

DOCUMENT RESUME

ED 130 310

CS 203 023

AUTHOR Little, Graham  
 TITLE Form and Function in the Written Language of Sixteen Year Olds.  
 PUB DATE 75  
 NOTE 147p.; M.Ed. Thesis, University of New South Wales

EDRS PRICE MF-\$0.83 HC-\$7.35 Plus Postage.  
 DESCRIPTORS \*Composition Skills (Literary); Educational Research; Expository Writing; \*Language Development; \*Language Styles; Masters Theses; Narration; Secondary Education; \*Syntax; \*Written Language  
 IDENTIFIERS Australia

ABSTRACT

Recording, narrative, exposition, and argument were hypothesized to present writing tasks of increasing cognitive and verbal complexity. This was investigated by obtaining writing samples in each mode from a stratified sample of 128 Australian sixteen year olds. The cognitive-complexity hypothesis was supported by data concerning the relative discriminative power of the four tasks, as reflected in the correlation between judges, the correlation between task scores and examination results, and the difference in performance of more and less advanced pupils as assessed by their schools. The verbal-complexity hypothesis was supported by significant differences among the tasks on words per phrase, phrases per clause, and clauses per sentence. The more advanced pupils showed greater flexibility in patterns of language use than did the less advanced pupils. (Author/AA)

\*\*\*\*\*  
 \* Documents acquired by ERIC include many informal unpublished \*  
 \* materials not available from other sources. ERIC makes every effort \*  
 \* to obtain the best copy available. Nevertheless, items of marginal \*  
 \* reproducibility are often encountered and this affects the quality \*  
 \* of the microfiche and hardcopy reproductions ERIC makes available \*  
 \* via the ERIC Document Reproduction Service (EDRS). EDRS is not \*  
 \* responsible for the quality of the original document. Reproductions \*  
 \* supplied by EDRS are the best that can be made from the original. \*  
 \*\*\*\*\*

ED130310

U.S. DEPARTMENT OF HEALTH,  
EDUCATION & WELFARE  
NATIONAL INSTITUTE OF  
EDUCATION

THIS DOCUMENT HAS BEEN REPRO-  
DUCED EXACTLY AS RECEIVED FROM  
THE PERSON OR ORGANIZATION DRIGIN-  
ATING IT. POINTS OF VIEW OR OPINIONS  
STATED DO NOT NECESSARILY REPRESENT  
OFFICIAL NATIONAL INSTITUTE OF  
EDUCATION POSITION OR POLICY

FORM AND FUNCTION  
IN THE WRITTEN LANGUAGE OF SIXTEEN YEAR OLDS

by

Graham Little B.A. (Hons.) Dip. Ed. (Sydney)

PERMISSION TO REPRODUCE THIS COPY-  
RIGHTED MATERIAL HAS BEEN GRANTED BY  
Graham Little

TO ERIC AND ORGANIZATIONS OPERATING  
UNDER AGREEMENTS WITH THE NATIONAL IN-  
STITUTE OF EDUCATION. FURTHER REPHO-  
DUCTION OUTSIDE THE ERIC SYSTEM RE-  
QUIRES PERMISSION OF THE COPYRIGHT  
OWNER.

A thesis submitted in partial fulfilment of  
requirements for the award of the degree of  
Master of Education (Hons.)

School of Education  
University of New South Wales

1975

CS 203 023

This thesis, entitled, "Form and Function  
in the Written Language of Sixteen Year Olds",  
has not been submitted for a higher degree to  
any other institution

Signed : .....

A handwritten signature in cursive script, appearing to read "John", is written over a dotted line.

To Rosemary

ACKNOWLEDGEMENTS

Gratitude for assistance in this project is expressed to the following.

The New South Wales Education Department, for access to test data and schools,

The participating schools, the Principals, Staff and Students, for ready co-operation,

The Staff Research Fund, Canberra College of Advanced Education, for helpful funding,

Allen Whitehurst, Ed. Gaskell, Anne-Marie Tarry and Katie Sutherland, for their work on evaluation and word-counts,

Dr. R. M. Boyd, for critical assistance and encouragement,

Jean Pearson and Jackie Cavill for the typing of the manuscript.

## ABSTRACT

A stratified sample of 128 sixteen year olds complete four functional types of writing tasks hypothesised to be, in ascending order of cognitive and verbal complexity, 1. Recording, 2. Narrative, 3. Exposition and 4. Argument. Scripts are evaluated by two independent judges and the scores are validated against examination and school assessment data. The cognitive-complexity hypothesis is upheld in terms of the comparative discriminative power of the tasks in terms of three criteria in combination: correlation between judges, correlation between task score and examination score, and ability of task to discriminate between more and less advanced subjects as assessed by their schools. The verbal-complexity hypothesis relating form to function is upheld by significantly different patterns of words per phrase, phrases per clause and clauses per sentence among the tasks. The number of words per phrase increases acceleratingly as the level of the task rises. The other measures rise and fall, exposition producing most phrases and argument most clauses. Subjects tend to limit the number of phrases they produce per sentence. The less advanced subjects prove to be 70% to 80% as adaptable as the more advanced, significant differences in structure and performance on the harder tasks being attributable to more generation of phrases by the more advanced. Results are compared with other findings, models of adaptation of form to function are developed, and educational implications are discussed.

## CONTENTS

	<u>Page</u>
PREFACE	1
I. PROBLEMS OF FUNCTION AND FORM	2
1. Functions of Language	2
2. Person, Subject Matter and Audience	2
3. The Moffett and Britton Models of Modes of Discourse	7
4. Review of Problems	20
II. THE COLLECTION OF DATA ON FOUR WRITING TASKS	24
1. The Tasks	24
2. A Writing Test	25
3. Working Hypotheses	29
4. The Subjects	29
5. Evaluation of the Scripts	34
6. Review of Collection of Data	38
III. DETERMINATION OF THE COGNITIVE LEVELS OF THE TASKS	40
1. General Rationale	40
2. Discrimination Indices	44
3. The Three Indices Considered Together	50
4. Results for Achievement Scores	52
5. Results for the Three Discrimination Indices	55
6. Interpretation of Results	58
7. Conclusions	61

CONTENTS (continued)

	<u>Page</u>
IV. A. METHOD OF DETERMINING THE COMPLEXITY OF VERBAL RESPONSE	66
1. Function and Form	66
2. Analysing the Forms of Language Used on the Tasks	68
3. Word-counts and their Limitations	69
4. Words	71
5. Sentences	71
6. Clauses	72
7. Phrases	73
8. "and"	76
9. Meaning of Word-counts	76
10. Method of Making Word-counts	78
11. Reliability of Word-counts	79
12. Use of the Method	80
 V. ADAPTATION OF FORM TO FUNCTION	 82
1. Working Hypotheses	82
2. Findings	83
3. Differences between Tasks for all Subjects	92
4. Differences between Subgroups	92
5. Differences between Tasks and Groups combined	93
a) Recording	95
b) Narrative	96
c) Exposition	98
d) Argument	100
e) A Possible Limit to Increase in Word Ratios	102
6. Refined Model of Adaptation of Form to Function	103
7. Illustrative Scripts	106

## CONTENTS (continued)

	<u>Page</u>
VI. APPLICATIONS AND IMPLICATIONS	110
1. Other Tasks	110
2. Other Subjects	114
3. Other Modes of Discourse	117
4. Implications	118
a) Cognitive Aspects	119
b) Structural Aspects	122
5. Developmental Model of Functions and Forms of Discourse	125
 BIBLIOGRAPHY	 126
 APPENDICES	
A. The English Reference Test	129
B. Statistical Tables for Word-counts	133

<u>TABLES</u>	<u>Page</u>
Subjects of the study	33
2. Relationships among achievement scores for 128 subjects	53
3. Three discrimination indices for four writing tasks	55
4. Statistically significant differences between discrimination indices for four writing tasks completed by 128 subjects	57
5. Positive, zero and negative results for working hypotheses about mean word-counts	85
6. Significantly different mean word-counts for four writing tasks for 128 subjects dividing into 64 less advanced and 64 more advanced subjects on the bases of their schools' assessments	88
7. Ratio-formulae for mean word-counts from Table 6	89
8. Mean sentence formulae for small samples of published writing	116
9. Mean word-counts for 128 subjects on four writing tasks	133
10. Mean word-counts for 64 advanced and 64 not advanced subjects on four writing tasks	134

FIGURES

	<u>Page</u>
1. A version of the Moffett model of modes of discourse	9
2. The Britton taxonomy of modes of discourse	15
3. Model of age-placement of tasks by use of discrimination indices	41
4. Diagrammatic representation of significant and non-significant differences among three discrimination indices for four writing tasks completed by 128 subjects	58
5. Preliminary findings on age levels as indicated by discrimination indices for various writing tasks	63
6. Theoretical and empirical models of adaptation of form to function by 64 more advanced and 74 less advanced sixteen year olds	84
7. Mean word-count ratios for 128 subjects on four writing tasks	91
8. Model of adaptation of form to function by 128 sixteen year olds, in terms of percentage increase or decrease of word-ratios in movement from task to task	94
9. Model of adaptation of form to mode of discourse, based on the writing of 128 sixteen year olds	105
10. Developmental model of adaptation of form to function	125

## Preface

The present study is an interdisciplinary study of aspects of communicative competence. It is based on the insights of Moffett (1968) and Britton (1972) into ways in which children develop ability to handle various forms of discourse at various stages of their lives. These in turn are based on insights concerned with literary criticism and rhetoric (Lief and Light, 1972), though they are also related to work in developmental psychology (e.g. Piaget, 1961). The connections between the literary and rhetorical strand and the developmental strand of studies have been established, but exploration of the relationships must be admitted to be in its early stages, with much remaining to be done. Neither of these lines of inquiry has yet established strong links with closer syntactic analysis as practised in developmental psycholinguistics (e.g. Ferguson and Slobin, 1973) which has so far tended to concentrate on early language acquisition, though some relevant work has been done on later language development (Hunt 1965, Loban 1969) and readability (Gilliand 1972). In the present study an attempt is made to link all three strands of investigation to inquire by both analytical and quantitative methods whether there are consistent patterns of verbal change in changing communication-situations for various groups of subjects. The study draws upon selected studies in these fields for data relevant to its purposes. It is as much concerned with methodology as with findings, as it seeks to establish whether its methods appear to be worth using in further studies.

## I. PROBLEMS OF FUNCTION AND FORM

### 1. Functions of Language

The underlying hypothesis of the present study is that people use language in different ways in different situations and with different degrees of competency at various stages of their development. The study endeavours to develop techniques of investigation in this area, and applies them in this first instance to limited ranges of language and subjects.

As background, the long history of rhetoric (Lief and Light 1972) and stylistics (Turner 1973), together with psycholinguistic evidence (Ferguson and Slobin 1973) and sociolinguistic evidence (Giglioli 1972) lends little support to any static view of language or its use. The evidence is that there are stable functional varieties of language, codes or styles ranging from the ritual to the highly informal, among which competent language users range at will according to the situation in which they are using language. Competence in such uses of language develops systematically through recognisable stages related to psychological and sociological conditions.

### 2. Person, Subject Matter and Audience

It is hypothesised that the verbal behaviour of even moderately competent language users varies systematically with variations in factors such as -

person (age, experience),  
subject matter,  
audience.

If such an hypothesis is valid, it ought to be possible to trace quite systematic variations of language under changing conditions. It would follow that it would be misleading to generalise about the language of a person or group of persons without sampling the language in a defined variety of such situations. A number of empirical studies lend weight to such an hypothesis.

Labov (1969) established that there were systematic interactions between personal and situational variables in the language behaviour of certain groups of Negro children. Such children, although regarded by their teachers as virtually "non-verbal" on the basis of school performance, proved to be highly fluent and logical in certain non-school situations. Labov also established (1970) that the pronunciation of English in certain groups of adult subjects changed systematically according to whether the situation was formal, e.g. in reading aloud, or informal, e.g. in casual conversation.

Such variations are subject to social conventions, as recognised by theories of modes of discourse upon which the present study is based. As Lief and Light (1972, Appendix) point out, such theories have their origins in

ancient times. The classical rhetorical modes of discourse are description, narration, exposition and argumentation, each of which involves particular cognitive and verbal strategies which have been recognised for centuries and still apply to modern discourse. For instance, straightforward narrative is relatively concrete, and likely to be ordered in chronological sequence, proceeding mainly by naming agents and actions. Argumentation is relatively abstract, and likely to be ordered in logical rather than chronological sequence, proceeding largely by defining subjects and indicating existential and causal relationships among them. In terms of the modes of discourse in which they excel, James Joyce is a successor to Homer in the narrative mode and Karl Popper is a successor to Aristotle in the mode of argumentation. Both the functions and forms of these two forms of discourse may be clearly differentiated by rhetorical analysis and seen to be in historical continuity.

Distinct modes of discourse entailing distinct cognitive strategies and verbal styles are not confined to those recognised in literary criticism. In everyday living, language-use is subject to -

"... the conventions and presuppositions made by 'the mutual acknowledgement of communicating subjects' in the particular form of linguistic behaviour (telling a story, philosophising, buying and selling, praying, writing a novel, etc.) ... "  
(Lyons 1963 pp. 83-84.)

Despite the newborn child's tremendous capacity for acquiring language, such mastery involves a gradual acquisition of and discrimination between the various modes of discourse used in the culture. The conventions of the use of these different modes in different situations represent an important part of the competency to be acquired. The person may, of course, fail to observe the conventions, but such failure will cause problems in communication. The conventions are, of course, by no means eternally fixed; otherwise the language would be frozen and not subject to historical change. The conventions do not change rapidly, however, as the historical continuities referred to testify.

Conventions of language are certainly complex. At one level they involve various verbal strategies for various kinds of subject-matter and audience. These might be exemplified by the differences between addressing a public meeting and telling a bed-time story to a small child. At another level (Chomsky, 1969) the conventions involve various tactics in semantics, grammar and phonemics (or graphemics). The public meeting and the bed-time story call for different realms of vocabulary, different grammatical structures, and even different inflexions and pronunciations. Similar differences are to be perceived when the equivalent modes of discourse are rendered in the written form of the language. Such differences are subject to closer analysis in subsequent sections of the present study.

Mastery of the various modes of discourse would be expected to occur gradually, and at different rates for different individuals and groups. The Piagetian hypothesis (Piaget 1961) that hypothetico-deductive thinking ("formal operations"), as assessed by verbal means, emerges at ages 11 to 15 is an example of such an expectation. The studies of Connell et. al. (1975) appear, however, to necessitate a refinement of the Piagetian hypothesis in terms of differences associated with subject-matter. Although formal thought about physical problems may well appear in Western societies among 50% of those aged 11 to 15 years, formal thought about verbally-presented humanities-type problems does not appear until later. Subjects still at school are 17 or 18 by the time that 50% of them can think formally in the humanities area. Much smaller percentages of 11 to 15 year-olds can handle formal thought on humanities problems, as can similarly small percentages of 17 to 18 year-olds who have left school. (Connell et. al. 1975 Chapter 5.)

Evidently interactions between subjects at various stages of development and discourse on different types of subject-matter are more complex than the Piaget type of model would suggest. Nevertheless such interactions appear to be amenable to techniques of investigation which make systematic distinctions and as a consequence yield systematic patterns of differential ranges of change.

Such investigations necessitate closer analysis of differences among various forms of discourse in relation to different stages of human development. In this context, the term "development" is used without prejudice to the issue of whether observed norms and deviations are the product of biological development or common patterns of learning, or a combination of the two.

3. The Moffett and Britton Models of Modes of Discourse

James Moffett (1968 (a):34,35,47 especially) took the imaginative leap, supported by the observations of teachers, of applying the categories of traditional rhetoric to child development and learning. His model of language development relates an "abstractive scale" to a "rhetorical scale" in distinguishing various modes of discourse one from the other. The abstractive scale deals with the level of abstraction of the discourse, ranging from immediate reporting of concrete events in the here and now to highly generalised philosophical theorising. The rhetorical scale deals with relationship to audience, ranging from the most intimate to the most remote and impersonal. The model envisages the young child as operating in concrete and intimate forms of discourse, and the highly mature person as operating not only at this level, but also on the most abstract and impersonal level. As the levels increase, the person's relationship to his own sense-experience and his audience becomes more distant, and thus changing

relationships between person, subject matter and audience  
are depicted.

Figure 1 presents a version of the Moffett Model of modes  
of Discourse based on a way of depicting the model developed  
by Cambourne (Interim A.C.T. Education Authority 1974 : 18)



The Moffett model concentrates more on the functions of language than its forms. Through various speaker-subject-audience relationships the model explores what each type of discourse is doing. It is implied generally and made explicit at various points that changes of function are associated with changes in the form of language. Most notably, the verb-forms change from is happening to happened, and then to happens and finally to will happen, together with other modals such as may. Such changes are to be expected, but this aspect of the theory has not been very fully developed.

In the following discussion of the model, selective use is made of Moffett's subtle discussion of the modes of discourse. The treatment is selective in terms of those points about each mode which appear to be most relevant to the particular concerns of the present study. In the discussion, the present author makes an attempt to develop some tentative hypotheses about cognitive differences between certain modes of discourse, and to explore possible verbal differences which may appear to be demanded if the different functions of the different modes are to be competently achieved.

Square 1 of the model represents the recording of the drama of what is happening, more or less for one's own sake. An eye-witness makes jottings about the scene before him at the guillotine during the French revolution. A small child prattles endlessly about what is happening in his play activities. A Leopold Bloom's stream-of-consciousness registers what is happening inside and outside his head during one day in Dublin.

Cognitively, square 1 represents the simplest process in the model. There is perceptual selectivity, but the reporting flows from events as they occur rather than from any more deliberately constructive process. Verbally, the process is the least demanding of all in the model. The audience is mainly or altogether the self, or a person on intimate terms with the communicator, and no great communication gap has to be overcome. The discourse can be effectively carried out by the simple naming of objects and events, and, significantly, the verbs are confined to the present tense, with no necessity to indicate complex time-relations or the abstract, timeless relationships of defining which are required in other modes of discourse. It suffices to record events, and there is no need to account for them in terms such as cause and effect.

Developmentally, this highly concrete and egocentric mode of discourse is seen as "basic" both in the sense of being an early development and a pre-requisite of the development of competency in other modes of discourse. It is not seen as a mode of discourse which is abandoned when subsequent forms emerge, but as a mode which remains in use throughout life. The model as a whole is thus one of a widening of range rather than of the replacement of one mode of discourse by another.

Square 2 of the model represents reporting the narrative of what happened, to a fairly familiar audience. The eye-witness of the events at the guillotine tells acquaintances about what happened to Marie Antionette. The child tells his family about his

encounter with a real or imagined big dog. A person writes to a friend to tell the latest news and gossip.

Cognitively, there is more of a challenge than in simple recording, because events are recalled and shaped according to relevance to a main line of action, rather than going on before one's eyes and being subject to mere perceptual selectivity. A little more than the mere naming of objects and events is needed to lend verisimilitude to the narrative for an audience which has not shared the experience directly. More complex time-relationships are more likely to require more complex verbs that make clear that A happened after B had happened and while C was happening, and so on. Spatial as well as temporal relationships will need to be made more explicit, as by the use of appropriate phrases and clauses, and to some extent, relationships of cause and effect may similarly need to be rendered in order to make the pattern of events clear.

Square 3 represents generalising, making an exposition of what happens, to a less intimate audience. The student of the French Revolution sums up a series of events by devising the label "Reign of Terror". The child generalises in the classroom some of his observations about dogs. The teacher sums up the child's progress in a report to the parent.

Cognitively, the process is more complex than narrative,

because it represents the summarising of trends and possible exceptions in a set of relevant narrative-type sequences.

The verbs are verbs of generalisation - X usually happens - which is a more abstract mode than that of verbs of narration.

There is a call for more explicit treatment of relationships, such as are involved in the defining of terms, general rules and possible exceptions together with some suppression of irrelevancies about the particular features of events which might be vital in narrative but are distractions in generalising.

The movement is towards more of a logical and less of a chronological sequence.

