's is slacker than Derek's.

Whose is tautest?

17. Neville's garden is shorter than Helen's.

Peter's is longer than Helen's.

Whose is shortest?

18. Roger's house is nearer than Margaret's.

Bob's is farther than Margaret's.

Whose is farthest?

19. George is sadder than Fiona.

Raymond is happier than Fiona.

Who is saddest?

20. Ian's car is slower than Eve's.

Harold's is faster than Eve's.

Whose is fastest?

21. Mike is sloppier than Karen.

Mike is tidier than Joe.

Who is sloppiest?

22. Simon's driveway is narrower than Iris'.

Simon's is wider than Henry's.

Whose is widest?

23. Tony is thinner than Elizabeth.

Tony is fatter than Gerald.

Who is thinnest?

24. Philip is duller than Elspeth.

Philip is brighter than Andrew.

Who is brightest?

25. Harry's hair is darker than Julie's.

Christopher's is darker than Harry's.

Whose is darkest?

26. Matt's garage is dirtier than Kathy's.

Fergus' is dirtier than Matt's.

Whose is cleanest?

27. Karl's house is colder than Morag's.

Julian's is colder than Karl's.

Whose is coldest?

28. Max's wine is rougher than Linda's.

Nigel's is rougher than Max's.

Whose is smoothest?

29. Steve's knife is blunter than Laura's.

Laura's is blunter than Mark's.

Whose is bluntest?

21. Mike is sloppier than Karen.

Mike is tidier than Joe.

Who is sloppiest?

22. Simon's driveway is narrower than Iris'.

Simon's is wider than Henry's.

Whose is widest?

23. Tony is thinner than Elizabeth.

Tony is fatter than Gerald.

Who is thinnest?

24. Philip is duller than Elspeth.

Philip is brighter than Andrew.

Who is brightest?

25. Harry's hair is darker than Julie's.

Christopher's is darker than Harry's.

Whose is darkest?

26. Matt's garage is dirtier than Kathy's.

Fergus' is dirtier than Matt's.

Whose is cleanest?

27. Karl's house is colder than Morag's.

Julian's is colder than Karl's.

Whose is coldest?

28. Max's wine is rougher than Linda's.

Nigel's is rougher than Max's.

Whose is smoothest?

29. Steve's knife is blunter than Laura's.

Laura's is blunter than Mark's.

Whose is bluntest?

30. Adam's tent is drier than Iris'.

Iris' is drier than Hugh's.

Whose is wettest?

31. Jonathan's bread is staler than Diana's.

Diana's is staler than Calum's.

Whose is stalest?

32. Bert is stupider than Veronica.

Veronica is stupider than Ronald.

Who is wisest?

[illegible]

[illegible] sentences for [illegible]

[illegible]

1. In the middle of this picture is a [illegible].
   It's a little [illegible] with [illegible] and [illegible]
   It's [illegible]
   The [illegible] is [illegible]

2. In the middle of this picture is a [illegible].
   It's only a little one, though, and [illegible]
   It's standing very still.
   The [illegible] is [illegible]

3. In the middle of this picture is a [illegible].
   [illegible]
   [illegible]
   The [illegible] is [illegible]

4. In the middle of this picture is a [illegible].
   He's a very old [illegible] with a long [illegible]
   He's [illegible]
   The [illegible] is [illegible]

5. In the middle of this picture is a [illegible].
   It's [illegible] with a black tail and white [illegible]
   It's [illegible]
   A [illegible] is [illegible] the [illegible]

6. In the middle of this picture is a [illegible].
   It's [illegible] a big [illegible] and a brown [illegible]
   He has a little [illegible]
   The [illegible] is being [illegible] by an [illegible]

<u>Appendix D</u>

Preamble Sentences for Experiments 6,7,8 and 9, together with a sample target sentence for each set.

1. In the middle of this picture is a sheep.

   It's a little grey sheep with big yellow horns and brown legs.

   It's walking slowly along.

   The sheep is following a tortoise.

