

A LONGMAN PAPERBACK



Cohesion in English

M.A.K. Halliday
and Ruqaiya Hasan



Cohesion in English

M. A. K. HALLIDAY

*Professor of Linguistics
University of Sydney*

RUQAIYA HASAN

*Senior Lecturer
School of English and Linguistics
Macquarie University*



LONGMAN
London and New York

Foreword

Throughout more than a century of outstanding progress in linguistics – and especially from the time of the *Junggrammatiker* – the most impressive and apparently most abiding successes have been in work at the elemental end of language structure: the description and relation of phonological units. Nor, when they were pressed into reluctant service, did the categories and insights evolved for phonology get us far in explicating linguistic organization at other ‘levels’, the morphological and syntactic. Moreover, even in the fruitful renaissance of syntactic studies during the third quarter of this century, work has been virtually confined to relations within the sentence. This limitation, though to some extent vigorously defended on theoretical grounds, has not in general been because no relevance to linguistic structure was seen in the relations between sentences, in the connections which resulted in the impression of well-formed paragraphs or longer stretches of discourse. But as with semantics – another and indeed closely related area which linguists have hesitated to enter, often justifying their dissociation on closely-argued theoretical grounds – it was not unreasonably held that relations ‘beyond the sentence’ involved a complex interplay of linguistics with other concerns such as rhetoric, aesthetics, and pragmatics, for which the theoretical foundations and framework were too shaky to support ambitious model building. And that in any case linguists had enough on hand to get their sentential house furnished.

Meanwhile, literary critics (for whom of course text structure has been a traditional concern) and social anthropologists (for whom text and tale constitute fundamental evidence) began themselves to look at the constructs evolved by de Saussure, the Prague School, and other linguists. One thinks for example of Lévi-Strauss, Dell Hymes, Roland Barthes, as outstanding exponents of structuralism in broad-scale textual analysis. And among linguists, there have always been those who have persisted in the

venture to subserve literary and other humanistic disciplines by extending their work to embrace stylistics and other aspects of textual studies. In this movement, Michael Halliday and Ruqaiya Hasan have long been especially active. The prose of Golding and the verse of Yeats are among the material subjected to valued linguistic scrutiny by the former, while the latter has made 'cohesion' her special field, beginning with a doctoral dissertation at the University of Edinburgh and continuing with influential papers while she worked for several fruitful years in the Communication Research Centre at University College London. During the whole of this period, the two authors have worked in close cooperation and mutual influence, acutely aware of areas in English studies of profound interest for both linguists and critics but rigorously explored to a large extent by neither.

We are singularly fortunate that we are able to correct some of these grave deficiencies in the description of English with the work of so uniquely equipped a team. As English has increasingly come into world-wide use, there has arisen a correspondingly increasing need for more information on the language and the ways in which it is used. The English Language Series seeks to meet this need and to play a part in further stimulating the study and teaching of English by providing up-to-date and scholarly treatments of topics most relevant to present-day English – including its history and traditions, its sound patterns, its grammar, its lexicology, its rich variety and complexity in speech and writing, and its standards in Britain, the USA, and the other principal areas where the language is used.

University College London
May 1975

RANDOLPH QUIRK

Preface

This book originated as one of a series of studies of the English language and modern English texts which were undertaken by the *Nuffield Programme in Linguistics and English Teaching* at University College London. The aim of these studies was to provide an account of aspects of contemporary English which would be both founded on theory and also applicable in practice: a description of the system, but one which, since it was based on evidence from texts of different varieties, including both spoken and written, would be useful in application to further text studies.

A relatively neglected aspect of the linguistic system is its resources for text construction, the range of meanings that are specifically associated with relating what is being said or written to its semantic environment. The principal component of these resources is that of cohesion. Cohesive relations are relations between two or more elements in a text that are independent of the structure; for example between a personal pronoun and an antecedent proper name, such as *John . . . he*. A semantic relation of this kind may be set up either within a sentence or between sentences; with the consequence that, when it crosses a sentence boundary, it has the effect of making the two sentences cohere with one another. The various kinds of cohesion had been outlined by M. A. K. Halliday in his writings on stylistics, and the concept was developed by Ruqaiya Hasan in her University of Edinburgh doctoral thesis.