Square 4 represents argumentation about what will happen or may happen or ought to happen if ... The historian or sociologist devises a theory of revolutions. This subsumes knowledge of the patterns of events in the French and other revolutions, and may go on to predict, in generalisations applied to current societies, which situations are likely to prove revolutionary and which are not. The high-school student applies Newton's laws of motion to problems set by the Science textbook and speculates about what would happen in outer space where there is no friction. Cognitively, this presents the greatest challenge of all four forms of discourse because it subsumes processes of observing, reporting and generalising into a process in which many events are involved in concepts and principles which are systematically related to one another, and which give rise to the generation of hypotheses, the making of deductions and their verification by further deductions and observations. Verbally

there is a call for precise definition and the setting out of logical and causal relationships in if ... then types of sequences. Such argumentation is not a matter of rendering personal experiences or assuming intimate social relationships, but turns its attention from the particular to the general, and is addressed to an impersonal audience such as an examiner, the general reader, or even posterity.

Broadly speaking, the model sees early childhood as the period of development of competence in square 1, the years of primary education as square 2, secondary as square 3 and tertiary education as square 4. There is no ultimate in such a developmental sequence, although one might nominate users of language of the status of a Shakespeare or an Einstein as developed to the level of the very top right-hand corner of the model.

In the Moffett model, squares 1, 2, 3 and 4 do not exhaust the possibilities of discourse, but merely outline the main trend of progression. For instance, it is possible to theorise intuitively and informally to oneself before shaping one's theories in the modes generally deemed appropriate for their public presentation. It is possible to render highly personal experiences to a wide public as in an autobiography or confessional type of poetry. As the model indicates, poetic forms of language may occupy all squares in the model, whether in imagist form of rendering a moment of sense-experience or the abstractive form of philosophical reflection. In short, all squares in the model may be occupied by some form of discourse involving a particular

combination between abstractive and rhetorical properties.

A person may change mode of discourse freely during any particular utterance, moving from abstraction to illustrate a point by anecdote and back to abstraction, and so on.

The present study concentrates upon squares 1, 2, 3 and 4 of the Moffett model as rendered above, with reference to both cognitive and verbal aspects of discourse. Relevant refinements have been made to the Moffett model by Britton (1971 : 251) in conjunction with colleagues working on the Schools Council Writing Research Project. In devising a taxonomy for the classification of samples of writing from a large number of students aged from 12 to 17, Britton and his colleagues evolved the following model, with acknowledgment to the Moffett model.

Figure 2 : The Britton Taxonomy of Modes of Discourse

1. TRANSACTIONAL	2. EXPRESSIVE	3. POETIC
1.1 <u>Informative</u>	1.2 <u>Conative</u>	3(1.1) Poetic (Inf.)
1.1.1 Record	1.2.1 Regulative	3(1.2) Poetic (Con.)
1.1.2 Report	1.2.2 Persuasive	3(2) Poetic (Exp.)
1.1.3 Generalised narrative/descriptive		
1.1.4 Analogic, low level of generalisation		
1.1.5 Analogic		
1.1.6 Speculative		
1.1.7 Tautologic		

This taxonomy distinguishes the Transactional modes of discourse from the Poetic on the grounds that in the transactional mode the communicator is playing a "participant" role. In such a role, the person is using language to get things done in real world, the world of action. The focus is upon the phenomena being dealt with in the world, and the language is a vehicle for such dealings. In the poetic mode, the person is playing a "spectator role", reflecting upon experience as distinct from getting things done in the world. The focus is more upon the thinking and feeling of the person reflecting upon experience and perhaps re-shaping his conceptions of it. In the poetic mode, the language is less a mere vehicle, and is more attended to in its own right, as verbal art. The difference in function is illustrated by different responses to the language used in the different modes of discourse. Readers are likely to be pleased if a favourite textbook is updated in content and simplified in language. Readers are not likely to be pleased if the work of a favourite poet were to be treated in such a way.

In the conative subset of the Transactional we have the language that is used to regulate behaviour, and to persuade. As in the Moffett model, the poetic may smack of any of these. Apart from refining these categories, Britton's contribution is to draw attention to the Expressive category, in which a person moves freely from the poetic to the transactional now giving information and now thinking and feeling about it, as

in a personal conversation. It is persuasively maintained by Barnes (1969), using transcripts of classroom language, that this personal mode of language-use is far more important to learning than has often been supposed.

The Moffett and Britton models are not unrelated to Halliday's analysis (1969) of seven "models" of language-use in terms of which children operate :

- I Instrumental
- II Regulatory
- III Interactional
- IV Personal
- V Heuristic
- VI Imaginative
- VII Representational

For present purposes, it is sufficient to establish that these three analyses are not inconsistent with one another, and enable the identification of modes of discourse in which there are different relationships between person, subject-matter and audience.

The Moffett model was based on general theory and the observation of children, largely by teachers. The Britton model was also based on wide-ranging theory, and was stimulated by the need to classify examples of discourse gathered in the field.

The findings of the Schools Council Writing Research Project as

reported by Martin (1975) confirm the Moffett-Britton type of model of language development in terms of the types of writing produced by various age groups. Among subjects aged 11 to 17 years there is distinct change in the proportions of discourse of various kinds as the age level rises. The writing of the youngest pupils sampled includes the greatest proportion of the more basic forms of discourse. Among slightly older pupils, low-level generalisation begins to predominate. Among older pupils still, higher-level theorising begins to make an appearance. In general, low-level transactional language appears more than might be expected: a huge proportion of school writing is the retailing of facts and low-level generalisations to the audience of teacher in the role of examiner. As Rosen points out (1973), few of the potential alternative audiences are addressed, suggesting that the school is in this sense narrowing rather than encouraging the expansion of the range of discourse in which students might become competent. Taking these qualifications into account, the Martin data appear to confirm a Moffett-Britton type of model of development in competence in various modes of discourse, in the sense that there is an association of particular modes of discourse with particular age levels. As noted, the general trend of the findings of Connell et. al. (1975) on the onset of formal thought on verbally-presented humanities-type problems as distinct from physical problems, presents a similar kind of association.

While the identification of distinct modes of discourse, in association with various age-groups, goes some way towards the verification of the Moffett and Britton types of model, it must

be admitted that techniques of verification are so far established by the relatively "subjective" techniques of literary criticism: identifying the communicator's role in and through the rhetorical modes employed. While literary-critical methods are perfectly valid in their field, the association of the modes of discourse with developmental levels calls for further data if the hypotheses are to be more widely verified. In particular, some measure is desirable of the different levels of cognitive demand hypothesised to be associated with various modes of discourse. The present study is partly devoted to exploring methods of identifying any such differences by means supplementary to an approach through rhetorical analysis.

The study is also devoted to related explorations of verbal features of the writing produced by subjects of given ages in composing in different modes of discourse. Both the Moffett and the Britton models imply identifiable verbal differences between modes, and sometimes make these explicit (e.g. in the Moffett account of the different verb-forms appropriate to recording, narrative and exposition). The analyses of verbal detail in different models of discourse is in an early stage of development, however. It therefore seems desirable to explore whether or not there are systematic relationships between developmental levels of subjects; mode of discourse; cognitive challenge of mode of discourse; and verbal forms used in the various modes of discourse. Data exist on verbal aspects of development such as changes in mean sentence-length and the incidence of various kinds of syntax at various age levels (Hunt 1965, Loban 1969),

but these data do not differentiate between various modes of discourse.

Further verification of this type of theory therefore appears to require techniques by which the cognitive demands of various forms of discourse can be identified, along with any specifically verbal demands that they make. These may then be related to the performances of persons of different levels of cognitive and verbal competency.

#### 4. Review of Problems

The present state of this line of inquiry into language competency and development appears to be as follows.

1. There is sufficient consensus on the identification of various modes of discourse for such distinctions to be operationalised in field-surveys which reveal that different age groups write different proportions of various discourse along the predicted lines.
2. Areas still imprecise include the closely-related problems of identifying -
  - a) levels of difficulty of various modes of discourse;
  - b) levels of competency in these modes;
  - c) age levels at which certain levels of competency typically occur in the various modes of discourse;
  - d) relationships between the form of language used (e.g. syntax) to its function (mode of discourse).

Approaches taken to these problems in the present study are as follows.

1. It is assumed that the question of competency must involve the critical evaluation of writing. "Communication" is an achievement-word implying that a recognisable objective has been at least to some extent attained. Even "Narrative" or "Exposition" is an achievement word, implying that if the term is applied, a story has been told or a generalisation presented, rather than that a meaningless mass of words or some other achievement discourse has been presented. While it is valuable to know that few 12 year olds begin to theorise in writing, and rather more 17 year olds (from among those 17 year olds still at school) begin to do so, it is also important to know how competent this theorising may be. Similarly, while it is valuable to know that the younger writers produce proportionately more narrative, it is important to know whether this is competent narrative compared to that produced by older writers. For these reasons, an attempt is made in the present study to evaluate the writing produced, not merely to categorise it.

2. The question of identifying age level norms is related to the question of evaluating competency. Age level norms for cognitive operations are generally found (e.g. Connell et.al. 1975) by asking questions which have right and wrong answers and identifying the age at which 50% of subjects give the right answer.

This method is not appropriate for norms in using language to communicate more than simple answers to questions, and ways need to be devised to ascertain levels for "better and worse" as well as "right and wrong" responses. A rationale is therefore developed in the present study and some measures are taken in order to explore whether such ways can be devised.

3. The question of the relationship of form to function needs further to be explored in any verification of the theory of development in modes of discourse. The Moffett-Britton models imply, if they do not elaborate, a theory of relationships between function and form such that a competent language user changes many detailed features of his language as he moves from mode to mode. What details are changed and in what ways are not yet clear. As work such as that of Hunt (1965) and Loban (1969) has established changes of language forms with age, without attending to changes of function, it seems desirable to study these factors in relation to one another.

In the above context, the present study attempts to grapple with modes of discourse under controlled or experimental-type conditions, rather than in naturalistic field studies, which have so far predominated. Despite the dangers of artificiality, and the caution which needs to be exercised in applying results of such studies to field circumstances, this procedure is followed in order to explore whether there are predictable or otherwise discoverable patterns in the data when the following conditions obtain.

1. Samples of writing are elicited which are definitely in certain contrasting modes of discourse.
2. The subjects are of identifiable levels of competency in use of English on criteria extrinsic to the writing elicited in the study.
3. The samples of writing are evaluated by a standard procedure.
4. Methods are employed to grapple with the problem of age-norms for data other than data in the form of right or wrong answers.
5. Forms of language used are analysed and their relationships to the function of the sample of writing in which they occur are explored.

These steps are taken so that data can be gathered to refute or refine the hypotheses put forward, or, if the data do not refute the hypotheses, to confirm them to that extent. In this context, the present author is not as ashamed as he would otherwise be to operate on a simplified, even crude version of the subtle theories of Moffett and Britton which gave rise to the study.

## II. THE COLLECTION OF DATA ON FOUR WRITING TASKS

### 1. The Tasks

The present study is based on four writing tasks, one for each of Moffett's forms of discourse as subsumed into Britton's taxonomy (see pp. 9 and 15 ). These four tasks are, it is hypothesised, examples of -

1. Recording (Britton's 1.1.1)
2. Reporting (1.1.2)
3. Generalising (1.1.3)
4. Theorising (Subdivided by Britton into 1.1.4 to 1.1.7)

The term "tasks" is used in recognition that the writing used in this study represented set test-type exercises for the subjects rather than more authentic and spontaneous writing which would be expected to be encountered in field studies. While the subjects kindly complied with the request to supply the kinds of writing asked for, it is not to be assumed that the writing is particularly personal. This is why it is thought appropriate to classify it under the Transactional heading within the sort of category that is appropriate to responses to an examiner rather than the category of either the personal or the poetic. There may, of course, be something of the personal in the writing nevertheless, but it is not assumed that the writing means much to the subjects or that it is the best they could produce if otherwise motivated. Because of this, it is not assumed that

generalisations based on this writing can readily be transferred to other writing situations. Nevertheless, it is possible that competencies and adaptations revealed in a test situation may yield clues of assistance in further analysis, including analyses of field data.

## 2. A Writing Test

The tasks were presented to subjects on sheets of paper labelled honestly if uninvitingly, "WRITING TEST". The aim was to see if different cognitive and verbal operations could be discerned in tasks selected as nearly as possible to represent the classic modes of discourse which form the eventual basis of the models under discussion. These samples of writing were taken in a forty minute English lesson-period in schools.

Each of these four items is presented and discussed in turn.

### TASK 1 : RECORDING : A TELEGRAM

"You had planned to stay with an aunt in Melbourne, arriving at Flinders Street Railway Station at 7 p.m. on Saturday next. But your father has had a slight accident causing him a brief stay in hospital under observation. He is not seriously hurt but the accident has caused a change in your plans. You will now be arriving at Essendon Airport at 7.30 p.m. on Sunday. Without bothering about names and addresses, write the main message of a telegram making clear in the fewest possible words your change of plans and the reason for it."

The test paper gave the subjects four lines to write on.

It is hypothesised that selecting some information from given data and putting it down in the simplest possible form represents what Moffett and Britton might mean by "Recording". Cognitively, it is hypothesised, the task is the least demanding of those presented. The task is merely to select a few relevant facts and put them down simply. Verbally, it is hypothesised, the task is also the simplest. A distinction should be made, however, between the "telegraphese" of the small child operating with two- or three-word sentences and the "telegraphese" which is subject to conventions of word omission in saving money on telegrams. The criterion of brevity means that relatively competent subjects will attain a more elegant brevity than less competent subjects, who will be inclined to waste words.

TASK 2 : REPORTING (NARRATIVE) : A BED-TIME STORY

"Write the beginning of a story suitable for reading to a small child. There is no need to complete the story, but try to 'set it up' and 'get it going'."

Subjects were given 18 lines to write on.

The task is one of simple narrative, the conventions of which are well known to most children. This task will henceforth be referred to as "Narrative" as a more precise term than "Reporting" for the particular kind of discourse involved. It is hypothesised

that the task is cognitively more demanding than reporting, as a story must be structured from memory and imagination rather than merely involving a selection of items presented, as in the reporting task. It is hypothesised that this task is verbally more complex than Recording. Narrative involves more than the simple naming of objects and actions. There needs, for competent performance of the task, to be sufficient descriptive art to make an acceptable story. There is a restraint, however, on the elaboration of language, because by convention one keeps language simple for a small child. It is hypothesised that the more competent subjects will observe this rule more than the less competent.

### TASK 3 : EXPOSITION : THE RULES OF A GAME

"Select a game you know well. Explain the main rules of the game."

Subjects were given 18 lines to write on.

It is hypothesised that the task, in the form in which it is presented, is an example of generalised narrative/descriptive information. It could be claimed that in Britton's term the task is in the Regulative category, but it is here considered that the task is more about "This is what happens", and "This is how it works", than "You must do this or that" as an actual example of regulating someone's behaviour by orders, commands and requests. Cognitively, it is hypothesised, the task is more difficult than narrative, because it requires the ordering of

rules into some coherent framework. A narrative sequence will not suffice to produce a description of a game which conveys the main rules in the requested short passage. Verbally, it is hypothesised, the task is more complex than narrative, because of the necessity to define, make temporal, spatial and causal relationships explicit and to note exceptions and the like. Nor is there the restraint that a telegram should be brief or that a small child is the audience. Hence it is hypothesised that the more competent language users will produce more complex texts than the less competent.

TASK 4 : THEORISING : OPINION ON AN ISSUE

"The Australian Government has introduced a 'points score' for television, to impel channels to present more Australian productions, particularly at peak viewing times. Is this a good move, or is it undesirable interference with programming and viewing?"

The topic was widely publicised at the time the writing was collected and it was assumed that sixteen year olds generally would be able to frame a view on the issue.

It is hypothesised that the task is probably in the Britton category 1.1.6, Speculative. This involves the framing of hypotheses, though not dealing with them in terms of systematic theory. The latter would be in a higher category. Cognitively the writer must do more than generalise about a game he knows. He must think up reasons for and/or against the policy and take

some coherent line on the matter to write effectively in this mode. The argument requires that the generalisations must have some social validity, and not merely be a matter of personal likes or dislikes. Verbally, it is hypothesised, this is the most complex of the tasks, as matters need to be defined, qualifications need to be made and connections between ideas set out. It is hypothesised that the more competent subjects will produce more complex texts than the less competent produce.

### 3. Working Hypotheses

To sum up, the hypothesis underlying the selection of tasks predicts cognitive challenge and verbal complexity :

1. Reporting
2. Narrative
3. Exposition
4. Argumentation.

It further predicts that where simplicity is the norm, as in Tasks 1 and 2, the more competent language users will produce simpler texts, and that where complexity is the norm, as in Tasks 3 and 4, the more competent subjects will produce more complex texts.

### 4. The Subjects

The subjects were selected to comprise a not unrepresentative stratified sample of sixteen year olds still at school.

At the time in which the samples of writing were taken, there would have been some 100,000 sixteen year olds in the state of New South Wales. Some 70,000 were in the fourth form (now known as Year 10, Year 1 being age 6, the first year of compulsory schooling). Some 52,000 were in government schools as distinct from non-government (or rather government-aided) schools. All fourth formers were studying English under the Secondary School Board's syllabus.

The students of some 40 schools were taking an experimental Reference Test in English, data from which is used in the present study. The 40 schools were selected as a representative (not random) sample of schools of different types and locations. The purpose of the Reference Test was to give scores in English in order to moderate school assessments in the subject. This is done in the context of phasing out an external examination and replacing it by a short test to moderate assessments between schools as a means of adhering to some common standard.

The English Reference Test comprised multiple-choice questions in comprehension and usage, an essay on literature studied, and an essay in response to a picture stimulus. It is further described together with notes on its construction and marking, in Appendix A.

Of the 40 schools taking the Reference Test four were chosen for the present study to represent different types of schools and localities. All four schools were non-selective, this being the typical form of Government secondary schools in New South Wales.

One school was an inner urban co-educational school drawing its population from a crowded district with many small houses and flats. Another school was a semi-rural co-educational school in a small town with some farming interests. However it was also a dormitory town for a nearby industrial city. Two schools were in fairly affluent suburbs; one a boys' school, the other a girls' school. It is not claimed that this selection of schools is a random sample, but simply that it includes representation of kinds of schools and districts common in New South Wales.

The four schools were asked to provide an Advanced group and a Not-Advanced group for testing. The classification of students as "Advanced" or not was a formal part of the system of school assessment under the English Syllabus in operation. The syllabus (Secondary Schools Board, 1972) defined the aim of English as the utmost personal competence in the use of the language. It emphasised demonstrated competence in listening, speaking, reading and writing over a range of contexts such as everyday communication in various media, literature and personal expression, and evaluation in terms of the grasp of meaning in response to and control of the forms of language. The syllabus set out a variety of objectives comprising the formally endorsed state-wide criteria for evaluation of student performance. The syllabus emphasised practical competence in the use of language and explicitly excluded knowledge of such matters as a theory of grammar from the objectives. The Advanced were defined as the top 25% in English of the age cohort in the state. Schools formally assessed

candidates for a School Certificate award in English in these terms as part of the process of assessment, another part of

which was, as noted above, the use of a reference-test in an endeavour to maintain common standards. The syllabus also described the advanced in the following terms -

"... they generally read well, and speak and write with fluency and facility. They understand and use a wide range of language, including the more formal kinds, and often show imagination in the use of language. Advanced Level pupils discern a greater depth of meaning than other pupils, and in the later Forms ... especially, show greater recognition of, and capacity for abstract thought. They enjoy reading widely, and reading increasingly complex texts with increasing insight, going beyond literal level of understanding such as following the plot, to an appreciation of the issues in the text as a whole." (Secondary Schools Board 1973 p.17.)

This description was backed up operationally by lists of suitable texts, comprehension-tests with norms of performance, and scripts written by students at this level, as was done in the Examiner's Report of the 1971 School Certificate Examination (Secondary Schools Board 1972).

In so far as these procedures led schools to use common standards of assessment, the subjects put forward as Advanced would be in the top quartile of the age cohort in English, and the remainder of the subjects would not be in that quartile. It was widely believed that schools were operating on common standards of cumulative assessment because there were few controversies or appeals about whether students were Advanced or not, and results of School Certificate English awards were usually predictable on the schools' and candidates' part. (This indeed being a reason why it came to be widely believed that the examination system

could with some confidence be reformed.)