2. In the middle of this picture is a giraffe.

   It's only a little one, though, and it hasn't any horns yet.

   It's standing very still.

   The giraffe is watching a hen.

3. In the middle of this picture is a soldier.

   He's wearing a tin hat with twigs and leaves stuck on it.

   He's also wearing a green jacket.

   The soldier is shooting a cowboy.

4. In the middle of this picture is a tiger.

   He's a very big tiger with a long tail.

   His mouth is open wide.

   The tiger is chasing a lion.

5. In the middle of this picture is a horse.

   It's brown with a black tail and white nose.

   It's quite fat.

   A bull is chasing the horse.

6. In the middle of this picture is a cowboy.

   He's wearing a big blue hat and a brown coat.

   He has a little moustache.

   The cowboy is being shot by an Indian.

Appendix D

Preamble Sentences for Experiments 6,7,8 and 9, together with a sample target sentence for each set.

1. In the middle of this picture is a sheep.

   It's a little grey sheep with big yellow horns and brown legs.

   It's walking slowly along.

   The sheep is following a tortoise.

2. In the middle of this picture is a giraffe.

   It's only a little one, though, and it hasn't any horns yet.

   It's standing very still.

   The giraffe is watching a hen.

3. In the middle of this picture is a soldier.

   He's wearing a tin hat with twigs and leaves stuck on it.

   He's also wearing a green jacket.

   The soldier is shooting a cowboy.

4. In the middle of this picture is a tiger.

   He's a very big tiger with a long tail.

   His mouth is open wide.

   The tiger is chasing a lion.

5. In the middle of this picture is a horse.

   It's brown with a black tail and white nose.

   It's quite fat.

   A bull is chasing the horse.

6. In the middle of this picture is a cowboy.

   He's wearing a big blue hat and a brown coat.

   He has a little moustache.

   The cowboy is being shot by an Indian.

7. In the middle of this picture is a dog.

It's a very pretty brown and white collie dog.

It has a big bushy tail.

A calf is chasing the dog.

8. In the middle of this picture is a duck.

It is brown with a pink chest.

It has a very long neck.

A pig is being watched by the duck.

9. In the middle of this picture is a soldier.

He's wearing a green uniform with a tin hat and big black boots.

He's kneeling down.

The soldier is being shot by a knight.

10. In the middle of this picture is a monkey.

It's a little grey one with a very long tail.

He's walking slowly along.

The monkey is following a bird.

11. In the middle of this picture is a man.

He's a farmer and he's carrying a big shotgun.

He's smoking a pipe.

A horse is watching the man.

12. In the middle of this picture is a bear.

It's a very big white polar bear.

It's mouth is open and it's growling.

The bear is following a deer.

13. In the middle of this picture is a knight.

He wears an orange jacket and a silver helmet.

He has blue trousers on.

The knight is shooting a soldier.

14. In the middle of this picture is a calf.

    It's a black and white calf.

    It's tail is sticking up in the air.

    The calf is being chased by a dog.

15. In the middle of this picture is a pig.

    It's a very special kind of pig.

    It's black with a pink stripe in its middle.

    A duck is watching the pig.

16. In the middle of this picture is a deer.

    It's a big brown deer, with a tiny tail.

    It has enormous horns.

    A bear is following the deer.

17. In the middle of this picture is an Indian.

    He has two feathers in his hair.

    He has an axe strapped to his belt.

    The Indian is shooting a cowboy.

18. In the middle of this picture is a bull.

    It's a great big black and white bull.

    He's got a ring through his nose.

    The bull is being chased by a horse.

19. In the middle of this picture is a bird.

    It's a big white pelican with a long beak.

    Its wings are stretched out.

    A monkey is being followed by the bird.

20. In the middle of this picture is a horse.

    It's a white carthorse with a grey tail.

    He's standing quite still.

    A man is being watched by the horse.

Sets 1,2,3 and 4 are the four practice sets, others are experimental.

Attention is drawn to the fact that the copyright of this thesis rests with its author.

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without the author's prior written consent.