The earlier chapters of this book were first published as *Grammatical Cohesion in Spoken and Written English, Part I*, by Ruqaiya Hasan, Communication Research Centre (University College London) and Longmans, Green & Co, *Programme in Linguistics and English Teaching: Papers*, No. 7, 1968. This contained Chapters 1, 2 and 3 in their original form. The later chapters were written in collaboration by Ruqaiya Hasan and M. A. K. Halliday, and were prepared for publication in the follow-up series (*Schools Council Programme in Linguistics and English Teaching: Papers*

Series II). However, instead of issuing this part separately, it was decided to revise the earlier chapters and to publish the two halves together as a book. The revision was undertaken by M. A. K. Halliday, who also added the last two chapters.

We should like to express our gratitude to several individuals and institutions for their cooperation and help. The Nuffield Foundation financed the original project within which the earlier part of the work was written. The Schools Council financed the successor project (*Schools Council Programme in Linguistics and English Teaching*, 1967-71); although the later part was not written directly under their auspices, since Ruqaiya Hasan had by then left the team, it had been planned to publish it in the series of papers emanating from this project, and we are grateful to them for allowing it to be withdrawn and published in its present revised form. The final version was written by M. A. K. Halliday during his tenure of a fellowship at the Center for Advanced Study in the Behavioral Sciences, Stanford, California, and we are most grateful to the Center for providing this opportunity.

We wish to thank Stephen Lushington, General Editor of the *Schools Council Programme in Linguistics and English Teaching: Papers Series II*, and a former colleague in the project, for his valuable help and comments throughout the preparation of the original manuscript. Other members of the Nuffield team – Kenneth Albrow, Eirian Davies, Peter Doughty, David Mackay and Brian Thompson – provided stimulating discussion, as did our colleagues on another related research project, Rodney Huddleston, Richard Hudson and Eugene Winter. To Marcia Insel we express our appreciation for her research and bibliographical assistance during the final revision. Students at the Linguistic Society of America's Linguistic Institute, at the University of Michigan, Ann Arbor, in summer 1973, made numerous helpful observations in the context of a course based on this material.

We much appreciate the interest shown by Randolph Quirk, friend, former colleague, and General Editor of the present series; and would like to take this opportunity of referring to the debt owed by everyone in the field of contemporary English to the work done by him and by his colleagues at the Survey of English Usage. Finally we thank the many people who have kindly enquired after the progress of the book. Their continuing concern has been a most valuable source of encouragement.

University of Essex
May 1975

MAKH
RH

Acknowledgments

We are grateful to the following for permission to reproduce copyright material:

Author's agents for the sonnet 'The Bad Thing' by John Wain; Gerald Duckworth & Company Ltd for 'The Hippopotamus' from *The Bad Child's Book of Beasts* by Hilaire Belloc; Granada Publishing Ltd for extracts from *Class, Codes and Control* Vol 1 by Basil Bernstein, published by Paladin Books; The Proprietor of The Greenwich Bookshop for extracts from *Royal Greenwich* by Olive and Nigel Hamilton, The Greenwich Bookshop 1969; the Author for an extract from the article 'Meeting Wilfred Pickles' by Frank Haley from *The Dalesman* September 1973; Author's agents for extracts from 'An Inspector Calls' from *The Plays of J. B. Priestley* Vol 3 published by William Heinemann Ltd. Reprinted by permission of A. D. Peters and Company and Author's agents, M. B. Yeats, Miss Anne Yeats, Macmillan of London & Basingstoke, Macmillan of Canada and Macmillan Publishing Company Inc for an extract from *The Autobiography of William Butler Yeats*. Copyright © 1916, 1935 by Macmillan Publishing Co Inc, renewed 1944, 1963 by Bertha Georgie Yeats.