In each school, some 35 Advanced and 35 not-Advanced students were given the Writing Test in a forty-minute lesson period conducted by the writer for one class and the class teacher for the other. Thus some 280 subjects were tested. From this pool of subjects with complete records a table of random numbers was used to select a sample as follows -

Table 1 : Subjects of the Study

	<u>Advanced</u>		<u>Not-Advanced</u>	
	<u>Boys</u>	<u>Girls</u>	<u>Boys</u>	<u>Girls</u>
Inner Suburban School	8	8	8	8
Rural School	8	8	8	8
Suburban School A	16	-	16	-
Suburban School B	-	16	-	16

Thus the sample used in this investigation totalled 128 subjects, 64 Advanced and 64 not Advanced in English, with equal numbers distributed among schools and between sexes.

It is supposed that this sample is not unrepresentative of sixteen year olds in New South Wales schools, with the following qualifications. Advanced subjects, who by definition represent the top quartile in attainment in English, comprise half the

subjects in the study. Consequently, the sample parameters would not apply to the population as a whole. Further, the numbers of Advanced and Not-Advanced girls and boys are equalised, when it is common observation that at this age, girls tend to be more advanced than boys in the kinds of verbal task with which the study is concerned (see e.g. Connell et.al. (1975)).

For the purpose of the present study, all that is claimed of the sample is that it is not unrepresentative of the more advanced and the less advanced sixteen year-old English Students in New South Wales government schools.

#### 4. Evaluation of the Scripts

Each of the four pieces of writing completed by each subject was identified by a code number which was provided to conceal whether the subject was a boy or a girl, came from one school or another, or was Advanced in English or not.

The amount of writing done in response to each task is detailed in Appendix B. The mean total amount written per subject was 345 words. It was hypothesised that quite short pieces of discourse in distinct modes would reveal significant variations. This was the finding of the Pilot Study (Little 1973). The approach is thus in contrast with that of Hunt (1965) who obtained 1,000 words per subject without discriminating between modes of discourse.

Two experienced secondary teachers and examiners evaluated the scripts independently. The judges were not told what the present study was about. They were simply asked to evaluate scripts on the basis of the English Syllabus and School Certificate Examination and Reference Test marking criteria, with which they were well familiar (see Appendix A). The principle involved is to read and evaluate a script as a whole, attending equally to "what is said" and "how it is said", and comparing the script, as a communication, with the other scripts read. The marking is thus a ranking operation rather than a pass-fail one. It is also an instance of multiple-impression marking as recommended, on the evidence as most reliable for essays by Britton et.al. (1966), Godshalk et.al. (1966) and Maling-Keepes and Rechter (1973).

The scripts for each Task were evaluated separately from those from each other Task.

Marking was on a five-point scale with the following forced distribution -

- Grade 5: Script in the best 10%
- Grade 4: Script in the next 20%
- Grade 3: Script in the middle 40%
- Grade 2: Script in the next 20%
- Grade 1: Script in the worst 10%

The technique was to read through a set of scripts to gain a general understanding of the range and quality of the writing, and then to re-read the scripts physically sorting them into heaps representing the five-point scale. The heaps were counted

to see whether they conformed to the required distribution, with re-reading and re-sorting to obtain the distribution if it was not attained in the prior sorting. As Table 2 (p.53) indicates, the markers conformed closely to the expected distribution.

The correlation between markers for each task was obtained by using the Pearson Product-Moment Correlation Coefficient. Values ranged from 0.58 for the Recording task to 0.75 for the Argumentation task (see Table 2, p.53 for details). These correlations are significantly different from zero at the .01 level.

Each subject's eventual score for the task was the sum of the two marks given, the highest possible score being 10 points. The reliability of the combined judgment of the markers would be higher than that of either taken singly (Godshalk et.al. 1966).

The degree of reliability as indicated by these data is, as noted, significantly above zero, and is higher than reliabilities gained by Godshalk et.al. (1966). As these authors point out, reliability is increased above the level indicated by the correlation between markers when the results from the different markers are combined, as in the present instance.

Responses to the tasks were also analysed in terms of clause, phrase and sentence length. The procedures and data for that

part of the study are set out in sections IV and V below. At this stage, the concern is with data on achievement or quality scores and their interrelationships, and the endeavour is to identify the cognitive levels of tasks and performances of tasks, prior to an attempt to relate these to other variables.

In order to focus upon these concerns, a review of collection of data follows.

## 5. Review of Collection of Data

In summary, data is collected as follows -

Four writing tasks, one each in Recording, Narrative, Exposition and Argument, are completed by a not unrepresentative sample of sixteen year olds, with equal representation among four schools, between the sexes, and between more and less advanced English students as assessed by their schools on state-wide syllabus criteria.

The scripts are evaluated according to these criteria by two independent, experienced judges using a standardised five-point distribution. Data is also available on subject's scores from an English Reference Test taken by all the schools.

It is hypothesised that -

I. Scores on the tasks will prove not unreliable, and not invalid in terms of relationships with other data.

II. The tasks are in the rank order of difficulty for the subjects -

1. Recording (easiest)
2. Narrative
3. Exposition
4. Argument (most difficult).

III. The subjects will produce the simplest language in response to Task 1 (Recording), more complex language in response to Task 2, more complex language still in response to Task 3 and the most complex for Task 4.

IV. The more advanced subjects will produce simpler language than the less advanced subjects on Tasks 1 and 2 and more complex language than the less advanced subjects on Tasks 3 and 4.

### III. DETERMINATION OF THE COGNITIVE LEVELS OF THE TASKS

#### 1. General Rationale

As noted on page 21, there are problems in placing a cognitive function at a developmental level when the function is not one which yields right or wrong answers. The rationale of the attempt to solve this problem in the present study is as follows.

The typical method for a task with right or wrong answers is that followed by Connell et.al. (1975) in their work on the onset of formal thought about verbally-presented humanities-type problems. Essentially, the method is to present the problem to subjects representing a number of age groups. The percentage of each age group giving the right answer is then determined. The task is placed developmentally at the age level at which 50% of the age group give the right answer. On this method, as used in the Connell study, 17 to 18 emerges as the age at which 50% of school students reveal ability to deal with formal thought about humanities-type problems presented verbally. Lower percentages of younger people show themselves able to operate at this level, and higher percentages of older people.

Generalising, the principle is that a task is placed at that age level at which it discriminates maximally between subjects.

As Ebel shows (1965 : 355) a group score of 50% correct yields maximal discrimination indices. A lower group score than 50% right means that the task is less discriminating because it is relatively difficult. A group score of higher than 50% means

that the task is less discriminating because it is relatively easy.

The kind of group score obtained when 50% of an age group gives a right answer is not the only type of discrimination index available, however. Other discrimination indices exist which might also be used for the purpose of finding the age levels at which cognitive functions may be maximally discriminating. Before discussing these, it seems appropriate to generalise the principles involved in a model.

Figure 3 : Model of Age-Placement of Tasks by Use of Discrimination Indices (D = Discrimination Index).

Level of Task	Age Level (Hypothetical)			
	4	8	12	16
High Level (Argument)	Very Low D	Low D	Moderate D	<u>High</u> D
Moderately High Level (Exposition)	Low D	Moderate D	<u>High</u> D	Moderate D
Moderately Low Level (Narrative)	Moderate D	<u>High</u> D	Moderate D	Low D
Basic Level (Recording)	<u>High</u> D	Moderate D	Low D	Very Low D

Note : It is assumed that the discriminations in the top left-hand corner are low because the task is too difficult for the subjects and that the discriminations in the bottom right-hand corner are low because the task is too easy for the subjects.

One index of discrimination other than the percentage of an age group getting a right answer to a problem is provided by the degree of reliability with which qualified judges can agree in discriminating levels of competency in the performance of a task. Other things being equal, a task which is too easy or too difficult for an age-group will present the judges with performances which are very much alike because they all tend to be either very competent or very incompetent. If the task is one in which some of the age group are competent and some are not, the judges will be presented with a wider variety of performances which they will better be able to discriminate. The fact that discrimination of this kind is maximal for a given age group provides some grounds for reasoning that the task is placed developmentally at that age level.

Thus, not only percentages gaining right answers, but correlations between markers could provide discrimination indices in Figure 3.

Another kind of index of discrimination is the relationship between performance on a task and some trusted criterion measure. For instance, an investigator may have evidence that a measure such as mental age or teachers' ratings of students' competencies in some field represents a reasonable criterion measure of the broad cognitive differences being investigated. The task is then to discover whether some particular cognitive operation under investigation bears a

close relationship to the criterion variable. One way of measuring such a relationship is to determine the correlation coefficient between the particular cognitive operation and the criterion variable. A task that was either too high or too low in level for a given age-group would, other things being equal, yield a lower correlation with the criterion measure than a task which discriminated maximally at that particular age level. Such correlation coefficients could provide further discrimination indices in Figure 3.

A third way of measuring the relationship between a particular cognitive operation and a criterion variable is provided by techniques such as determination of the significance of the difference between means for groups selected as relatively high and relatively low on the criterion variable. Indices of magnitude and significance of such differences could provide a further discrimination index in Figure 3.

Such indices are, of course, indirect measures. They depend on "other things being equal", and would require verification by one another and further verification in such matters as whether discrimination is relatively low because performances are relatively good or relatively bad on the whole. Nevertheless it may be argued that the evidence provided by a number of indirect measures combined, circumstantial and open to alternative interpretations though it may be, is probably worth gathering as a way of verifying subjective impressions about degrees of difficulty; of opening up a difficult field; and possibly

suggesting more precise hypotheses and techniques in the process.

The part of the model being verified in the present study is the column for sixteen year olds in Figure 3, though some reference is made to similar data on other age groups (pp.114ff).

## 2. Discrimination Indices

In the present study, three discrimination indices are used to gain evidence on the relative discriminative power of the four writing tasks. The indices are as follows -

### 1. Inter-marker Reliability.

This index is provided by the Pearson product-moment coefficient for the marks awarded by the two independent judges.

### 2. Correlation between Score on Task and Reference Test.

This index is provided by the coefficient for the score on each task and the English Reference test taken by all subjects.

### 3. Difference between Mean Score of Advanced and Not-Advanced Subjects on the Task.

In the circumstances of a forced standard distribution of scores for each task for the subjects as a whole, the mean difference itself provides a suitable measure, but the size of an index such as  $t$  in the  $t$ -test of significance of difference between means could also provide such an index.

In more detail for each criterion, the rationale is as follows -

Index 1 : Inter-Marker Reliability.

In terms of the Mcffett-Britton model for the four forms of discourse studied and related findings (e.g. Martin 1975, Connell et.al. 1975) it is predicted that the four tasks would discriminate between sixteen year old subjects in the ascending order,

1. Recording (discriminating, but least discriminating of the four tasks),
2. Narrative,
3. Exposition,
4. Argument (most discriminating of the four tasks).

It would be expected that most sixteen year olds could write a competent telegram based on the given data. For such subjects, there is no great cognitive challenge in such a task, though some may well handle it better than others. On the whole, however, the markers will be presented with a mass of competent telegrams, and will have more difficulty than with other tasks in discriminating reliably between the scripts. With the slightly more challenging task of Narrative, there will be more difference, more discrimination. There will be more difference and discrimination still on the more challenging task of Exposition, and the most difference and discrimination will occur in the case of Argument.

This, it is suggested, is a functional explanation of why a

formal essay has so often been part of public examinations and tests in English for people of about the age of the present subjects. Simpler tasks are not such good discriminators because they are too easy. A formal essay in the transactional mode, to use Britton's terminology (see page 15), is a good discriminator at this level because it represents, so to speak, a developmental growing-point among the subjects such that about half have and half have not developed the necessary competence. Such formal, transactional essay-writing would be an indiscriminating test for eight year olds because it is too difficult. For them, Recording or Narrative would be a better discriminator. By the same token, tautologic theory at Britton's level 1.7 would provide poor discrimination among sixteen year-olds because it would be far too hard, representing as it does an undergraduate or graduate type of performance, although some younger people may well be found who could handle such discourse at least in some limited field.

The main prediction is thus that Argument will provide the best discrimination by providing the highest correlation between the two markers. Supplementary predictions are that the correlations between markers will be next highest for Exposition, next highest for Narrative and lowest for Recording. The correlations are given in Table 3, page 55 following the discussion of the other discrimination indices.

## Index II : Correlation between Score on Task and Reference Test.

The English Reference Test used in the present study is set out in Appendix A. It was designed to give an index of competency in English in terms of the Syllabus under which students and teachers were working at the time. The test was aimed at discriminating maximally between the more and less advanced subjects on the same criteria by which subjects were assessed for competency in English by their schools. This was done in order to verify school assessments. The reliability and validity of this type of test is discussed elsewhere (Little 1973).

If a task is a good discriminator of competency in a form of discourse at the sixteen year old level, it ought to correlate maximally with such a reference test. If the task is either too hard or too easy to provide good discrimination, it ought to correlate minimally with such a reference test.

The predictions for this criterion are therefore similar to those for the first criterion: Recording should yield the lowest correlation with the reference test, Narrative a higher correlation, Exposition a higher correlation still, and Argument the highest of the four correlations.

Obviously Criterion 2 is linked with Criterion 1 because the marks awarded by the two markers are involved in both. Even so, the cases are somewhat different. The marks correlated for Criterion 1 are those of the two markers taken separately.

Those for Criterion 2 are -

1. the marks given by the two markers combined, the total marks presumably being more reliable than those of each marker taken singly and
2. score on the Reference Test, which is independent of all other measures taken in the present study.

Thus, although there is the common element of marks from the two markers, there are other data involved which give wider opportunity for the hypotheses to be refuted.

Index III : Difference between Mean Scores of Advanced and Not-Advanced Subjects on the Task.

For each of the four tasks, the means for the Advanced and Not-Advanced groups, identified as such by their schools on the broad criteria set out by the English syllabus in operation throughout the state, have been determined.

If the task is a good discriminator at the sixteen year old level, the difference between the means of the Advanced and Not-Advanced groups ought to be maximal. If the task is not a good discriminator in this sense, the mean-difference should be minimal.

Again the marks awarded by the two markers are involved, as for Criterion 1 and 2. As in the case of Criterion 2, however, it is the combined marks and not the marks of each marker taken separately which are involved. Additional data is provided by

the school's identification of subjects as Advanced or Not-Advanced, which is data independent of both the two markings, the the Reference-Test, and gives further opportunity for the hypotheses involved in the study to be refuted.

It is relevant to note again that the quality-scores for each task were awarded by two markers using a forced distribution with a mean of 3 points on a 5-point scale (see p.35 ). This makes the mean of the eventual mark 6 points out of a 10-point scale. Because the distribution for each task is the same, the difference between the means of the Advanced and Not-Advanced subjects gives a measure of the relative degrees of discrimination of the tasks. If there is no discrimination, the means of the groups will each be 6 points. If there is some discrimination, the mean of the Advanced will be higher in proportion to the degree to which the mean of the Not-Advanced group is lower. The forced distribution means that for every case of an Advanced subject gaining a higher score there will be a corresponding case of a Not-Advanced subject gaining a lower score.

The predictions are that the size of the difference between mean quality scores for the Advanced and Not-Advanced groups will occur in the rank order: least difference for Recording, more for Narrative, more still for Exposition and most for Argument.

### 3. The Three Indices Considered Together

It is here maintained that the three criteria provide indirect measures of the relative discriminatory power of the four tasks at the sixteen year old level, and potentially at other age levels, at which the predictions would take the form of different rank orderings.

It is possible that on some indices there may be no significant difference. For instance, it is possible that by one of the criteria, Argument may prove to be the most discriminating task, but there may be no significant difference between the discriminative power of, say, Recording and Narrative. The hypothesis that Argument is the most discriminating task would be upheld, and the hypothesis that Narrative is more discriminating than Recording.

In this sense, the major hypothesis can be broken down into minor hypotheses as follows.

Predictions regarding the Recording tasks are as follows -

1. Recording will be less discriminating than -
  - a) Narrative
  - b) Exposition
  - c) Argument.

This is really three hypotheses, any one or any combination of which could be confirmed or refuted.

2. The above three hypotheses are applied to three separate discrimination indices, making the total number of hypotheses for the task nine, any one of

which could be confirmed or refuted. Thus, for four tasks, the study is dealing with  $4 \times 9 = 36$  hypotheses, any of which could be confirmed or refuted. As the direction of difference is hypothesised, one-tail tests of significance are employed (Popham 1967).

3. The possibilities of confirmation and refutation are as follows -

a) The Null Hypothesis may be upheld, there being no significant difference between two tasks compared, in respect of the criterion upon which they are being compared. In the context of the present study this is taken to mean that the tasks are equal with respect to that particular criterion.

b) The possibility remains open that the tasks are significantly different in the predicted direction on some other criterion. The first, "neutral" result does not cancel out the second, "positive" result.

c) The third possibility is that the tasks are significantly different in the direction opposite to that predicted. This would be taken to be a significant "negative" result if significant on a two-tail test.

Thus, while the present study uses indirect measures for which alternative explanations are always possible, there are rational grounds of placing some credence on results in the form of a combination of measures hypothesised to be in concordance in revealing differences in a consistently predicted direction,

and open to "negative", "neutral" and "positive" results.

This is stated in the context in which a number of independent sources of data are involved -

1. the assessment of the subjects by their schools as Advanced or Not-Advanced;
2. the evaluation of their performance on the tasks by two separate judges;
3. the analysis of their language by two research assistance other than the judges;
4. the provision of the English Test by a separate research body.

The present study co-ordinates these data, the present author having planned the procedures but having no hand in making the particular assessments, evaluations or analysis which provide the actual data. In addition, none of the participants was informed of the particular hypotheses being tested until after the data was processed.

#### 4. Results for Achievement Scores

Findings for this part of the study are presented first in terms of means and correlations of the various achievement scores, from which the discrimination indices are derived for subsequent analysis.

TABLE 2 : RELATIONSHIPS AMONG ACHIEVEMENT SCORES FOR 128 SUBJECTS

Relationship	No. of Subjects	English Reference Test	Sum of Quality Scores for the 4 Tasks	Quality Scores for Each Task			
				Recording	Narrative	Exposition	Argument
Pearson product-moment correlation ( $r$ ) between the two judges	128	-	-	0.58**	0.57**	0.69**	0.76**
$\bar{x}$ with English Reference Test	128	-	0.58**	0.21 *	0.38**	0.47**	0.54**
Mean (and standard deviation) of Scores	64 Advanced 64 Not-Advanced	53.8 (10.7) ** 46.4 (10.9)	27.1 (5.17) ** 21.0 (4.88)	6.31 (1.86) ** 5.70 (1.99)	6.92 (1.72) ** 5.05 (1.73)	6.78 (2.07) ** 5.14 (1.66)	7.01 (1.89) ** 5.00 (1.69)

\* Statistic significant at the .05 level.

\*\* Statistic significant at the .01 level.

Significance was determined by the  $t$  test in each case (Popham 1967 chapter 10).

Table 2 indicates results of considerable significance, statistically and in their import in the context of the hypotheses and methods employed in the study.

The reliability of the markers, as indicated by the correlation of the two marks for each task, was significantly higher than zero, at the .01 level of confidence. By implication, the reliability of the sum of the two marks would be higher than that of either of the two marks.

The correlation between the sum of scores on the four tasks and the English Reference Test is also significantly different from zero at the .01 level of confidence. So is the correlation between each task taken singly and the English Reference Test, except that correlation for Recording, which is significant at the .05 level.

The English Reference Test, the sum of scores on the four tasks and the score for each task all discriminate between the Advanced and Not-Advanced groups at the .01 level for significance of difference between means.

These data are consistent with the hypothesis that the evaluation of the scripts was not unreliable, and that the measures of competence on the tasks, on the test and on the school assessment are, though independent of one another, not uncorrelated. In particular, it would appear that scores on the tasks have some validity as indicators of competency in English. The study is

thus dealing not with random or chance data, but with statistically significant relationships in the predicted directions.