Table of Contents

<i>Foreword</i>	v
-----------------	---

<i>Preface</i>	vii
----------------	-----

I Introduction

I.1 The concept of cohesion	I
I.1.1 Text	I
I.1.2 Texture	2
I.1.3 Ties	3
I.1.4 Cohesion	4
I.2 Cohesion and linguistic structure	6
I.2.1 Texture and structure	6
I.2.2 Cohesion within the sentence?	7
I.2.3 Cohesion and discourse structure	10
I.2.4 Cohesion as a semantic relation	11
I.3 Cohesion and linguistic context	14
I.3.1 The domain of cohesive relations	14
I.3.2 Text and situation	19
I.3.3 Components of the context of situation, and register	21
I.3.4 The place of cohesion in the linguistic system	26
I.3.5 The meaning of cohesion	28

2 Reference

2.1 Endophoric and exophoric reference	31
2.2 Types of reference	37
2.3 Personal reference	43
2.3.1 Semantic distinctions in the personal system	45
2.3.2 Speech roles and other roles	48

2.3.3 Some special kinds of personal reference	52
✓ 2.3.3.1 Extended reference, and text reference	52
2.3.3.2 Generalized exophoric reference	53
2.3.4 Personal pronouns, possessive determiners and possessive pronouns	54
2.3.5 Cataphoric reference	56
2.4 Demonstrative reference	57
2.4.1 The selective nominal demonstratives: <i>this, these, that, those</i>	59
2.4.1.1 Near and not near: <i>this/these</i> versus <i>that/those</i>	60
2.4.1.2 Singular and plural: <i>this/that</i> versus <i>these/those</i>	62
2.4.1.3 Head and modifier: <i>this</i> , etc, as pronoun versus <i>this</i> , etc, plus following noun	62
2.4.1.4 Extended reference and reference to 'fact': <i>this</i> and <i>that</i>	66
2.4.1.5 Anaphoric and cataphoric demonstratives	68
2.4.2 <i>The</i>	70
2.4.3 Demonstrative adverbs	74
2.4.4 A final note on demonstratives	75
2.5 Comparative reference	76
2.5.1 General comparison	77
2.5.2 Particular comparison	80
2.5.3 A note on <i>so, such</i> and <i>as</i>	84

3 Substitution

3.1 Substitution and ellipsis	88
3.1.1 Substitute and reference	88
3.1.2 Types of substitution	90
3.2 Nominal substitution	91
3.2.1 The meaning of substitute <i>one/ones</i>	92
3.2.2 Conditions of use of the nominal substitute	95
3.2.3 The word <i>one</i> other than as substitute	98
3.2.3.1 Personal pronoun <i>one</i>	98
3.2.3.2 Cardinal numeral <i>one</i>	98
3.2.3.3 Indefinite article <i>one</i>	100
3.2.3.4 'Pro-noun' <i>one</i>	102
3.2.4 Summary of uses of <i>one</i>	104
3.2.5 Nominal substitute <i>same</i>	105
3.2.5.1 <i>say the same</i>	107
3.2.5.2 <i>do the same</i>	108
3.2.5.3 <i>be the same</i>	109

3.2.6 Difference between <i>the same</i> and <i>one(s)</i> as nominal substitutes	110
3.3 Verbal substitution	112
3.3.1 The meaning of the verbal substitute <i>do</i>	113
3.3.2 Conditions of use of the verbal substitute	117
3.3.3 The word <i>do</i> other than as substitute	123
3.3.3.1 Lexical verb <i>do</i>	124
3.3.3.2 General verb <i>do</i>	124
3.3.3.3 Pro-verb <i>do</i>	125
3.3.3.4 Verbal operator <i>do</i>	127
3.3.4 Summary of uses of <i>do</i>	128
3.4 Clausal substitution	130
3.4.1 Difference between clausal and other types of substitution	130
3.4.1.1 Substitution of reported clauses	131
3.4.1.2 Substitution of conditional clauses	134
3.4.1.3 Substitution of modalized clauses	134
3.4.2 Similarity among the types of clausal substitution	135
3.4.3 Some related patterns	137
3.4.3.1 Response forms	137
3.4.3.2 Other uses of <i>so</i> and <i>not</i>	138
3.4.4 Summary of uses of <i>so</i>	139

4 Ellipsis

4.1 Ellipsis, substitution and reference	142
4.2 Nominal ellipsis	147
4.2.1 Ellipsis within the nominal group	147
4.2.2 Presupposition of nominal elements	150
4.2.3 Types of nominal ellipsis	153
4.2.3.1 Specific deictics	155
4.2.3.2 Non-specific deictics	157
4.2.3.3 Post-deictics	159
4.2.3.4 Numeratives	161
4.2.3.5 Epithets	163
4.3 Verbal ellipsis	167
4.3.1 Ellipsis within the verbal group	167
4.3.2 Lexical ellipsis	170
4.3.3 Operator ellipsis	174
4.3.4 Presupposition of verbal group systems	176
4.3.4.1 Polarity	176
4.3.4.2 Finiteness and modality	180

4.3.4.3 Voice	182
4.3.4.4 Tense	186
4.3.5 Summary of verbal ellipsis	192
4.3.6 Verbal ellipsis and the clause	194
4.4 Clausal ellipsis	196
4.4.1 Modal and propositional	196
4.4.2 No ellipsis of single elements	202
4.4.3 Ellipsis in question-answer and other rejoinder sequences	206
4.4.3.1 Direct responses (1): yes/no questions	208
4.4.3.2 Direct responses (2): WH- questions	210
4.4.3.3 Indirect responses	212
4.4.3.4 A note on zeugma	214
4.4.3.5 Other rejoinders	214
4.4.4 Ellipsis in 'reporting-reported' sequences	217
4.4.4.1 Indirect WH- questions	217
4.4.4.2 Indirect yes/no questions	218
4.4.4.3 Indirect statements	219
4.4.4.4 Ambiguity between indirect statements and indirect questions	220
4.4.4.5 Reports and facts in relation to clausal ellipsis	221
4.4.5 Clausal ellipsis and clause complexes	222

5 Conjunction

5.1 Conjunction and other cohesive relations	226
5.1.1 Structural equivalents of conjunctive relations	227
5.1.2 Types of conjunctive expression	230
5.2 Some common conjunctive elements	233
5.2.1 The 'and' relation	233
5.2.2 Coordinate <i>and</i> and conjunctive <i>and</i>	235
5.2.3 Other conjunctive elements: <i>but</i> , <i>yet</i> , <i>so</i> , and <i>then</i>	237
5.3 Types of conjunction	238
5.4 Additive	244
5.5 Adversative	250
5.6 Causal	256
5.7 Temporal	261
5.8 Other conjunctive items (continuatives)	267
5.8.1 <i>now</i>	268
5.8.2 <i>of course</i>	269
5.8.3 <i>well</i>	269

5.8.4 <i>anyway</i>	270
5.8.5 <i>surely</i>	270
5.8.6 <i>after all</i>	270
5.9 The cohesive function of intonation	271

6 Lexical cohesion

6.1 The class of 'general nouns'	274
6.2 Types of reiteration	277
6.3 Lexical relations as cohesive patterns	282
6.4 Collocation	284
6.5 The general concept of lexical cohesion	288

7 The meaning of cohesion

7.1 Text	293
7.1.1 Length of text	294
7.1.2 Definitiveness of the concept of text	294
7.1.3 Tight and loose texture	295
7.1.4 Imaginary texture	297
7.2 The general meaning of cohesion	298
7.3 The meaning of the different kinds of cohesion	303
7.3.1 General principles behind the different types	304
7.3.2 Reference	308
7.3.3 Substitution and ellipsis	314
7.3.4 Lexical cohesion: reiteration and collocation	318
7.3.5 Conjunction	320
7.3.6 Summary	322
7.4 Cohesion and the text	324
7.4.1 Texture within the sentence	325
7.4.2 The texture of discourse	326
7.4.3 The role of linguistic analysis	327

8 The analysis of cohesion

8.1 General principles	329
8.2 Summary of cohesion, and coding scheme	333
8.3 Sample texts	340
<i>Bibliography</i>	357
<i>Index</i>	367

Chapter 1

Introduction

1.1 The concept of cohesion

1.1.1 *Text*

If a speaker of English hears or reads a passage of the language which is more than one sentence in length, he can normally decide without difficulty whether it forms a unified whole or is just a collection of unrelated sentences. This book is about what makes the difference between the two.

The word **TEXT** is used in linguistics to refer to any passage, spoken or written, of whatever length, that does form a unified whole. We know, as a general rule, whether any specimen of our own language constitutes a **TEXT** or not. This does not mean there can never be any uncertainty. The distinction between a text and a collection of unrelated sentences is in the last resort a matter of degree, and there may always be instances about which we are uncertain – a point that is probably familiar to most teachers from reading their students' compositions. But this does not invalidate the general observation that we are sensitive to the distinction between what is text and what is not.