### 5. Results for the Three Discrimination Indices

Table 3 gives the three discrimination indices for each writing task upon which the rationale of determination of the cognitive levels of tasks is based.

Table 3 : Three Discrimination Indices for Four Writing Tasks

Tasks	Discrimination Indices with Actual Order (1) to (4)		
	I. Correlation between Markers	II. Correlation between score on tasks and Reference Test	III. Difference between means of Advanced and Not-Advanced Groups (Means given in brackets, Advanced first)
In predicted order of discrimination, lowest (1) to highest (4)			
(1) Recording	(2) 0.58	(1) 0.21	(1) 0.61 (6.31-5.70)
(2) Narrative	(1) 0.37	(2) 0.38	(3) 1.87 (6.92-5.05)
(3) Exposition	(3) 0.69	(3) 0.47	(2) 1.64 (6.78-5.14)
(4) Argument	(4) 0.76	(4) 0.54	(4) 2.01 (7.01-5.00)

The general trends of the data are in line with the predictions : perfectly so except for the reversal of ranks 1 and 2 on the Index I and ranks 2 and 3 on Index III.

To accept these rankings without some test of statistical significance would, however, be less satisfactory than to apply some appropriate statistical tests. The tests applied and the results are as follows.

The significance of the differences between correlations was determined by finding the confidence interval of each correlation at the .05 level of confidence, using the standard error formula for a one-tail test with the number of cases 128 (Popham 1967 p. 92). For instance, at the designated level of confidence the first two correlations in the first column of Table 2 (Coefficients 0.58 and 0.57) have the same probable upper and lower limits so that the difference between them is not significant. But the third correlation (0.69) has a higher probable lower limit than the probable upper limit of the other two correlations, so that the difference between this correlation and the others is significant at the .01 level. Similar comparisons were made for the remainder of the correlation coefficients, and the results are given in Table 4 below.

For the difference between means, the following approach was taken. Using the confidence-interval method for the means of the Advanced group (A) and the Not-Advanced group (N) it was noted whether the mean for the Advanced group on Task 1 (A1) was significantly lower than the same group's mean on Task 2 (A2), which proved to be the case ( $A1 < A2$ ). Then it was determined whether the mean for the Not-Advanced group (N1) was significantly higher on Task 1 than on Task 2 (N2), which also proved to be the case ( $02 < 01$ ). As  $A1 - N1 < A2 - N2$ , it was concluded that the difference between group means on Task 1 was significantly less than the difference between means on Task 2. The same procedure was followed for the other tasks.

Table 4 gives results of tests of significance of differences between discrimination indices. The entry "0" means that there was no significant difference. The entry "+" means that the difference was in the predicted direction and significance at the .05 level. (A "minus" entry reserved for any occurrence of a significant negative difference, but as there was no such occurrence the entry does not appear in the table.)

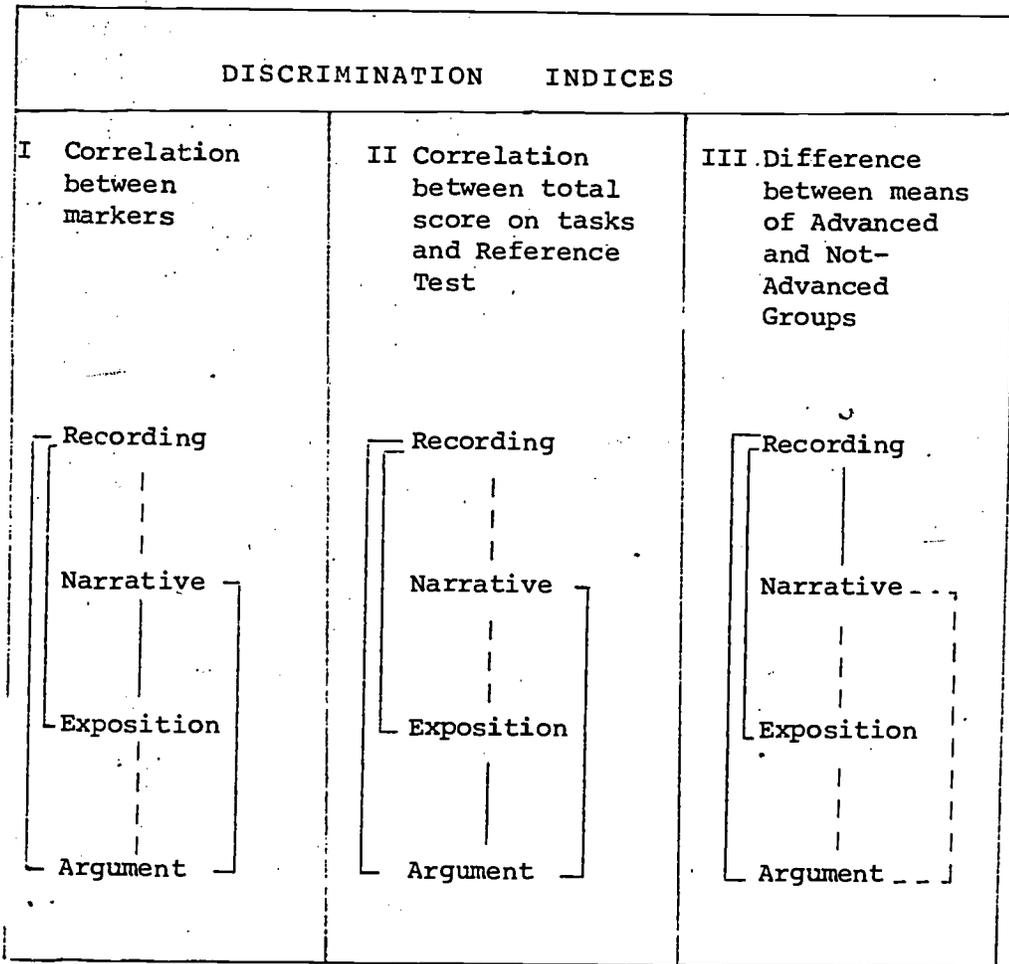
Table 4 : Statistically significant differences between discrimination indices for four writing tasks completed by 128 subjects.

"+" = difference in the predicted direction significant at the .05 level.

"0" = difference not significant.

Tasks Compared	DISCRIMINATION INDEX		
	Correlation between markers	Correlation between total scores on tasks and Reference Test	Difference between means of Advanced and Not-Advanced groups
Recording and Narrative	0	0	+
Recording and Exposition	+	+	+
Recording and Argument	+	+	+
Narrative and Exposition	+	0	0
Narrative and Argument	+	+	0
Exposition and Argument	0	+	0

Figure 4 : Diagrammatic representation of significant differences (continuous lines) and non-significant differences (broken lines) between three discrimination indices for four writing tasks completed by 128 subjects.



## 6. Interpretation of Results

On each criterion, each task may be compared with any other, producing for each task nine indices of discrimination. A significant difference between one task and another in the predicted direction is taken as positive evidence. The lack

of a significant difference is taken as neutral evidence: on the given criterion the tasks were not significantly different; i.e. they were equal as discriminators. "Neutral" evidence is not taken as "negative" evidence. If tasks are not different on one criterion, they may yet differ on another criterion. Thus it can be stated on what criteria tasks were significantly different as discriminators and on what criteria they were in effect, equal. In other words, the statement that two tasks are different means that they are significantly different on at least one of the presented criteria.

a) Recording

Although the Recording task proved, as predicted, to be a statistically significant discriminator, it also proved, as predicted, to be the least discriminating of the four tasks.

It proved to be less significantly discriminating than both Exposition and Argument on all three criteria employed. Two of the criteria failed to separate it from Narrative as a discriminator, but the third criterion (difference between Advanced and Not-Advanced groups) did separate it from Narrative, in favour of Narrative as the better discriminator. Thus, on seven of the nine indices, Recording proved to be an inferior discriminator, two indices being neutral on whether it was inferior to Narrative.

Taking all criteria and indices into account, it may be concluded that the evidence is that Recording is the least discriminating

of the tasks. It is strictly in the sense of the above statements that it is taken to be so in the remainder of the present study. The same rationale applies to conclusions for the other tasks.

b) Narrative

As indicated for Recording, two indices are neutral on whether Narrative is a better discriminator than Recording, but the third index comes down in favour of Narrative.

For the remaining six indices the evidence on Narrative is as follows. On Criterion I, Narrative is significantly less discriminating than both Exposition and Argument. On Criterion II it is significantly less discriminating than Argument. The remaining indices do not separate Narrative from Exposition or Argument.

The sum of evidence - four out of seven indices positive, the remainder neutral - is to the effect that Narrative is more discriminating than Recording and less discriminating than the other two tasks.

c) Exposition

Six of the nine indices point to Exposition as a better discriminator than Recording and Narrative but an inferior discriminator to Argument. Only Index III separates Exposition from Argument and only Index I separates it from Narrative, but in all comparisons at least one indicator places it in the predicted

rank, while others are neutral.

d) Argument

On six of the nine criteria, Argument is the best discriminator, the remaining three indices failing to separate it from Exposition in two cases and Narrative in one. On the whole, Argument is thus the most discriminating task, taking all criteria into account.

7. Conclusions

Of 18 comparisons (six comparisons of each task with every other on three criteria), 11 show significant differences in the predicted direction, and seven fail to discriminate between tasks as discriminators. There are no negative instances (significant differences in the direction not predicted). With these qualifications in mind, it is concluded that, as predicted for sixteen year olds, there are objective evidential grounds for stating that the Recording task, although it is discriminating, is the least discriminating of the four tasks, that Narrative is more discriminating, Exposition more discriminating still and Argument the most discriminating of all.

If the proposal is accepted that discrimination indices yield an indirect measure of age level placement of a task, then Argument, as the most discriminating task, would appear to be nearest to the age level of the sixteen year old subjects, followed by Exposition, then Narrative, and finally, Reporting.

As the data shows that there is more difference in performance between the Advanced and the Less Advanced the more difficult the task, there are objective grounds for stating that the tasks other than Argument are less discriminating because they are, developmentally "younger" tasks, competencies acquired earlier by the subjects as a whole. Argument appears to be the task that enables greatest discrimination because the subjects are in the process of mastering it, some having done so and others not having done so to any considerable extent.

This hypothesis may be further tested by scrutiny of representative scripts. Some are presented at the end of Part V below.

The conclusion is that on the combined criteria the tasks are in the rank order of discriminative power for sixteen year olds that was predicted, and that to this extent the Moffett-Britton model of developmental stages in modes of discourse is verified by methods other than purely literary-critical methods.

It appears that the methods of indirect measurement used produce sufficiently consistent and significant results to provide some basis for further studies in which performance (other than performance in terms of a right or wrong answer) can be placed on age scales on the basis of their power to discriminate.

In this instance the hypotheses and methods were tested only for one age group. In further studies it is hoped to apply such

techniques to other tasks and other age groups. It is relevant to report that in an unpublished pilot study related to the present investigation (Davis 1975), discrimination indices were found by Criterion I. of the present study (correlation between markers) for Narrative and Exposition written by fifty 8 year olds and fifty 12 year olds. For 8 year olds the correlation was higher for Narrative and lower for Exposition. For 12 year olds, the correlation was higher for Exposition and lower (though only slightly so) for Narrative.

In terms of the model presented in Figure 3 (p. 41) findings so far confirm the predictions. Figure 5 repeats that model with the addition of such data as is at present available.

Figure 5 : Preliminary findings on Age Levels as indicated by discrimination indices for various age levels and writing tasks. Data in the first two columns is from Davis (1975) with permission, using correlation between markers as a discrimination index, and data in the third column is from the present study, using three criteria combined (see p. 50).

D = Discrimination. The arrows show indices compared which have confirmed predictions.

TASK	AGE LEVEL		
	8	12	16
Argument	Very low D	Moderate D	High D
Exposition	Low D	High D	Moderate D
Narrative	Moderate D	Moderate D	Low D
Recording	High D	Low D	Very Low D

To this extent it appears that competence in the handling of Recording precedes competence in Narrative, which precedes competence in Exposition, which precedes competence in Argument. Narrative may be placed somewhere around the eight year level, Exposition around the twelve year level and Argument around the sixteen year old level, as a general rule, the exceptions to which also need to be defined.

More data needs, of course, to be gathered in further studies before much confidence can be placed in the model set out in figure 3, but so far the hypotheses appear to be confirmed and the techniques seem to be capable of yielding highly-patterned and statistically significant results. It is suggested that by the kind of process of iteration represented in the present study

a more precise, objective and comprehensive view of development in competence to handle modes of discourse than we now possess could be built up.

In the remainder of the present study, it is accepted that the tasks of Recording, Narrative, Exposition and Argument are in ascending order of cognitive challenge in the sense defined on page 61. On this basis, relationships between the cognitive level of these functions of language and the forms of language used to serve these functions are explored.

#### IV A METHOD OF DETERMINING COMPLEXITY OF VERBAL RESPONSE

##### 1. Function and Form

Accepting that the tasks of Reporting, Narrative, Exposition and Argument are in ascending order of cognitive challenge, the present study proceeds to explore whether this order of cognitive difficulty is matched by a corresponding order of verbal complexity.

In the discussion of the Moffett and Britton models of forms of discourse (p.7 ff.) it was hypothesised that successive modes of discourse from the most concrete and intimate to the most abstract and remote call upon successively more complex feats of cognitive operations and verbal elaborations. Hunt (1965) and Loban (1969), dealing with written and spoken language respectively, have established a connection between increasing age of children and the length and complexity of their sentences. This was so although the mode of discourse was not systematically controlled. Even so, as Hunt points out (1965 p. 3) young children tend to write personalised narrative no matter what the ostensible topic, while older children are more likely to write other forms of discourse as well. The present investigation seeks verification of the hypothesis that there is not merely a general association of more complex discourse and more complex verbalisation with age, but that increase in age is associated with more complex forms of verbalisation arising from the changing

nature of the discourse undertaken. On this hypothesis, simpler forms of discourse associated with simpler forms of verbalisation do not disappear, but remain in the person's repertoire to be engaged in as appropriate. On such a view, simple discourse and verbalisation are not to be dismissed as "immature", even though immature persons may be limited to these. Simpler forms remain appropriate to some communication situations, as more complex forms evolve as appropriate to others. Verbal competency thus involves discrimination of when to use simple form, and when complex.

Such variations are illustrated by Arthur's (1973) case of a small girl who tells the story of Cinderella racyly in a friendly conversation, and changes diction and sentence-form to produce a much more elaborate and formal version when asked to tell the story into a tape recorder. The interpretation is that even though the child has not yet learned to read, she has acquired informal conversational and formal storytelling forms of discourse, and switches readily from one to the other as appropriate. Such variation is also reflected in a pilot study for the present investigation (Little 1973) in which eight year olds described some pictures and explained the rules of a favourite game. Analysis of taped transcripts revealed that the children produced significantly more elaborate phrases and used significantly more subordinate clauses in the explaining situation.

It must be confessed that there is not as yet any highly

systematic theory of the forms of language associated with various forms of discourse among developing users of the language. In a sense, this part of the present study is "inductive", in contrast to the more "deductive" earlier sections of the study, in which deductions from established models provided the framework of the investigation. On general theoretical grounds one would suppose that grammatical or syntactic differences would play some role in the matter, as is evidenced by readability studies in which grade levels can be assigned to discourse in view of factors such as sentence length (Gilliand 1972). It is these, rather than semantic differences which are under investigation in the present study. If systematic patterns are found at the grammatical level, it may be possible in future explorations to investigate semantic variables along similar lines, in relationship to what is known of the cognitive and the grammatical factors at work.

## 2. Analysing the Forms of Language Used on the Tasks

As analysis of language at sentence level, clause level and phrase level have produced various patterns of significant results (Hunt 1965, Loban 1969, Little 1973), it was decided to analyse the language produced for the tasks at each of these levels, as well as taking account of the incidence of "and", on the hypothesis that its frequent use is a sign of syntactic immaturity. This decision also reflects the state of knowledge in the field in casting the net as widely as possible, so that various possible trends may appear because as yet there is no

systematic theory relating modes of discourse to verbal complexity.

3. Word-Counts and their Limitations

Word-counts may be objected to on the grounds that words have such different functions that to count them as if they are equivalents may be misleading. Some differences of function include the following -

a) Structural and Lexical Words

In the structural type of grammar associated with Fries (1952) a distinction is made between structural words which carry little "dictionary meaning" and lexical words which carry more of such meaning. Nonsense writing illustrates the difference.

'Twas brillig and the slithy toves  
Did gyre and gimble in the wabe ...

(Carroll)

The structural words are 'Twas ... and the ... did ... and ....the'. Structural words are all words except nouns, verbs, (other than auxiliary verbs), adjectives and adverbs. They form a limited, finite class not open to new coinages and provide a sentence frame into which lexical words may be fitted.

In the example, the lexical words happen to be nonsense words. In the frame given by the structural words, brillig is clearly a noun, slithy an adjective, gyre and gimble are verbs and wabe another noun. The class of lexical words is is more numerous than the class of structural words, and

open to new coinages, which are not confined to nonsense words. These words name the objects, actions and qualities that the discourse is about.

On word-counts, sentences high on structural words and low on lexical words and sentences with an opposite pattern will be equivalent, though it may be doubted whether they really are so. A long sentence in very plain diction would by word-counts be the equivalent of a long sentence in very ornate diction. Word-counts could be said to obscure differences in focussing on superficial similarities.

b) Common and rare words and uses of words.

Tom had a big dog is composed of the same number of words as A dangerous weapon is irony, and again it may be doubted whether the sentences are really equivalent. The differences are clearly related to the level of abstraction, the use of metaphor, and the inversion of normal word order.

Word-counts therefore ignore certain grammatical and semantic differences which may be of major significance to the communication as a whole. This does not invalidate word-counts, however, any more than the lack of colour invalidates black-and-white photography or that in measuring the volume of objects one is ignoring their weight and texture and various other properties. The point is to see how far word-counts yield results of significance, without supposing they represent the only significant factor at work in discourse.

#### 4. Words

Counting words involves defining "word". The approach adopted was to regard as a word what subjects showed to be a word by leaving a space between letters in their writing. As the focus was upon simplicity and elaboration, contractions such as don't, being more complex than do, were counted as two words; hyphenations such as half-asleep were counted as two words; and complex verbs such as would have been finished were counted as four words.

#### 5. Sentences

The sentence is notoriously difficult to define, but the decision was made in the present study to take as a sentence what the subject offered as a sentence on the evidence of the punctuation. In the case of the Recording task, in telegram form, a sentence was taken to be a set of words which a normal writer of English would recognise as one by placing a full stop or the word STOP in the appropriate position. On the other tasks, semi-colons and similar punctuation signs were taken as not representing the end of a sentence, as is the case in various readability formulae (e.g. Flesch 1948), but as revealing relationships between utterances more complex than those signalled by a full stop. The greater complexity would thus be registered by counting the words on both sides of the semi-colon as part of one sentence.

Observation of the writing sampled shows that there are few run-on sentences and fewer punctuation signs such as semi-colons.

The sixteen year old subjects are very close indeed to normal adult practice in marking off sentences, and it was considered that less harm would be done by accepting their prose as offered than by introducing arbitrary principles of sentence division, except in the relatively simple case of the telegram.

The alternative was considered of operating in terms of Hunt's T-unit (Hunt 1965) which is essentially a principal clause together with its dependent subordinate clauses. In Hunt's form of analysis, a co-ordinate principal clause starts a new T-unit, so that the sentence Tom had a big dog and Jane had a little kitten would be two T-units. Hunt uses this device to distinguish between long sentences produced by immature subjects using the construction "and ... and ... and ..." and mature subjects using other constructions. In the present study, this aspect of the texts is checked by counting the incidence of and, making it possible to consider this factor while operating with the normal conventions about the sentence.

## 6. Clauses

Clauses present no special difficulties of definition. The conventional grammatical notion of a clause as marked by a finite verb and its associated words is followed throughout. This produces principal clauses, adjectival clauses and noun clauses, together with co-ordinate clauses of all types. A sentence such as "He believed that he was right" is counted as two clauses, dividing between believed and that. The sum of words in clauses so defined equals the sum of words in the sentence.

## 7. Phrases

The Pilot Study for the present enquiry found significant differences between describing and explaining on the basis of a word-count of phrases. In case there were differences between patterns of discourse at phrase, clause and sentence levels, a phrase-count was included.

For the purposes of the present study, a "phrase" is defined as follows - a), b) and c).

- a) A phrase is a set of words other than a clause (the clause including a finite verb, and the phrase not doing so).
- b) A phrase is a set of words acting as an adjectival or adverbial unit, such as the groups underlined in the following passage.

One day, Goldilocks went for a walk in the forest,  
taking a basket of food with her.

One day is a typical time phrase of which later on,  
at six o'clock, the results having been published (as an  
absolute phrase) are examples.

Taking a basket is a typical phrase beginning with a non-  
finite verb, in this case a participle. Such phrases may  
also begin with infinitives, of which to see her grandmother  
would be an example.

The remainder of the phrases underlined in the passage are  
typical propositional phrases: for a walk, in the forest,  
of food, with her.

Single adjectives, adverbs, infinitives or participles count as single words in some large unit. It is only when more than one word operates as a unit that a phrase is taken to occur. Thus the word there in He sat there is counted only as a word, but in the sentence He sat on a log the phrase on a log functions in place of the single word there and is counted as a phrase. Similarly singing in She walked along, singing, is counted as a word, but singing to herself or singing a nursery rhyme is counted as a phrase. Again, She went to sleep is counted as single words (the to being taken as the infinitive verb-marker, and not a preposition) but in She decided to eat her sandwiches to eat her sandwiches is taken as a phrase.

As so far defined, then, the phrase-counts include absolute phrases, phrases beginning with non-finite verbs and phrases beginning with prepositions.

- c) So far the definition of "phrase" has been that of conventional grammar. For the purposes of the present study an unconventional step is taken in including one further unit in the definition. This unit is what is left over when all the phrases as so far defined are identified. In the sentence One day, Goldilocks went for a walk in the forest, taking a basket of food with her, all words may be allocated to phrases as conventionally defined except the main subject and the verb, Goldilocks went. For present purposes this, too, is regarded as a phrase, so that all words are allocated to phrases for

the purposes of the study. This enables the following formula (suggested by Hunt's formulae for clauses and other units, 1965) to operate :

$$\begin{array}{rcccl} \text{No. of} & & \text{No. of} & & \text{No. of} \\ \text{words} & & \text{phrases} & & \text{words} \\ \text{per} & \times & \text{per} & = & \text{per} \\ \text{phrase} & & \text{sentence} & & \text{sentence} \end{array}$$

An alternative to counting the main subject and its verb as a phrase would be to divide that unit into a Noun Phrase and a Verb Phrase as is done in some grammars (e.g. Fries 1952 Lyons 1970). In such cases the subject Goldilocks (or if there were adjectives, the set of words Pretty little Goldilocks) would count as a phrase, and so would the verb went (or a set of words such as went gaily). This alternative is not taken for the following reasons.

If an expression such as Goldilocks went were to be divided into a noun phrase and a verb phrase, consistency would seem to demand that other noun phrases should be divided from their neighbouring words. Noun phrase objects would then have to be divided from their finite verbs, non-finite verbs or prepositions. Then these finite verbs, non-finite verbs and prepositions would stand as one-word units, the count reverting towards a word-count. However, a word-count is already being taken and the need is for a count of significantly larger units.

For these reasons, phrases are defined as above, adding to

the traditional grammatical category the main subject and verb of each clause an an additional "phrase". The decision is, of course, arbitrary, as it must be, but as consistent results were obtained by using this approach in a pilot study (Little 1973), the approach is used in the present instance.

#### 8. "And"

As Hunt (1965) and Loban (1969) draw attention to the use of a string of "ands" as a sign of syntactic immaturity, a count was made of the incidence of "and" in the scripts, to give an index of the maximum extent to which such constructions could account for the clauses generated.

#### 9. Meaning of Word-counts

The minimal sentence is the one-word sentence, which on the present method of analysis is also a one-word phrase and a one-word clause.

One way of elaborating upon the minimum utterance is to lengthen phrases by such means as adding adjectives and adverbs and compounding nouns and verbs. This may be regarded as essentially a matter of more describing.

Another way of elaborating upon the minimum utterance is to increase the number of phrases by means of using groups of words such as preposition-plus-object and the like.

This may be regarded as essentially a matter of bringing in more relationships between phenomena as a more elaborate way of describing.

Clauses will increase in length as the product -

$$\begin{array}{r} \text{Words} \\ \text{per} \\ \text{phrase} \end{array} \times \begin{array}{r} \text{Phrases} \\ \text{per} \\ \text{clause} \end{array} = \begin{array}{r} \text{Words} \\ \text{per} \\ \text{clause} \end{array}$$

Clause length may thus be attributed to the length and/or number of phrases as the case may be.

Clauses will increase in number when minimal statements are enlarged by adding further statements in the form of a conjunction (whether "and" or another conjunction) plus a verb and (usually) its subject. This may be regarded as essentially a matter of bringing in more relationships between events. If the other forms of elaboration may be justly referred to as simpler and more complex describing, this form of elaboration may justly be referred to as a greater tendency towards explaining.

Sentences will increase in length as the product -

$$\begin{array}{r} \text{Words} \\ \text{per} \\ \text{phrase} \end{array} \times \begin{array}{r} \text{Phrases} \\ \text{per} \\ \text{clause} \end{array} \times \begin{array}{r} \text{Clauses} \\ \text{per} \\ \text{sentence} \end{array} = \begin{array}{r} \text{Words} \\ \text{per} \\ \text{sentence} \end{array}$$

Sentence length may be attributed to each contributing factor.

The minimal limit of an utterance is one word, but there is no upper limit except the limit subjects place upon their own utterances. The technique allows exploration of whether subjects tend to place various limits upon utterances in various modes of discourse.

By the study of phrase length, phrases per clause, clause length and clauses per sentence, sentence length may be analysed in terms of combinations of sub-units. For instance, two sentences may be of identical length, a fact which may be of some interest.

But it is more interesting to know whether they are also similar in other respects. One sentence might be long because of a string of ands while the other may be long because it is one grand over-arching construction using many adjectives, adverbs and compound verbs. By such means, both similarities and differences may be traced.

#### 10. Method of Making Word-counts

In more detail, the method of making word-counts was as follows.

The number of words per script was counted, "ands" being circled as this count was being made. The totals for words and "ands" were noted. Then each sentence was marked off by a vertical line at its end, using a pencil of a certain colour, with the number of sentences noted. Next, the clauses were marked off by a vertical line at its end, using another colour. The total number of lines was noted, as indicating the total number of clauses. A one-clause sentence was registered as 1 sentence and 1 clause; a two-clause sentence as 1 sentence and 2 clauses, and so on. Phrases were marked off similarly, in a third colour, and the total number of vertical lines counted to obtain the number of phrases.

The following is an example of the method, using, instead of different colours, one dividing line for marking the end of a phrase, two for a clause and three for a sentence.

A script with a completed word-count looks as follows.

Once upon a time, / there was a little boy / called Jack. ///  
He lived / in a little cottage / in a big forest / with his  
widowed mother. /// They were very poor // and often did not  
have enough food, // so one day / Jack's mother decided / to  
sell their cow ///

Words : 46  
Sentences : 3  
Clauses : 5  
Phrases : 12  
"and" : 1

In the case of "embedding" (for instance, the time-phrase one day interrupting the construction so Jack's mother decided) only one division-marker was inserted. Taken literally, this would mean that so one day would be a phrase and Jack's mother decided another. In fact the units are one day and so Jack's mother decided. No steps were taken to indicate such revised allocations of words, however, as the interest was in mean clause and phrase length, which would be the same whether the words were allocated one way or another.

For each text, the figures noted were used to calculate the following, the figures being those for the passage above.

Words per text : 46  
Words per sentence:  $46 \div 3 = 15.33$   
Words per clause :  $46 \div 9 = 5.999$ , rounded to 6.00  
Words per phrase :  $46 \div 12 = 3.83$   
Words per "and" : 46

#### 11. Reliability of Word-Counts

The rationale of word-counts was discussed at length with the assistants who undertook them, and photocopied samples of writing by one subject on the four tasks were worked through and discussed. On a "test run" involving a total of some 700

words there were only five disagreements over some 200 divisions. All cases of disagreement arose from someone missing a phrase division. The differences were discussed and resolved, and another set of photocopies was worked through, there being no disagreement on the result.

Further statistics were not calculated but periodic random checking by the present author indicated that the error or disagreement rate appeared to be very low indeed. If fatigue or inattention affected accuracy, it would be in missing divisions at the phrase level, where the demands on attention would appear to have been greatest. It is hypothesised that otherwise, errors would tend to be random.

#### 12. Use of the Method

The method involves defining "phrase" according to traditional grammar, adding the main subject of a clause and its verb as an additional "phrase" for the purpose of the investigation. "Clause" is defined traditionally and, the subjects being fairly mature writers, the sentence is taken to be what they show it to be by punctuation.

Words per phrase, per clause, per sentence and per "and" are counted for each script. By combination of scores, means for each task for all 128 subjects, and for the 64 more advanced and the 64 less advanced subjects, are determined.

Differences between means can be tested for statistical significance to determine whether, as hypothesised -

1. the tasks are in an ascending order of complexity of verbal response, as they are in an ascending order of cognitive difficulty 1. Recording, 2. Narrative, 3. Exposition, 4. Argument,
2. the subjects as a whole flexibly adapt the form of language to its function, the more advanced subjects being more flexible than the less advanced subjects.

## V. ADAPTATION OF FORM TO FUNCTION

### 1. Working Hypothesis

The main working hypothesis is that form follows function: the higher the cognitive level of the discourse the more complex the verbalisation. The cognitive levels of the four examples of discourse have been established in section III of the present study, and the complexity of the verbalisation is established on the basis of the word-counts set out in section IV. In this context, the working hypothesis implies that for the subjects as a whole, word-count-ratios will be lowest for Recording, higher for Narrative, higher still for Exposition and highest of all for Argument.

A subsidiary working hypothesis is that the more advanced subjects, as identified by school assessments, are more flexible in their use of language than the less advanced subjects. This implies that in the tasks requiring some constraint of language (Recording and Narrative) the word-counts for the more advanced will be lower than the word-counts for the less advanced. In tasks with no such constraint (Exposition and Argument) the word-counts for the more advanced will be higher than the word-counts for the less advanced.

These hypotheses are put in the form of a theoretical working model in Figure 6.

## 2. Findings

Full results for word-counts are presented in Appendix B. Findings are here set out selectively, in three stages, as follows -

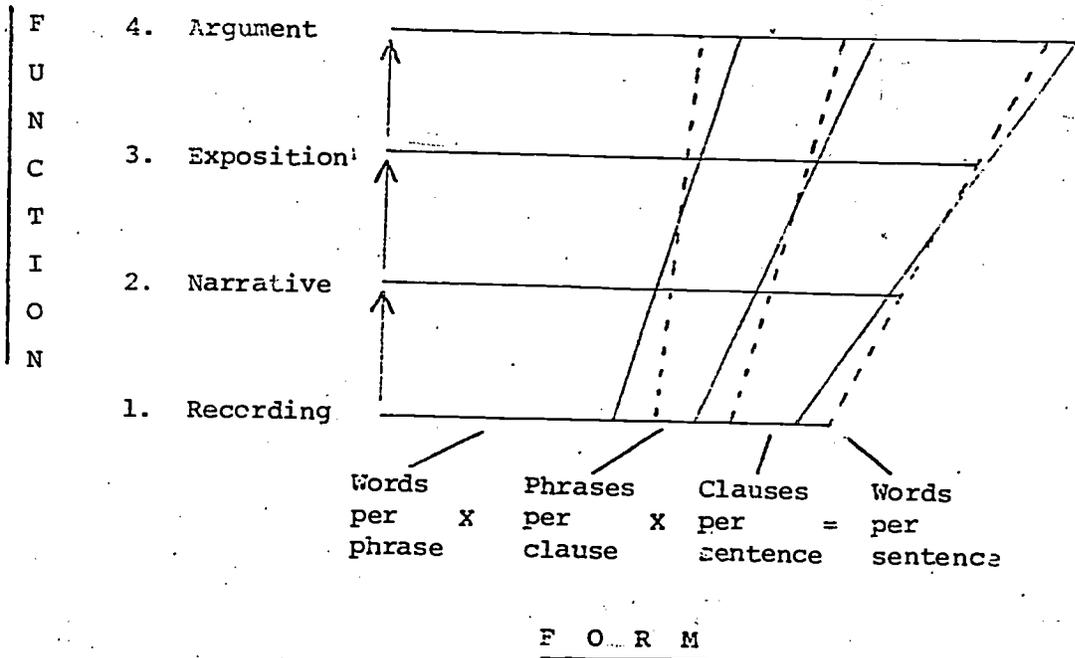
1. Theoretical and empirical models of word-counts are compared (Figure 6, p.84), indicating the extent to which the raw means of word-counts conform to the pattern predicted.
2. Indication is given (Table 5, p.85) of which means are significantly different by the t-test (Popham 1967), enabling the eventual presentation of findings to be confined to data significant in this sense.
3. Eventual findings are reported in terms of
  - (a) tabulated means (Table 6, p.88),
  - (b) ratio-formulae based on these (Table 7, p.89),
  - (c) a graphical representation (Figure 7, p.91).

This procedure is chosen as it indicates the high degree of patterning evident in the raw data as a background to the development of a more refined model based upon only those findings which are statistically significant.

Figure 6 :

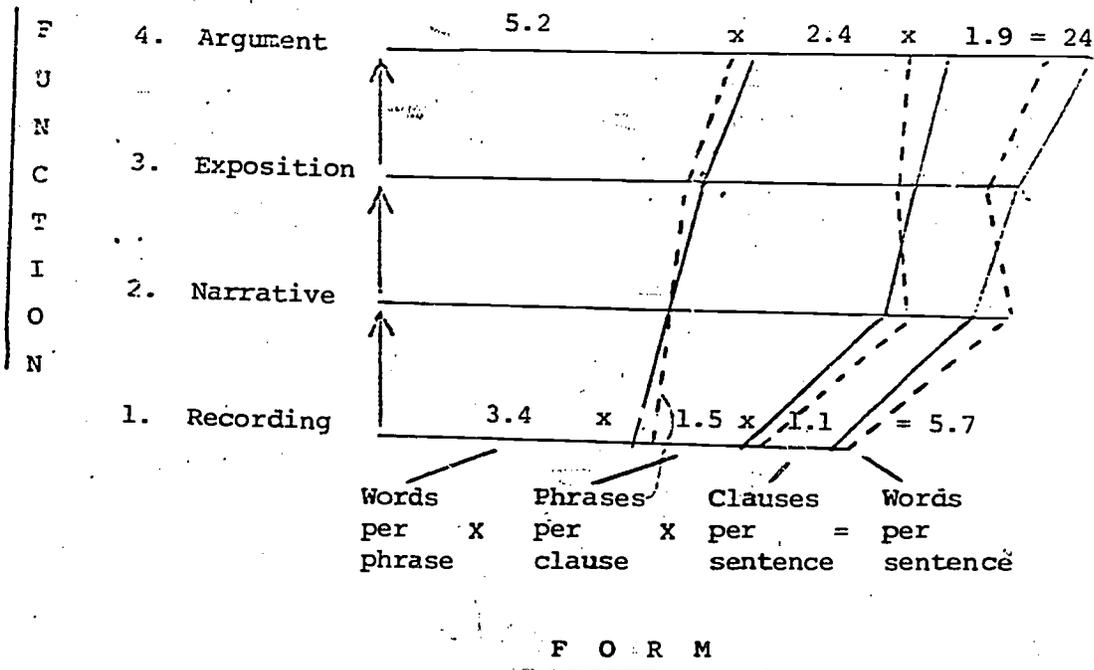
Theoretical and empirical models of adaptation of form to function by 64 more advanced (/) and 64 less advanced (✓) 16 year olds.

(a) Theoretical model.



(b) Empirical model.

Illustrative figures given for more advanced subjects in Argument and Recording.



The Empirical and Theoretical models prove to be very similar.

There is a strong tendency for the word-ratios to increase as the cognitive level rises, and for the more advanced subjects to use a wider range of structures than the less advanced: larger structures on the more difficult tasks, and smaller on the less difficult. In this sense, the major and minor working hypotheses, and beyond them, the theories underlying the enquiry, are supported by the findings. The findings so far presented are, however, raw means untested for statistical significance of differences between tasks and groups. Table 5 presents a more detailed analysis of results in terms of whether specific predictions of the working hypotheses are verified at a .05 or .01 level of confidence.

Table 5 : Positive (+), Zero (0) and Negative (-) results for working hypotheses about mean word-counts on four writing tasks for 128 subjects, dividing into 64 more advanced and 64 less advanced.

Positive results are significant differences in the predicted direction. Zero results represent no significant difference. Negative results indicate a significant difference in the direction opposite to that predicted. "+" and "-" differences are significant at the .01 level, except those annotated (.05). Negative results are significant by a two-tail test.

Subjects	Predictions for means	Words per phrase	R E S U L T S	
			Words per clause	Words per sentence
All 128	1. Recording lower than			
	2. Narrative	+	+	+
	2. Narrative lower than			
	3. Exposition	+	+	0
	3. Exposition lower than			
	4. Argument	+	-(.05)	+
	1. Recording lower than			
	3. Exposition	+	+	+
	1. Recording lower than			
	4. Argument	+	+	+
	2. Narrative lower than			
	4. Argument	+	+	+
64 more advanced (A)	A. lower for 1. Recording	0	0	0
64 less advanced	A. lower for 2. Narrative	0	0	0
	A. higher for 3. Exposition	0	+(.05)	0
	A. higher for 4. Argument	0	+	+

In Table 5 it may be seen that sixteen of the eighteen specific predictions about relative word-counts on the various tasks for the subjects as a whole are confirmed by the findings. The exceptions are -

1. that there is no significant difference in mean sentence length for 2. Narrative and 3. Exposition;
2. that clause length for Exposition is significantly larger than for Argument.

These data are taken generally to confirm the working hypothesis that the higher the cognitive level, the more complex the language, the two exceptions being noted and to be incorporated in the refined model presented below.

While differences between tasks are thus generally significant, differences between groups are generally much less significant. Only three of the twelve specific predictions about differences between mean word-counts for the more and the less advanced groups are confirmed by the findings. As shown in Figure 6 such differences are in the direction predicted, but Table 5 indicates that most of them do not rise to the .05 level of significance. There are no significant differences in the direction opposite to that predicted. The three positive findings concern clause and sentence length on the more difficult tasks. These are noted for inclusion in the refined model presented below.

The following Table culls from the findings such mean word-counts as are associated with significant differences-

1. among tasks for all 128 subjects;
2. between the 64 more advanced and the 64 less advanced subgroups.

Where there are no significant differences between the subgroups, the mean for all subjects is presented in the Table.

Table 6 : Significantly different mean word-counts for four writing tasks for 128 subjects dividing into 64 less advanced (Not A) and 64 more advanced (A) on the basis of their schools' assessments.