This suggests that there are objective factors involved – there must be certain features which are characteristic of texts and not found otherwise; and so there are. We shall attempt to identify these, in order to establish what are the properties of texts in English, and what it is that distinguishes a text from a disconnected sequence of sentences. As always in linguistic description, we shall be discussing things that the native speaker of the language 'knows' already – but without knowing that he knows them.

A text may be spoken or written, prose or verse, dialogue or monologue. It may be anything from a single proverb to a whole play, from a momentary cry for help to an all-day discussion on a committee.

A text is a unit of language in use. It is not a grammatical unit, like a clause or a sentence; and it is not defined by its size. A text is sometimes

envisaged to be some kind of super-sentence, a grammatical unit that is larger than a sentence but is related to a sentence in the same way that a sentence is related to a clause, a clause to a group and so on: by CONSTITUENCY, the composition of larger units out of smaller ones. But this is misleading. A text is not something that is like a sentence, only bigger; it is something that differs from a sentence in kind.

A text is best regarded as a SEMANTIC unit: a unit not of form but of meaning. Thus it is related to a clause or sentence not by size but by REALIZATION, the coding of one symbolic system in another. A text does not CONSIST of sentences; it is REALIZED BY, or encoded in, sentences. If we understand it in this way, we shall not expect to find the same kind of STRUCTURAL integration among the parts of a text as we find among the parts of a sentence or clause. The unity of a text is a unity of a different kind.

1.1.2 Texture

The concept of TEXTURE is entirely appropriate to express the property of 'being a text'. A text has texture, and this is what distinguishes it from something that is not a text. It derives this texture from the fact that it functions as a unity with respect to its environment.

What we are investigating in this book are the resources that English has for creating texture. If a passage of English containing more than one sentence is perceived as a text, there will be certain linguistic features present in that passage which can be identified as contributing to its total unity and giving it texture.

Let us start with a simple and trivial example. Suppose we find the following instructions in the cookery book:

[1:1] Wash and core six cooking apples. Put them into a fireproof dish.

It is clear that *them* in the second sentence refers back to (is ANAPHORIC to) the *six cooking apples* in the first sentence. This ANAPHORIC function of *them* gives cohesion to the two sentences, so that we interpret them as a whole; the two sentences together constitute a text. Or rather, they form part of the same text; there may be more of it to follow.

The texture is provided by the cohesive RELATION that exists between *them* and *six cooking apples*. It is important to make this point, because we shall be constantly focusing attention on the items, such as *them*, which typically refer back to something that has gone before; but the cohesion is effected not by the presence of the referring item alone but by the presence

of both the referring item and the item that it refers to. In other words, it is not enough that there should be a presupposition; the presupposition must also be satisfied. This accounts for the humorous effect produced by the radio comedian who began his act with the sentence

[1:2] So we pushed him under the other one.

This sentence is loaded with presuppositions, located in the words *so*, *him*, *other* and *one*, and, since it was the opening sentence, none of them could be resolved.

What is the MEANING of the cohesive relation between *them* and *six cooking apples*? The meaning is that they refer to the same thing. The two items are identical in reference, or COREFERENTIAL. The cohesive agency in this instance, that which provides the texture, is the coreferentiality of *them* and *six cooking apples*. The signal, or the expression, of this coreferentiality is the presence of the potentially anaphoric item *them* in the second sentence together with a potential target item *six cooking apples* in the first.

Identity of reference is not the only meaning relation that contributes to texture; there are others besides. Nor is the use of a pronoun the only way of expressing identity of reference. We could have had:

[1:3] Wash and core six cooking apples. Put the apples into a fireproof dish.

Here the item functioning cohesively is *the apples*, which works by repetition of the word *apples* accompanied by *the* as an anaphoric signal. One of the functions of the definite article is to signal identity of reference with something that has gone before. (Since this has sometimes been said to be its only function, we should perhaps point out that it has others as well, which are not cohesive at all; for example none of the instances in (a) or (b) has an anaphoric sense:

[1:4] a. None but the brave deserve the fair.

b. The pain in my head cannot stifle the pain in my heart.

For the meaning of *the*, see 2.4.2 below.)

1.1.3 Ties

We need a term to refer to a single instance of cohesion, a term for one occurrence of a pair of cohesively related items. This we shall call a TIE. The relation between *them* and *six cooking apples* in example [1:1] constitutes a tie.