Task	M E A N			
	Words per phrase	Words per clause	Words per sentence	Words per "and"
1. Recording	3.6	5.2	5.8	"and" not used
2. Narrative	3.9	10.8	19.0	30 (maximum)
3. Exposition	4.2	12.3 (Not A) 11.8 (A) 12.8		50 (maximum)
4. Argument	5.1	11.6 (Not A) 10.3 (A) 12.3	23.0 (Not A) 21 (A) 24	

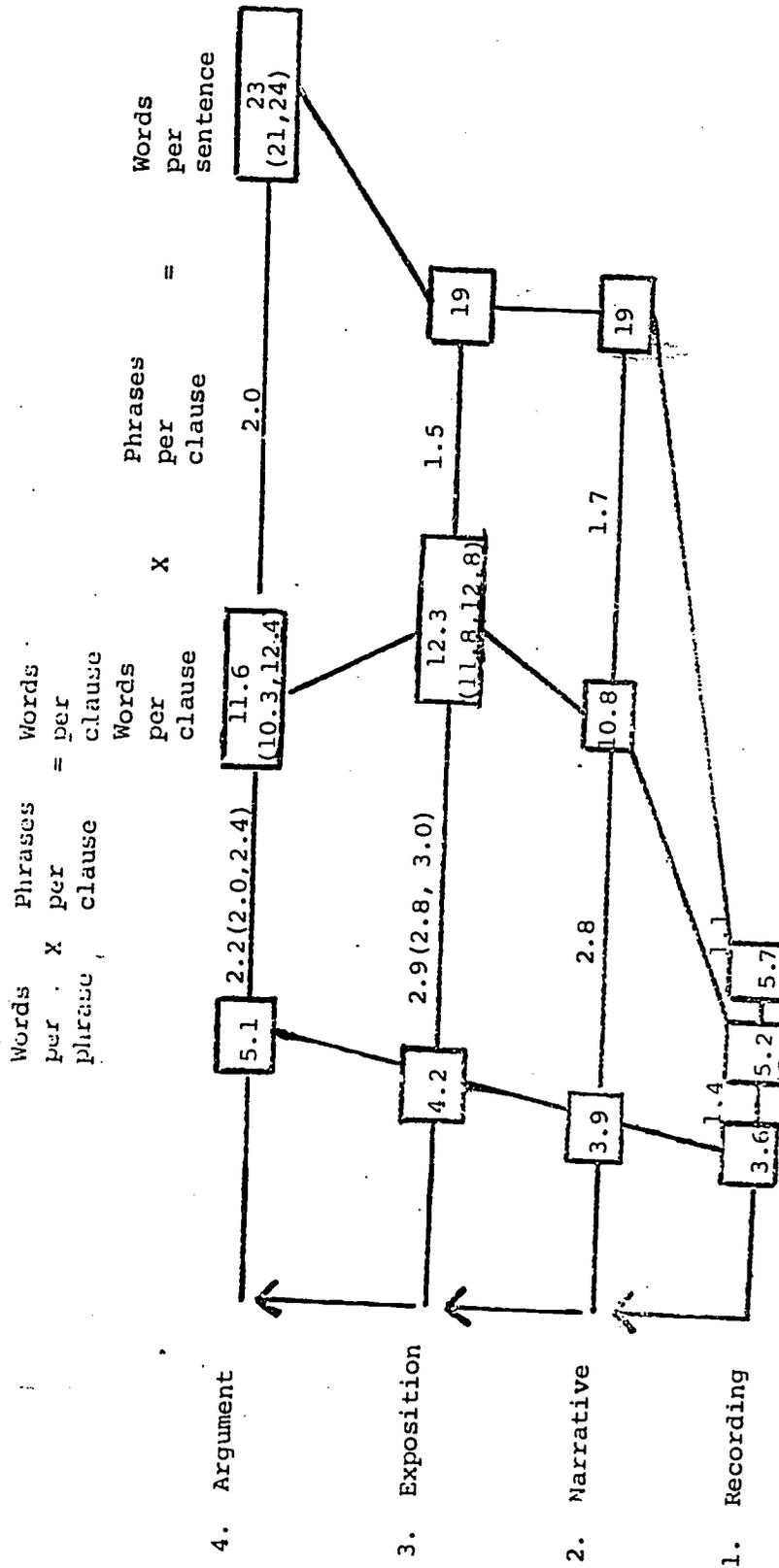
Relationships within such data are clarified by use of the ratio-formulae shown in Table 7 which is derived from Table 6.

**Table 7** : Ratio-formulae for mean word-counts on four writing tasks completed by 128 subjects, subdividing into 64 more and 64 less advanced groups when differences between groups are significant.

Task	Subjects	F O R M U L A E			
		Words per phrase X	Phrases per clause X	Words per clause =	Words per sentence =
1. Recording	128	3.6	X 1.4	= 5.2	= 5.8
				5.2	
2. Narrative	128	3.9	X 2.8	= 10.8	= 19
				10.8	
3. Exposition	128	4.2	X 2.9	= 12.3	= 19
				12.3	
	64 Not A	4.2	X 2.8	= 11.8	= 19 (-)
				11.8	
64 A	4.2	X 3.0	= 12.8	= 19 (+)	
			12.8		X 1.5
4. Argument	128	5.1	X 2.2	= 11.6	= 23
				11.6	
	64 Not A	5.1	X 2.0	= 10.3	= 21
				10.3	
64 A	5.1	X 2.4	= 12.4	= 24	
			12.4		X 2.0

For convenience, the results of Table 7 are set out graphically in Figure 7, p. 91.

Figure 7 :  
 Mean word-count ratios for 128 subjects on four writing tasks.  
 Figures in brackets are for less advanced subjects, then more advanced subjects,  
 and are given when the difference is statistically significant.



### 3. Differences between tasks for all subjects

The findings generally confirm that for the present subjects and tasks, verbal structures enlarge systematically as the cognitive level of discourse rises through the four tasks.

The only exceptions are -

1. that Exposition produces the longest clauses of all the four tasks, the prediction having been that the longest clauses would be produced in Argument;
2. that Exposition does not, as predicted, produce longer sentences than Narrative.

An explanation of these exceptions is offered below in an hypothesis about different demands placed on the phrase and the clause in different modes of discourse within a governing limit of the number of clauses per sentence.

### 4. Differences between subgroups

Tables 5, 6 and 7 and Figures 6 & 7 reveal that the significant differences between the more and the less advanced subgroups are confined to clause and sentence length on the more difficult tasks.

Table 7 reveals that these significant differences are attributable to a single factor: a greater number of phrases per clause on these tasks for the more advanced subjects. The

observed significant differences in clause and sentence length do not arise from significant differences in the number of words per phrase or the number of clauses per sentence. These ratios, though they show some differences in the predicted direction, are not significantly different for the subgroups. It is therefore concluded that the observed differences between subgroups arise essentially from the difference in number of phrases per clause.

These data also verify the hypothesis of greater flexibility of language on the part of the more advanced subjects. The differences on the simpler tasks fall short of statistical significance, but as there are significant differences on the harder tasks, the general hypothesis is upheld. The more advanced subjects thus use a significantly wider range of structures than the less advanced subjects. The extreme values from Table 7 and Figure 7 indicate that the less advanced subjects vary relevant structures some 70% to 80% as much as the more advanced subjects.

##### 5. Differences between tasks and groups combined

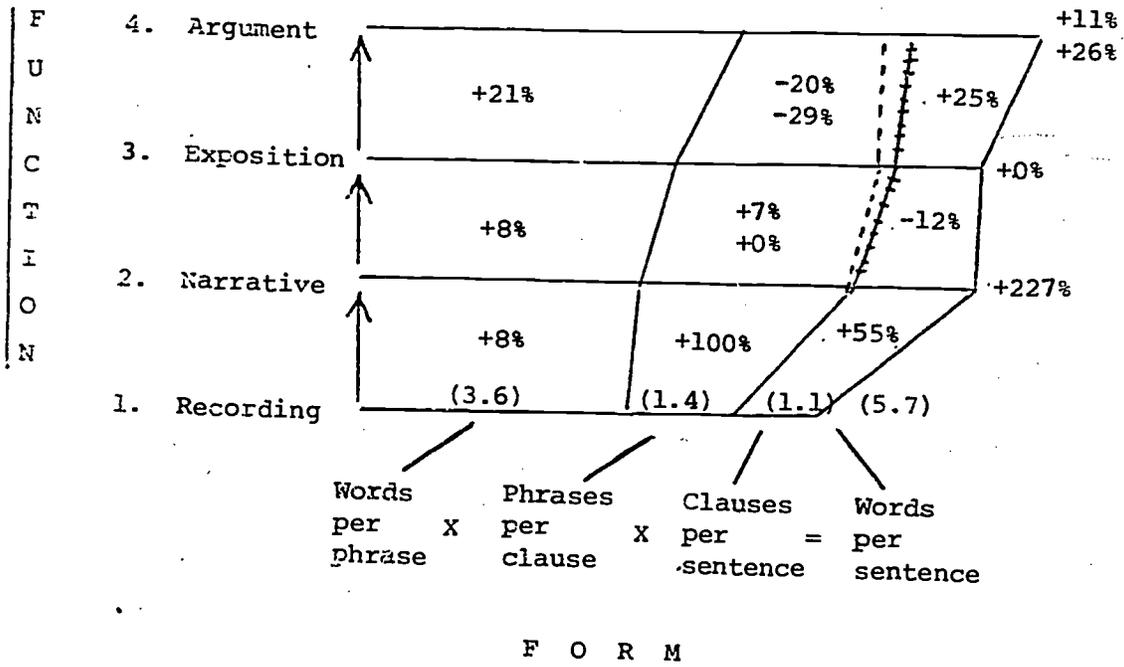
Figure 8 (p.94) shows findings in terms of percentage of change in verbal structures as the groups move through four modes of discourse in ascending order of cognitive level. Again where there is no significant difference between subgroups, the mean for all subjects on the task is used. Where there is a significant difference between subgroups, the different means are used.

Figure 8 :  
Model of adaptation of form to function by 128 16 year olds,  
in terms of percentage increase or decrease of word ratios  
in movement from task to task.

- / all subjects.
- less advanced subjects.
- × more advanced subjects.

When two figures are given for the same change, the first is for the less advanced subjects and the second for the more advanced.

The figures in brackets are the mean ratios for Recording for all subjects.



Results are now discussed task by task.

a) Recording

The formula for Recording is as follows -

$$\begin{array}{ccccccc} 3.6 & & 1.4 & & 1.1 & & 5.8 \\ \text{words} & & \text{phrases} & & \text{clauses} & & \text{words} \\ \text{per} & \times & \text{per} & \times & \text{per} & = & \text{per} \\ \text{phrase} & & \text{clause} & & \text{sentence} & & \text{sentence} \end{array}$$

There are 5.2 words per clause. "And" is not used at all.

In Task 1, Recording, all structures are minimal for both the more and the less advanced groups, which do not differ significantly in this respect.

There are minimal though significant differences in quality scores for the two groups (Table 2 ) but these are not associated with any significant structural difference evident in the findings. There is a hint in the data that structures are more economical for the abler group, but this tendency does not rise to the .05 level of significance. Evidently, differences in quality are to be attributed to differences other than the structural features of the scripts; perhaps, semantic features.

Structurally, the Recording task shows compact phrases of a mean of 3.6 words, compact clauses of 5.2 words and a predominance of one-clause sentences resulting in a mean sentence length of only 5.7 words. There is a low but fairly "even" production of phrases per clause (1.4) and clauses per sentence (1.1) when compared to the other tasks.

It would appear that the subjects as a whole share a competency in selecting a few pertinent facts from presented data and in recording these in telegram form. Difference in quality of such writing between the subgroups are not attributable, however, to structural differences, because these are not significantly different for the subgroups. Telegram form is obviously restrictive, but other Recording also seems to be restricted in form (see p.107).

b) Narrative

The formula for Narrative is as follows --

3.9		2.8		1.7		19
words		phrases		clauses		words
per	x	per	x	per	x	per
phrase		clause		sentence		sentence

There are 10.8 words per clause. "And" is used no more than once in about every 30 words.

In movement from Task 1, Recording, to Task 2, Narrative, the subjects increase all structures. The phrases increase by 8%, the number of phrases per clause by 100%, the clauses increase over 100% and the number of clauses per sentence increases by 55%. As a product of these increases, sentence length increases by almost 230%.

In this movement from Recording to Narrative, emphasis shifts from a low but relatively "balanced" production of phrases per clause and clauses per sentence, to a much greater production of phrases per clause (from 1.4 to 2.8) and a lesser increase

in production of clauses per sentence (from 1.1 to 1.7).

The word "and" is used no more than once in every 30 words, signifying that a maximum of one clause in three might begin with that conjunction.

It thus appears that in Narrative as produced by the subjects, there are slightly longer phrases than in Recording, signifying changes such as arise from using more adjectives and adverbs or compounding nouns and verbs. There are many more phrases per clause, indicating more use of prepositions, non-finite verbs, and other phrase-starters such as may be used for time-phrases and absolute phrases. This involves more indications of relationships between phenomena; relationships such as are indicated by expressions such as later on, in the forest, singing merrily, to chop some wood. Because of this increase in the generation of phrases, clauses are much longer. As a result of these factors and the increase in the number of clauses, sentences are much longer, reaching a mean<sup>o</sup> of 19 words. In this mode of discourse, one in every two or three clauses would be a main principal clause, fewer than one in three clauses would begin with "and" and one in three or more clauses would begin with a conjunction other than "and".

As there is no significant difference in language structure between the more and the less advanced subgroups, the significant difference in quality scores between the groups, which is greater than for Recording, must be attributable to some other factor, such as semantic differences. This is stated in the context of differences such that the more advanced write more

simply than the less advanced, but the differences do not rise to the required level of statistical significance.

Ability to move from the compressed language required in the Recording task to the more descriptive and explanatory language required in the Narrative task is evidently an ability shared by the subjects as a whole.

c) Exposition

The formula for Exposition is as follows -

i) All subjects :

$$\begin{array}{ccccccc}
 4.2 & & 2.9 & & 1.5 & & 19 \\
 \text{words} & & \text{phrases} & & \text{clauses} & & \text{words} \\
 \text{per} & \times & \text{per} & \times & \text{per} & = & \text{per} \\
 \text{phrase} & & \text{clause} & & \text{sentence} & & \text{sentence}
 \end{array}$$

There are 12.3 words per clause "And" is used no more than once in about every 50 words.

ii) More advanced subjects :

As above, except that there are 3.0 phrases per clause and 12.8 words per clause.

iii) Less advanced subjects :

As above, except that there are 2.8 phrases per clause and 10.3 words per clause.

When the subjects move from Task 2 Narrative to Task 3, Exposition, sentence length does not change significantly but significant changes occur within the sentences.

Of all four modes of discourse, Exposition appears to generate most activity at the phrase level. In Exposition, phrase length increases by 8% compared to Narrative, and 17% compared to Recording. The number of phrases per clause is greatest for all four writing tasks : 3.0 for the more advanced and 2.8 for the less advanced. Hence clauses are the longest for this task compared to all others. The difference between groups in the number of phrases per clause is significant, and accounts for the difference between groups for clause length on this task.

Compared with Recording and Narrative, Exposition involves more elaboration and multiplication of phrases. This implies that within the phrase there are more modifiers and/or more compound nouns and verbs. It implies that within the clause there are more phrases indicating spatial, temporal and other relationships. There is thus more use of descriptive words and phrases.

In Exposition, clauses decline in number while phrases increase. There are fewer clauses per sentence than in Narrative (1.5 as compared to 1.7), and the difference is statistically significant.

In Exposition, about half the sentences would be one-clause sentences, and half two-clause sentences. Some two out of three clauses would be main principal clauses. If "and" were used maximally to begin clauses, the remainder of the clauses would be fairly evenly divided between clauses beginning with "and" and clauses beginning with other conjunctions. If "and" were not maximally to begin clauses, clauses beginning with conjunctions other than "and" would somewhat predominate the remainder of the clauses.

The fact that the more advanced subjects produce more phrases per clause than the less advanced is associated with a significant difference in favour of the more advanced in the quality scores for Exposition. While such association does not necessarily prove a causal connection, the hypothesis is tenable that a greater ability to explicate relationships by marshalling more phrases into the clause is a reason for the superiority of the Exposition of the more advanced subjects.

In summary, it may be said that in Exposition as Task 3 on a four-task rising scale of cognitive challenge, phrase length rises and phrase generation maximises. The generation of clauses, however, falls to a level below that of Task 2, Narrative. Evidently, in Exposition, the phrase rather than the clause is where the action is, and greater generation of phrases appears to be related to better-quality Exposition.

d) Argument

The formula for Argument is as follows -

i) All subjects :

$$\begin{array}{rcccccc}
 5.1 & & 2.2 & & 2.0 & & 23 \\
 \text{words} & & \text{phrases} & & \text{clauses} & & \text{words} \\
 \text{per} & \times & \text{per} & \times & \text{per} & = & \text{per} \\
 \text{phrase} & & \text{clause} & & \text{sentence} & & \text{sentence}
 \end{array}$$

There are 11.6 words per clause. "And" is used no more than once in 50 words, as for Exposition.

ii) More advanced subjects :

As above, except that there are 2.4 phrases per clause, 12.4 words per clause, and 24 words per sentence.

iii) Less advanced subjects :

As above, except that there are 2.0 phrases per clause.

10.3 words per clause and 21 words per sentence.

Argument produces significantly longer sentences than any other task. This is not a product of an increase in all structures, but is another example of shift of emphasis from one kind of structure to another.

Argument produces by far the longest phrases (5.1 words per phrase) and the most clauses per sentence (2.0). But compared to both Narrative and Exposition there is a toning-down of the number of phrases per clause (2.8 for Narrative, 2.9 for Exposition but only 2.2 for Argument). There is a significant difference in Argument between the two subgroups, however. The less advanced reduce this factor to 2.0 but the more advanced reduce it only to 2.4. This difference is associated with better quality scores for the more advanced subjects.

Argument thus makes the greatest demands upon phrase length; moderate to low demands on phrase generation (with a lesser reduction in this factor by the more advanced); and the greatest demand on clause length.

It was noted above that Recording places a low but relatively even demand upon the production of clauses and phrases.

Narrative greatly increases the demand for phrases and slightly

increases the demand for clauses. Exposition makes maximal demand for phrases and reduces the demand for clauses. With Argument, the balance is redressed, so that demand for clauses rises and demand for phrases falls to produce a more "even" balance (2.2 phrases per clause and 2.0 clauses per sentence).

There is thus a reciprocity between phrase and clause generation in a movement from Exposition to Argument: more phrases but fewer clauses in Exposition, and fewer phrases but more clauses in Argument. In each case, greater production of phrases on the part of the more advanced subjects appears to be the sole structural difference between the more and the less advanced subjects, and is an apparent cause of longer clauses and better quality exposition and argument.

e) A possible limit to increases in word ratios.

The initial working hypothesis places no limit upon the increasing of word-ratios as the cognitive level of the discourse rises. This would seem to imply, "the more words the better", which is evidently not the case either in general, or in the sampled writing. The better scripts do not multiply words endlessly, but conform to a distinct mean and standard deviation for each task.

If there is some limit to the production of words, as there seems to be, it is not the number of words per sentence, which varies systematically. A possible governing factor is, however, the relationship of the number of words per phrase to the number of phrases per sentence. These factors are not constants, but vary in similar ways. The number of words per phrase rises regularly

as the cognitive level of the discourse rises, and so does the number of words per sentence, with the exception that it is virtually the same for Narrative and Exposition. The ratio of the number of words per phrase to the number of phrases per sentence proves to be between 4.5 and 5.0 for Narrative, Exposition and Argument, being lower because of the compressed telegram form in Recording. The lower limit being a blank page, or perhaps one word, the upper limit would seem to be five phrases per sentence.

It thus appears that subjects end sentences after a maximum of 4.5 to 5.0 phrases. If there are many phrases per clause in the particular mode of discourse, then there are not many clauses per sentence. If there are fewer phrases per clause, there may be more clauses per sentence. Such a "governing factor" is incorporated into the refined model which follows.

#### 6. Refined model of adaptation of form to function.

Empirical findings lead to the following revised forms of the major and subsidiary working hypotheses with which the present analysis of the relationship of form to function began. The refined model makes the most economical number of assumptions compatible with fitting the data, and is presented less as a conclusion than as an hypothesis for further investigation.

1. The number of words per phrase, phrases per clause, clauses per sentence, words per clause and words per sentence rises as the cognitive level of the discourse rises through the tasks

1. Recording, 2. Narrative, 3. Exposition and 4. Argument.

- a) within the limit of five phrases per sentence;
- b) with the exceptions that
  - i) Exposition calls forth most phrases per clause, and
  - ii) Argument calls forth most clauses per sentence.