We can characterize any segment of a text in terms of the number and

kinds of ties which it displays. In [1:1] there is just one tie, of the particular kind which we shall be calling REFERENCE (Chapter 2). In [1:3], there are actually two ties, of which one is of the 'reference' kind, and consists in the anaphoric relation of *the* to *six cooking apples*, while the other is of a different kind and consists in the REPETITION of the word *apples*, a repetition which would still have a cohesive effect even if the two were not referring to the same apples. This latter type of cohesion is discussed in Chapter 6.

The concept of a tie makes it possible to analyse a text in terms of its cohesive properties, and give a systematic account of its patterns of texture. Some specimen analyses are given in Chapter 8. Various types of question can be investigated in this way, for example concerning the difference between speech and writing, the relationship between cohesion and the organization of written texts into sentences and paragraphs, and the possible differences among different genres and different authors in the numbers and kinds of tie they typically employ.

The different kinds of cohesive tie provide the main chapter divisions of the book. They are: reference, substitution, ellipsis, conjunction, and lexical cohesion. A preliminary definition of these categories is given later in the Introduction (1.2.4); each of these concepts is then discussed more fully in the chapter in question.

1.1.4 Cohesion

The concept of cohesion is a semantic one; it refers to relations of meaning that exist within the text, and that define it as a text.

Cohesion occurs where the INTERPRETATION of some element in the discourse is dependent on that of another. The one PRESUPPOSES the other, in the sense that it cannot be effectively decoded except by recourse to it. When this happens, a relation of cohesion is set up, and the two elements, the presupposing and the presupposed, are thereby at least potentially integrated into a text.

This is another way of approaching the notion of a tie. To return to example [1:1], the word *them* presupposes for its interpretation something other than itself. This requirement is met by the *six cooking apples* in the preceding sentence. The presupposition, and the fact that it is resolved, provide cohesion between the two sentences, and in so doing create text.

As another example, consider the old piece of schoolboy humour:

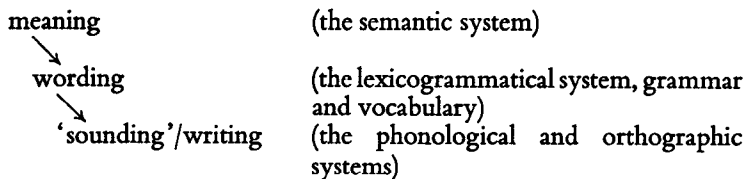
[1:5] Time flies.

– You can't; they fly too quickly.

The first sentence gives no indication of not being a complete text; in fact it usually is, and the humour lies in the misinterpretation that is required if the presupposition from the second sentence is to be satisfied. Here, incidentally, the cohesion is expressed in no less than three ties: the elliptical form *you can't* (Chapter 4), the reference item *they* (Chapter 2) and the lexical repetition *fly* (Chapter 6).

Cohesion is part of the system of a language. The potential for cohesion lies in the systematic resources of reference, ellipsis and so on that are built into the language itself. The actualization of cohesion in any given instance, however, depends not merely on the selection of some option from within these resources, but also on the presence of some other element which resolves the presupposition that this sets up. It is obvious that the selection of the word *apples* has no cohesive force by itself; a cohesive relation is set up only if the same word, or a word related to it such as *fruit* (see Chapter 6), has occurred previously. It is less obvious, but equally true, that the word *them* has no cohesive force either unless there is some explicit referent for it within reach. In both instances, the cohesion lies in the relation that is set up between the two.

Like other semantic relations, cohesion is expressed through the stratal organization of language. Language can be explained as a multiple coding system comprising three levels of coding, or 'strata': the semantic (meanings), the lexicogrammatical (forms) and the phonological and orthographic (expressions). Meanings are realized (coded) as forms, and forms are realized in turn (recoded) as expressions. To put this in everyday terminology, meaning is put into wording, and wording into sound or writing:



The popular term 'wording' refers to lexicogrammatical form, the choice of words and grammatical structures. Within this stratum, there is no hard-and-fast division between vocabulary and grammar; the guiding principle in language is that the more general meanings are expressed through the grammar, and the more specific meanings through the vocabulary. Cohesive relations fit into the same overall pattern. Cohesion is expressed partly through the grammar and partly through the vocabulary.