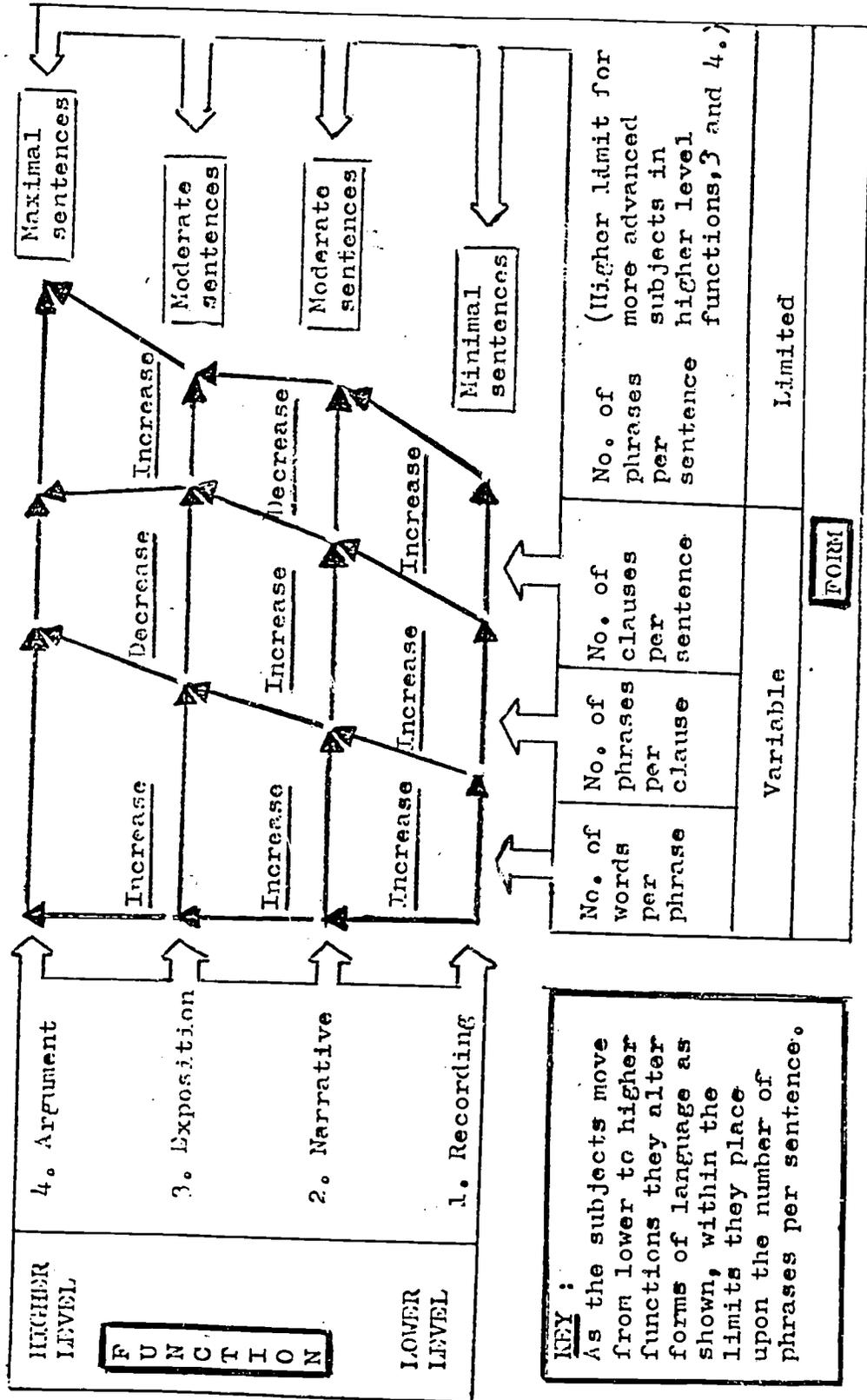
2. More advanced subjects produce more phrases per clause than less advanced subjects on the tasks Exposition and Argument, and this difference is associated with a difference of quality of the resultant scripts, the more advanced being superior.

Deductions from these hypotheses would logically account for the following -

- 1. Minimal structures throughout in Recording.
- 2. An increase in all structures in Narrative.
- 3. An increase, in Exposition, of phrase length, phrases per clause and clause length, but a decrease in clauses per sentence. The balance of increases and decreases would not exactly predict, but suggest the result that sentence length is, as a product, no longer than for Narrative.
- 4. An increase, in Argument, of phrase length and number of clauses, but a decrease in number of phrases. The increases and decreases would be compatible with, though not exactly predict, a decrease in clause length and an increase in sentence length when compared with Exposition.
- 5. More phrases associated with better Exposition and Argument on the part of the more advanced.

More precise predictions would require numerical values to be given to the parameters. Here the concern is with relativities, in order to develop the following generalised model.

Figure 9: Model of adaptation of linguistic form to mode of discourse (function) based on the writing of 128 sixteen year olds.



This model represents the conclusions of this part of the study: that the higher the cognitive level of the discourse the more complex the verbalisation of the types and within the limits shown; and that the more advanced subjects are more flexible in adapting form to function, essentially because of a greater ability to generate phrases in the higher levels of discourse.

### 7. Illustrative Scripts

The following are scripts which conform closely to the means for each task, in terms of both quality as evaluated by the two markers, and mean word-counts for the particular mode of discourse. The scripts are transcribed with original spelling and punctuation, noting that they were written under test conditions with little or no time for revision.

#### (a) Recording

The following telegrams approximate to the mean values of 3.6 words per phrase, 5.2 words per clause and 5.8 words per sentence. They are typical of mediocre examples in conveying the essential information, but in rather more words than necessary. Inferior examples tended to convey even less information and better examples the same information in fewer words.

- (i) Arriving at Essendon Airport 7.30 p.m. on Sunday. Sorry for the delay but Dad had a little accident and I was held up.
- (ii) Unable to make destination as planned. Father had little accident, nothing to worry about. Arriving Essendon Airport 7.30 p.m. Sunday.

(b) Narrative

The examples are (i) somewhat above average, and (ii) average in quality. They approximate the mean word-counts for Narrative of 3.8 words per phrase, 10.8 per clause and 19 per sentence.

- (i) Once upon a time, there lived a mean old witch. Her home was a burrow between the twisted roots of a huge pine tree. Although this witch had a wicked smile which showed her large pointed teeth and caused her huge purple nose to wobble, she had a very lonely life. A long time ago, the forest was full which many folk from a town nearby loved to have picnics in. Then a horrible little Goblin in a bright yellow suite with green shoes cast a bad curse over the whole forest and turned the prettiest girl in the villiage into a witch.

The forest under the curse became a place of darkness. There were no more flowers with pretty green stalks and golden petals and no one ever laughed any more. The poor witch lived by herself but wasn't really a bad witch.

- (ii) There once was a cow who lived on a farm his name was "Spiro".

He was very happy there until one day a truck run him over. The owner saw no use for this lame and distorted cow but his son loved it passionately.

This caused much confusion in the Vamvoukalis household. With his father seeing no financial value in the cow "Spiro" but the son praising him. No doubt bitterness followed, and with some intellegence one could decide on the route the father would take.

(c) Exposition

The following scripts are (i) slightly better, and (ii) slightly worse than average in quality, and approximate to the mean word-counts of 4.1 words per phrase, 12.3 words per clause and 19 words per sentence. The proliferation of phrases characteristic of Exposition is well evident.

(i) Ludo

Each player is supplied with four counters and they are placed in a starting box. One dice is used and when a player throws a six they get another turn. A player needs to get a six to leave their starting box and once out must travel by dice throw round the board to reach their home which is near where they started. If a player's counter is landed on by another player's counter the player who was landed on must return to their starting box. If a player has two counters out and they get a six and another number they may move the six and other number separately to try to send another player back to his starting box. The game continues in this manner until one player has all his counters home.

- (ii) The game that I know is called soccer. It is played with a ball and two teams consisting of eleven players. The aim of the team is to use tactics to try to score in the opponents goal. In this game the whole body of the player could (indecipherable) except for his hands. The goal keeper has the only privilege to use his hands.

In playing the game the opponents have to take the ball of each other. A player is restricted to push with his hands or savagely attack the opponent. He is allowed to push slightly with his shoulder. Another rule is that if a player except the goalkeeper touches the ball with his hands the other team will be awarded what is known as a free-kick or possession of the ball. When this is done in the goal keepers area a penalty is awarded to the team. The team is given a free shot at the goal with the goalkeeper only guarding the goal. All other players except the player taking the penalty must be out of the goal's area.

(d) Argument

The following examples are (i) slightly above average and (ii) average in quality. They approximate to the means of 5.1 words per phrase, 12.4 words per clause and 23 words per sentence. The topic is the points-system for Australian television content. The proliferation of clauses characteristic of Argument is well evident.

- (i) I believe this is undesirable because many Australian productions do not meet the liking of the public. While some Australian productions are quite good, others are not of good standing or quality.

In many households people watch productions from other countries because they believe channels would not waste money on shows people are not going to watch, so shows from other countries must be of some quality.

If they kept this "points system" I hope they bring in good quality Australian productions. One Australian show, Number 96, supposedly a good quality show is one, which I do not think particularly good at all. If the "points system" causes more shows to come on television like this I can not see myself watching television so much as I do now.

- (ii) The system is a good move as it will represent a step forward for the Australian film industry. Australian people quite often complain about the poor quality of Australian productions, without realising that without showing these "poor quality" shows Australia has no way to improve upon them. It is said that we learn from our mistakes and this is especially true of the film industry. If a big production is not a good seller there is a serious financial problem as the cost of film making is exorbitant.

## VI. APPLICATIONS AND IMPLICATIONS

It appears that the working hypotheses derived from the Moffett and Britton models of modes of discourse are broadly upheld for the sixteen year old subjects on the four writing tasks. Before considering implications, it would seem appropriate to consider applications of the model to other tasks and subjects.

### 1. Other tasks

Data are available for the 128 subjects of the present study for two other writing tasks. Though acquired in the course of other pursuits, these data provide opportunity for the further verification of hypotheses about the adaptation of form to function in the writing of the subjects.

The subjects completed a task known as the "Aluminium Passage". This is a rewriting exercise devised by Hunt and O'Donnell (Hunt 1970) and slightly adapted for Australian conditions by Johnston (1973). The exercise is intended to provide standardised data on the T-unit and sentence length. The task presents a string of minimal statements, each in a separate sentence, and subjects are asked to rewrite the passage in more normal, flowing prose. On the present analysis of modes of discourse, this task is hypothesised to be an example of Recording. Although the product resembles Exposition, in the setting out of facts about a subject, the process is one of selection and arrangement of presented data rather than its generation, such as in Exposition of the rules of a game. As an example of

Recording, the Aluminium Passage usefully avoids the restriction of telegram form imposed in the Recording tasks deliberately used in the body of the present study.

It is hypothesised that for the sixteen year old subjects such a Recording task would prove minimally discriminating and produce simpler verbal structures than Narrative, Exposition and Argument.

On the criterion of differentiation between the more and the less advanced subjects, using the same judges and procedures as for the other tasks, the Aluminium Passage proved to be more discriminating than the telegram-type of Recording task, but less discriminating than the other tasks, as hypothesised. Also as hypothesised, its word-count ratios proved to fall between those for Recording on the telegram task and for Narrative. The values were 3.8 words per phrase, 2.3 phrases per clause, 3.6 words per clause, 1.6 clauses per sentence and 13.3 words per sentence. A small random sample (16 scripts) was taken to determine the mean T-unit, which the task was designed to measure. The mean T-unit proved to be 9.9 words, with a range from 6.5 to 13.3 words as the mean T-unit per script. The similarly calculated T-unit for the Argument task as completed by the same subjects proved to be 16.3 words, with a range from 8.0 to 22.4 words. The mean T-unit for the Argument (16.3 words) is thus higher than the highest extreme for Recording (13.3 words). While no firm

conclusions are drawn from these small samples, they do suggest that the Aluminium Passage may be an example of Recording, and that for the same subjects, a different mode of discourse such as Argument may produce much higher levels of verbal complexity. The sampling of language by the Aluminium Passage would need to be regarded with some caution in the sense that while results may be valid for the particular form of Recording, they may not be valid for other modes of discourse.

Data are also available for response by the same 128 subjects to the picture of silhouetted figures dancing (see Appendix A). This task is hypothesised to be an example of Exposition, with some admixture of Narrative. The task involves more than Recording, as the data is not merely presented for selection and arrangement, but has to a considerable extent to be generated. Subjects are required to elaborate upon the presented stimulus by inventing some account of what the dancing figures are doing, thinking and feeling. Although the requirements of the task could to a considerable extent be achieved by Narrative, the task would seem to be one of Exposition, as the shaping of the response as Narrative is not a requirement.

In accord with the hypotheses derived from the Moffett and Britton models, it would be expected that the task would discriminate between subjects and produce word-

count ratios at a level close to that of Exposition (rules of a game). The procedures of evaluation were those of the English Reference Test, which while broadly comparable, are not exactly comparable to those of evaluation by the two judges in the main part of the present study. Even so, the task discriminated between the more and the less advanced subjects at the .01 level of significance, which at least does not contradict the hypothesis about discrimination level. The mean word-counts are directly comparable, however, and proved to be 3.8 words per phrase, 3.1 phrases per clause, 11.7 words per clause, 1.6 clauses per sentence and 18.5 words per sentence. These values are close to those established for Exposition on the rules-of-a-game task. The value of 3.1 phrases per clause is larger than but close to (not significantly different from) that of 2.9 for Exposition of rules of a game. Otherwise, the values are slightly lower than for Exposition, falling between those established for Exposition (rules of a game) and Narrative (bedtime story).

It would thus appear that on writing tasks other than the main four under study, the present subjects respond consistently in adapting form to function. In an additional Recording task without the restrictions of the telegram form, the discrimination index and the word-counts are between those already established for Recording in telegram form and for Narrative. On an additional Exposition task capable of some Narrative treatment, the

word-count ratios generally fall between those already established for Exposition and Narrative. On this task, the discrimination index, though not strictly comparable, is high, as would be predicted from the hypotheses.

These data suggest a consistent pattern of change in discrimination-indices and word-count ratios on the part of the same subjects moving between six different writing tasks. It would appear that tasks could, in further investigations, be classified by mode of discourse so as to lead to verifiable predictions about discrimination indices and phrase, clause and sentence forms for subjects at a given level of verbal development.

## 2. Other Subjects

The methods are applicable to other age-groups. As noted on p.63, Davis (1975) applied some of the techniques of the present study to Narrative and Exposition by eight year olds and twelve year olds. Using inter-marker reliability as a discrimination index, Narrative was found to be the better discriminator among eight year olds and Exposition among twelve year olds, as the developmental model would predict. Patterns of phrase, clause and sentence length were similar to those for the present sixteen year old subjects, though values were slightly lower for the twelve year olds and lower still for the eight year olds. It was noted that under pressure to write Exposition, the eight year olds tended to contract rather than expand their language.

These preliminary findings are reported only to establish that the method is applicable to younger subjects in attempted further verification of the developmental model.

The model is similarly applicable to older subjects. Godshalk et.al. (1966) present discrimination and validity indices for five writing tasks completed by some 250 seventeen and eighteen year old subjects and evaluated by five markers. The results are tabulated on pages 50 to 84 of the study. Predictably, in terms of the present model, the reliability indices are relatively low for a relatively simple recording task (describing a notable feature of the home town to a pen-pal). Also predictably, the highest discrimination indices tend to be those for a task in argument (analysing character and evaluating argument in a student speech). Two more straightforward expository tasks fall between the extreme discrimination indices associated with the above tasks. The exception would appear to be a narrative task, which is more discriminating for such subjects than might be predicted on the present model. The nature of this task is, however, highly "theoretical", in that a story has to be built around a common object. If this requirement be interpreted as requiring a demonstration of the significance of the object in people's lives, as would appear to be the case, the task is one of Argument-Narrative rather than Narrative of a more straightforward type, and the high discrimination would be accounted for. Correlations with scores on the

"ASAT" test as a criterion measure are in accord with the present model, but those with "SAT" as the criterion measure are not. Such data at least suggests that the hypotheses and techniques of the present study may be applicable to older subjects and a wider range of tasks.

They are also applicable to mature published writing. Small samples were taken of published writing in the Narrative, Exposition and Argument modes. The relevant formulae for the published writers were as follows.

Table 8 : Mean sentence formulae for small samples of published writing.

Mode of Discourse	Words per Phrase	X	Phrases per Clause	X	Clauses per Sentence	=	Words per Sentence
Narrative	3.8		2.6		1.5		15.5
Exposition	3.8		6.8		1.2		31
Argument	3.8		3.6		1.8		25

The sample is no doubt too small for any definite conclusions to be based upon it but the differences are noted to be broadly similar to those between tasks for the sixteen year olds, with a very similar structure for Narrative, a much more spectacular generation of phrases at the expense of clauses on the part of the published Exposition, and a return to a somewhat more balanced production of phrases and clauses in Argument. The suggestion is that the changes the sixteen year olds make in moving from task to task are less spectacular versions of the differences observable in published writing in the various modes.

It is noted that the published writers consistently use a mean phrase length of 3.8 words, and that if the ratio of this value to sentence length gives an index of the limit to which word-units are multiplied, the highest value is 8, compared to 5 for the sixteen year olds. These data are regarded not as definitive findings or conclusions but as indications that the method is applicable to writing generally, and suggests hypotheses for further investigation.

### 3. Other modes of discourse

In a pilot study, Little (1973) applied some of the present techniques to the spoken language of eight year olds. The subjects used minimal structures in a Recording task (describing a picture) and more elaborate structures in an Exposition task (describing the rules of a game). The techniques are evidently also applicable to the spoken tongue, and could enable comparisons between spoken and written English on the part of the same subjects on similar and different tasks.

Only four kinds of discourse were selected for study in the present instance, and under test rather than field conditions. The techniques could, however, be applied to further tasks to cover the whole Britton taxonomy, and to field data such as that indicated by Martin (1965).

As material that is written is material that can be read, the present techniques are applicable to "Readability" (Gilliand 1972). Sentence length is used as a major element in Readability formulae, and on the basis of data from Gilliland (1972 : 100), the present subjects write sentences associated with readability at anything from below Grade II to Grade XII level. The suggestion is that the making of distinctions between modes of discourse, and exploration of indices at phrase and clause level might lead to more discriminating indices of readability. The "listenability" of spoken material could also be investigated in such ways.

In summary, the techniques would appear to be applicable in studies of listening, speaking, reading and writing at the various age levels. It is also possible that the approach to age levels in terms of discrimination indices might also be applicable to non-verbal developmental tasks.

#### 4. Implications

Apart from developing techniques for use in further enquiries, the present study confirms the Moffett-Britton type of model of development in modes of discourse, and relates structural features of language to the function and cognitive level of the discourse.

The educational implications of this type of model have already been worked out fairly elaborately at the functional level. Moffett (1968) bases a whole Kindergarten-to-College language

arts curriculum upon this model, and Britton (1972), Barnes (1969), Martin (1975), Rosen (1973) and Smith (1972) are among well-known proponents of the model. The model is very much part of the Bullock Report (H.M.S.O. 1975). In Australia, the first curricula to be explicitly formulated in terms of this type of model are the Primary and Secondary English Curriculum Workshop papers of the A.C.T. Interim Education Authority (1964, 1965), although other systems and syllabuses (e.g. N.S.W. Secondary Schools Board, 1972) have made use of the Moffett and Britton types of models. More specific implications are as follows.

a) Cognitive aspects

The main educational implication is that of developmental sequence. If a child is functioning verbally at the level of Recording, teachers need not be surprised if the child is not functioning so well in Narrative or Exposition or Argument. The next developmental step for such a child would normally be in the direction of handling Narrative. The handling of Narrative would be a condition of developing ability to handle Exposition, and Exposition would be a condition of developing ability to handle Argument. While various children will move through these stages at various rates, the sequence appears to be universal. Broadly speaking, the Recording and Narrative stages belong to early childhood, though this does not mean that no preliminary developments towards Exposition and Argument are taking place. Exposition is developing more in the later primary and early secondary years. Argument is developing, apparently earlier for

physical-type issues than humanities-type issues, in the later years of secondary schooling. These broad generalisations need much qualification and further verification, but the outline appears to be emerging fairly clearly. It is emphasised that the "earlier" forms of communication such as Recording and Narrative do not fall out of use as other modes develop, and are not incapable of further development in the later years. People use and need all manner of modes of discourse, and all are deserving of cultivation.

The observations of Martin (1975), Newsome (1975), Cambourne (1971) and others (Dunkin and Biddle 1974 : chapter 9 ), to the effect that the overwhelming bulk of school talk is teacher talk in the expository mode, and that the overwhelming bulk of school writing in the expository mode to the teacher in the role of examiner, suggest that the school may typically be a narrowing rather than a broadening language environment, giving little scope to Recording and Narrative on the one hand, or Argument on the other. When it is apparent (Barnes 1969 ) that a good deal of concept learning involves Recording, Narrative and Argument (the latter perhaps brief and fumbling), it would seem that the best use of modes of discourse may not be made in typical classroom practice. Indeed if the school is so set on stopping all modes of discourse except the expository, the interpretation could be put on the situation that unconsciously, teachers are trying to turn the children into replicas of themselves in a narrow teacher-role rather than trying to help develop language competency in all its fulness and range.

Diagnostic observation of pupils and curriculum planning to meet their developmental needs involves observation of their verbal behaviour over a range of modes of discourse. Inadequate sampling of modes can only be misleading when the developmental patterns for the various modes are different. Little idea will be gained of the verbal competencies of small children by sampling their ill-developed powers of argument, as little idea of the verbal competencies of graduates will be obtained by asking them to compose a telegram. The different discriminative powers of different modes of discourse at different stages of development thus need to be kept in mind.

With such a model in mind, curriculum planning has a sounder basis for the selection of reading and listening materials and the promotion of activities in speaking, writing - and thinking - than an approach which failed to make such discriminations. In view of the apparent lateness of development of ability to handle formal reasoning about humanities-type problems, upper secondary and undergraduate curricula in these areas might deserve closer scrutiny for the appropriateness of the language demands made.

The implications are not confined to the subject of English. All subjects use language, often with the greatest weight being placed on the modes of exposition and argumentation. This is good reason for further critical enquiry into the demands placed upon language "across the curriculum" (Barnes. 1969 ) at all stages of development.

b) Structural aspects

The distinction between cognitive and verbal aspects of performance on the tasks is merely one of convenience in distinguishing between the investigation into discrimination indices (cognitive) and word-counts (verbal). In reality, the verbal may be part of of the cognitive, the two having a relationship such as that between strategy and tactics.

The hypothesis is put forward that the differences between modes of discourse are far more cognitive than they are verbal per se.

Apart from one-word utterances, normal sentences involve a kernel statement about an agent and an action: in a word, an event (John laughed). Such a kernel statement may be expanded by adding phrases, which set out relationships between phenomena connected with the main event (John laughed at the joke).

Phrases of both types may be expanded in length by more description (John the Welshman laughed heartily at the ridiculous joke). One-clause sentences may be expanded by indicating the relationships of one set of events with another. (... laughed heartily at the ridiculous joke which Bill told him).

Recording involves all of these four functions - making kernel statements, describing and showing relationships between phenomena and between events - and employs words, phrases and clauses within sentences to do so. There is nothing structural in Narrative that is not in Recording, and so on for the other

modes of discourse. The basic structures are present in the simplest mode, and it is only their combinations which change as the mode of discourse changes, together with its cognitive level.

There are, of course, certain verbal structures particularly associated with the various forms of discourse. Narrative may use ritual expressions such as once upon a time, for instance. Exposition omits the subject in instructions (Beat 6 eggs ...) and often enumerates points. Argument is assisted by connectives which show the relationships between whole sentences, such as however. Yet these do not constitute each mode, which fundamentally uses the same verbal structures as every other, but in different combinations and with different emphases,

This is so for the sixteen year old subjects, and is hypothesised to be so developmentally. Once small children have reached the stage of Recording they show that they have in repertoire the basic phrase and clause structures. What changes as they acquire higher forms of discourse is the extent to which they combine them.

In this sense the structural differences between the tasks are really cognitive. In recording, describing, relating phenomena and relating events are minimal. In narrative, there is greater demand for all. In exposition, there is more pressure on describing, much more on relationships between phenomena and less

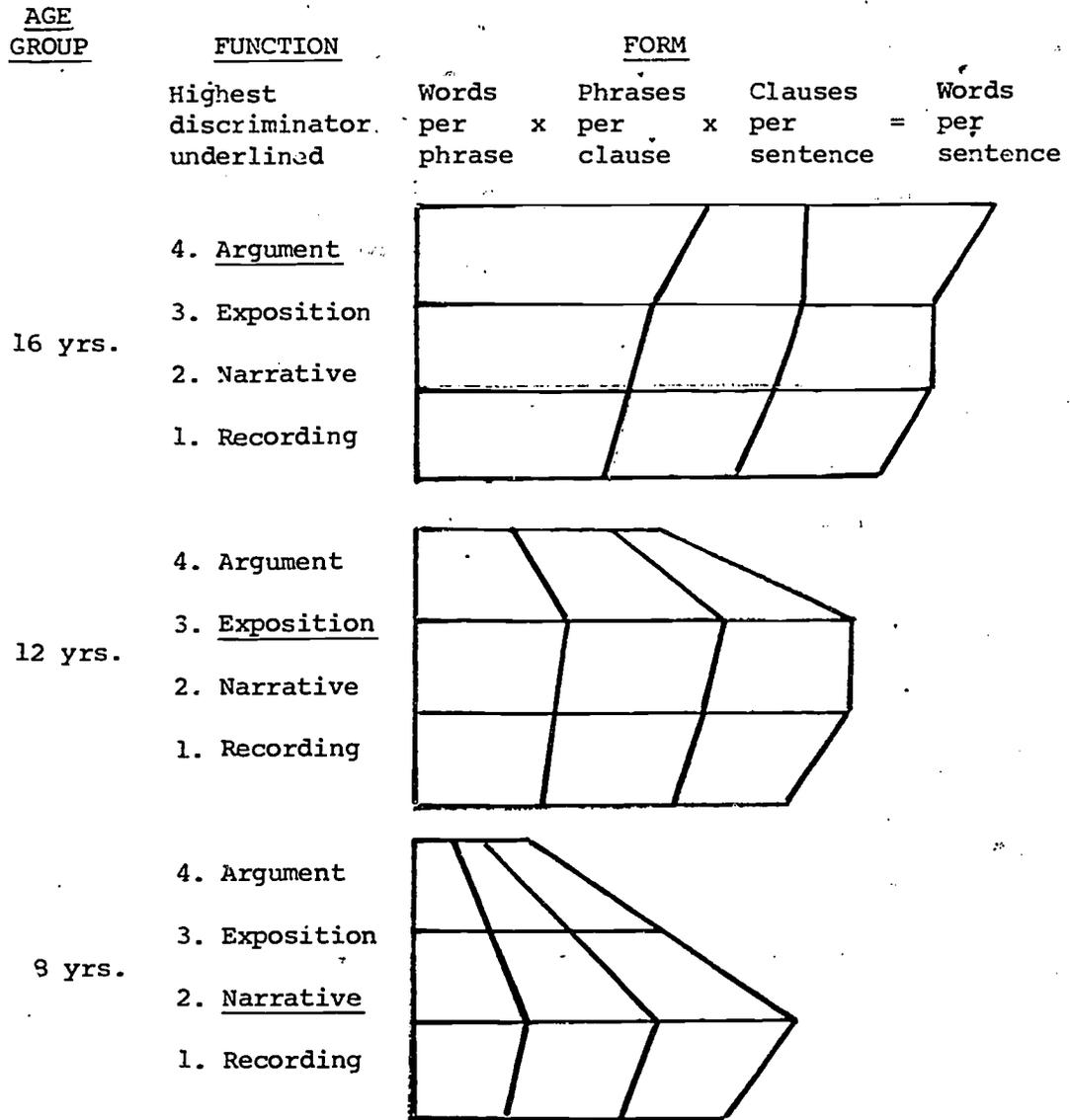
on relationships between events. In Argument, there is still more pressure on describing, less on relationships between phenomena and more on relationships between events.

The hypothesis that a person at the stage of development where Recording is possible, though there is little grasp of the other forms of discourse, is capable of the higher modes in the sense of possessing the requisite verbal structures. What is lacking is the cognitive power to relate many phenomena to one another in more and more elaborately-organised patterns.

On this hypothesis, differences in modes of discourse and in developmental levels related to them are differences of powers of thought more than verbal differences as such. The evidence is that the teaching of formal grammar does not affect the competency of writing (Elley 1971, Bray 1971), but that the subject-matter of the discourse (eg. whether physical or humanities-type problems) does affect the competency, (Connell et.al.1975). Such findings are compatible with a model which would lead to the prediction that learning experiences promoting more complex thought about subject-matter of interest and concern would do more to develop powers of discourse than formal instruction directly concerned with language structures.

Such interpretations must be deemed speculative until there is more evidence, and in the interest of providing hypotheses to test, the study closes with a developmental model of the four forms of discourse with reference to both discrimination indices and structural features of the discourse.

Figure 10 : Developmental model of adaptation of form to function.



## BIBLIOGRAPHY

- Arthur, B. Teaching English to Speakers of English.  
Harcourt, Brace Jovanovich, New York, 1973.
- Australian National Advisory Committee for U.N.E.S.C.O.  
The Teaching of English : Australian UNESCO  
Seminar. Australian Government Printer,  
Canberra. 1973
- Barnes, D. (Ed.) Language, the Learner and the School.  
Penguin. Harmondsworth. 1969.
- Bray, S.A.C. English Language Studies in Secondary Schools  
Sydney Teachers' College Monograph No. 1.  
Sydney. 1971.
- Britton, J. et.al. Multiple Marking of English Compositions  
Schools Council Examinations Bulletin No. 12.  
H.M.S.O. London. 1966
- Britton, J. Language and Learning. Allen Lane. The Penguin  
Press. London. 1970.
- Britton, J. "Whats the Use? A Schematic Account of Language  
Functions". Educational Review, Vol.23 No.3. 1971.
- Cambourne, B.L. A Naturalistic Study of Language Performance  
in Grade I Rural and Urban School Children. Ph.D.  
dissertation. The James Cook University of North  
Queensland. 1971.
- Chomsky, N. "Language and the Mind". Psychology Today,  
Feb., 1969.
- Connell, W. et.al. 12 to 20 : Studies of City Youth.  
Hicks Smith and Sons. Sydney. 1975.
- Davis, J. "A special Study in Contemporary English Language VI",  
School of Liberal Studies, Canberra College of Advanced  
Education. 1975.
- Dunkin, M.J. and Biddle, B.J. The Study of Teaching. Holt,  
Rinehart and Winston. New York. 1974.
- Ebel, R.L. Measuring Educational Achievement. Prentice Hall,  
Englewood Cliffs, New Jersey. 1965.
- Elley, W.B. "Research into the Teaching of English". English  
in Australia, No. 16. 1971.

- Ferguson, C.A. and Slobin, D.I. Studies in Child Language Development. Holt, Rinehart and Winston. New York. 1973.
- Flesch, R.F. "A New Readability Yardstick". Journal of Applied Psychology, Vol.32. 1948.
- Fries, C.C. The Structure of English. Harcourt, Brace and World. New York. 1952.
- Giglioli, P.P. (Ed.) Language and Social Context. Penguin, Harmondsworth. 1972.
- Gilliand, J. Readability. University of London Press. London. 1972.
- Godshalk, F.I. et.al. The Measurement of Writing Ability. College Entrance Examination Board. New York. 1966.
- Halliday, M.A.K. "Relevant Models of Language". Educational Review, Vol. 22. 1969.
- Her Majesty's Stationery Office. A Language for Life. ("The Bullock Report"). London. 1975.
- Hunt, K.W. Grammatical Structures Written at Three Grade Levels. National Council of Teachers of English. Research Report No. 3. 1965.
- Hunt, K.W. Syntactic Maturity in School Children and Adults. Monographs of the Society for Research in Child Development, Vol. 35, No.1. 1970.
- Interim Australian Capital Territory Schools Authority. Secondary English Workshop Report. Canberra 1974 and Primary English Workshop Report. 1975.
- Johnston, W.R. Kellog Hunt's T-Unit. Centre for Research in Learning and Instruction. New South Wales Department of Education. 1973.
- Labov, W. "The Logic of Nonstandard English". Georgetown Monographs on Language and Linguistics, Vol. 22. 1969. (Excerpts reprinted in Giglioli 1972).
- Labov, W. "The Study of Language in its Social Context". Studium Generale, Vol. 23. 1970. (Reprinted in Giglioli 1972).
- Lief, L. and Light, J.F. The Modern Age : Literature. Holt, Rinehart and Winston. New York. 1972.
- Little, G. "Grammar, Task and Socioeconomic Status". M.Ed. Papers in Sociology of Education, School of Education, University of New South Wales. 1973.

- Little, G. "What the School Certificate Examination has Taught Us". An English-History Bulletin. New South Wales Department of Education. 1974.
- Loban, W.D. The Language of Elementary School Children. National Council of Teachers of English. Research Report No. 1. 1963.
- Loban, W.D. Problems in Oral English. National Council of Teachers of English. Research Report No. 5. 1969.
- Lyons, J. Chomsky. Fontana Modern Masters. London. 1970.
- Lyons, J. Structural Semantics. Blackwell. Oxford. 1963.
- Maling-Keepes, J. and Rechter, B. English and its Assessment. Australian Council for Educational Research. 1973.
- Martin, N. "Talking about Writing". English in Australia, No. 30. 1975.
- Moffett, J. Teaching the Universe of Discourse. Houghton Mifflin. New York. 1968 (a).
- Moffett, J. A Student-Centered Language-Arts Curriculum: Grades K-13. Houghton Mifflin. New York. 1968 (b).
- Newsome, B. "Bernard Newsome on Curriculum". English in Australia, No. 30. 1975.
- New South Wales Secondary Schools Board. Syllabus in English for Forms I-IV. Sydney. 1972 and Report on the School Certificate Examination in English. 1972.
- Piaget, J. "The Stages of the Intellectual Development of the Child". Bulletin of the Menninger School of Psychiatry. March, 1961.
- Popham, W.J. Educational Statistics: Use and Interpretation. Harper and Row. New York. 1967.
- Rosen, H. "Written Language and the Sense of Audience". Educational Research, Vol. 15, No. 3. 1973.
- Smith, L.E.W. Towards a New English Curriculum. Dent. London. 1972.
- Turner, G.W. Stylistics. Penguin. Harmondsworth. 1973.

APPENDIX A

THE ENGLISH REFERENCE TEST

APPENDIX A : THE ENGLISH REFERENCE TEST

The English Reference Test was taken by all subjects as part of a study undertaken by the Centre for Investigation into Measurement and Evaluation of the New South Wales Education Department. The test was part of a project to investigate the efficacy of a short reference-test, held two-thirds of the way through the school year, as a moderator of school assessments, in the planned event of phasing out a public examination.

The terms of reference for the test arose from the current English Syllabus (New South Wales Secondary Schools Board, 1972). It aimed to test competence in reading and writing. The terms of reference for school assessments derived from the same syllabus, except that oral English was explicitly included.

The form of the test derived from previous School Certificate Examinations, on which it had been found (Little 1974) that a short three-question form correlated 0.82 with a longer five-question form.

The reading section comprised two passages of poetry and two of prose, with a mean of eight multiple-choice questions asked about each. The set of questions was compiled from questions established to be of statistically significant discriminators

on similar subjects. Three sets of questions were in comprehension and the fourth set were about improving the expression of a faulty passage. Half of this material was the same for the advanced and the less advanced subjects, to assist in determining a common scale for results at the two levels. Candidates were advised to spend 20 minutes on this section, and on the whole completed it satisfactorily in the time.

The remainder of the test comprised two essays, candidates being advised to spend 35 minutes on each.

The first essay involved response to a picture of a peasant figure silhouetted dancing against the skyline. The candidate was asked to imagine that he or she was one of the figures and to write about what he or she was doing, thinking and feeling. This question was common to the advanced and less advanced subjects.

The second essay involved writing about literature studied. For advanced subjects, the question hinged upon an abstract statement about the value of reading to be applied to the subject's own reading. For the less advanced, the question involved identifying a story that had impressed the reader, asking for reasons why it was so impressive.

Responses to the first section were machine-marked. The essays were marked corporately (i.e. with markers working

together at a marking centre) by a method involving

1. separate marking of each question;
2. general briefing sessions on criteria of evaluation and use of scales, under a system of impression marking rather than analytical marking, impression marking being recommended on the bases of empirical studies by Britton (1966), Godshalk, et.al. (1966) and Maling-Keepes and Rechter (1973);
3. group and individual readings of sample scripts marked by senior examiners;
4. trial marking by examiners (results not being counted), for the sake of comparisons of markings by various examiners and checking by seniors;
5. the marking proper.

The marking was conducted by teams of three examiners freely comparing and discussing scripts and referring problems to seniors, who would consult with other teams and seniors. Seniors also continually checked marking by monitoring statistical results and reading sample marked scripts. Any marking deviating from the agreed criteria was put through the system again.

These techniques of essay-marking are in accord with recommendations made on the basis of studies of reliability noted above. The encouragement of examiners to work and talk together continually was, however, an innovation. This step appeared to raise output and reduce boredom and fatigue

without any evident diminution of reliability. Studies of the reliability of this type of marking in previous School Certificate examinations had established that an essay marked in this manner provided as good a discrimination index as a set of multiple-choice items such as that outlined above (Little 1974).

Marks were collated by computer, and computer techniques were used to derive an English result consisting of the School Assessment as moderated by the Reference Test, i.e. assessments from schools which (on the evidence of the test) had overestimated or underestimated the competence of their candidature were adjusted accordingly. This result was compared, on a variety of criteria, with the result derived from the school assessment as moderated by a full-scale examination. The results were sufficiently close to justify the replacement of the examination with the Reference Test in the ensuing year.

APPENDIX B

STATISTICAL TABLES

FOR WORD-COUNTS

Table 9 : Mean word-counts for 128 subjects on four writing tasks.  
(Standard deviations in brackets).

TASK	WORDS PER PHRASE	WORDS PER CLAUSE	WORDS PER SENTENCE	WORDS PER TASK	WORDS PER "AND"	NO. OF CASES WITHOUT "AND"
Recording	3.57 (1.28)	5.23 (2.60)	5.77 (3.31)	13.1 (5.46)	0 (0)	128
Narrative	3.85 (0.74)	10.8 (3.01)	18.8 (7.76)	126 (37.8)	33.8 (21.5)	8
Exposition	4.15 (0.66)	12.30 (3.75)	18.57 (7.09)	119 (40.2)	50.8 (32.5)	16
Argument	5.07 (1.20)	11.63 (3.35)	23.02 (8.61)	87 (40.1)	48.3 (32.8)	46

Table 10 : Mean word-counts for 64 advanced and 64 not advanced subjects on four writing tasks.  
(Standard deviations in brackets).

TASK	GROUP	WORDS PER PHRASE	WORDS PER CLAUSE	WORDS PER SENTENCE	WORDS PER TASK	WORDS PER "AND" OMITTING CASES OF NO "AND"	NO. OF CASES WITH NO "AND"
Recording	Advanced	3.42 (1.01)	4.99 (2.87)	5.72 (3.93)	12.8 (4.62)	0	64
	Not Advanced	3.72 (1.48)	5.47 (2.45)	5.82 (2.54)	13.3 (6.36)	0	64
Narrative	Advanced	3.85 (0.65)	10.6 (2.71)	18.3 (7.10)	139 (39.4)	34.1 (28.0)	0
	Not Advanced	3.85 (0.81)	11.0 (3.27)	19.4 (8.34)	114 (36.6)	32.7 (19.6)	8
Exposition	Advanced	4.20 (0.71)	12.8 (4.11)	19.4 (8.19)	129 (45.4)	60.6 (43.5)	16
	Not Advanced	4.10 (0.61)	11.8 (3.27)	17.8 (5.63)	109 (29.9)	50.0 (35.5)	0
Argument	Advanced	5.17 (1.12)	12.4 (3.78)	23.9 (7.62)	102.2 (45.1)	53.5 (34.8)	20
	Not Advanced	5.05 (1.30)	10.3 (2.72)	20.8 (9.29)	71.8 (34.9)	43.2 (21.4)	26