

SOME EMERGENT DISCOURSE CONNECTIVES IN ENGLISH:

GRAMMATICALIZATION VIA RHETORICAL PATTERNS

Diana M. Lewis

Worcester College, University of Oxford

Submitted to the Faculty of English Language and Literature, University of Oxford

for the degree of Doctor of Philosophy

Trinity Term 2000



ABSTRACT

Some emergent discourse connectives in English. Grammaticalization via rhetorical patterns

Diana Lewis
Worcester College

D.Phil
Trinity Term, 2000

Discourse connectives are metatextual comments that signal discourse coherence relations. They can be realized by sentence adverbials that have their roots in verb phrase adverbials and have followed a path of development that is well attested both for English and cross-linguistically. This study investigates how and why it occurs. It claims that the development belongs to a wider phenomenon of unidirectional internal semantic change, that this change involves context-induced reinterpretation, and that both the immediate discourse context and the wider rhetorical context can be instrumental in bringing it about.

Using diachronic and synchronic data in a corpus-linguistic approach, the frequency and distribution of the adverbials *after all*, *in fact*, *at least* and *of course* are investigated. These are found to follow similar paths of development at different rates and to varying extents. Each undergoes some increase in frequency, subjectification and abstraction, shift of discourse plane and categorial reanalysis. Each acquires at least one connective function to express rhetorical relations such as concession, contrast, justification or elaboration. These relations are defined using the framework of Rhetorical Structure Theory.

The analysis identifies, in the history of the expressions, contexts of co-occurrence with particular relations that are argued to generate connective implicatures which later crystallize. During long periods the expressions may have stable but defeasible implicatures in the relevant contexts. These contexts include rhetorical structures spanning two or more clause complexes and often consisting of quasi-conventional sequences of rhetorical relations typical of argumentation. They may be described as incipient discourse constructions or rhetorical idioms. The emergence of new discourse connectives is seen to share many of the features attested in the grammaticalization of lexical material. It is argued that these phenomena are best accounted for in a single, usage-based theory of internal semantic change.

To the memory of Sofia

Acknowledgements

I should like to thank my supervisor, Professor Jean Aitchison, for all her support. Heartfelt thanks also to Robin Barnsley, Reem Bassiouney, Julia Briggs, Subarno Chattarji, Brendan Fleming, Lynne Lewis, Raymond Lewis, Gillian Ramchand, Gabriella Saibene, Bi Scott, Mike Scott, Hiromi Tanaka and Ana María Tejos, who have all contributed more than they know to this work.

TABLE OF CONTENTS

Abstract	ii
Acknowledgements	iii
Key to symbols and abbreviations used in examples	viii
Introduction	1
1. Discourse connectives	9
1.1 Defining ‘discourse marker’ and ‘discourse connective’	9
1.2 Previous approaches to discourse markers	15
1.2.1 Functional approaches	17
1.2.2 Formal approaches	22
1.3 Discourse markers and discourse planes	27
1.4 Syntax and Semantics of DCs	34
1.4.1 Syntactic category	34
1.4.2 Semantic classification of S-Adv	39
1.4.3 Sub-categorization of speaker-oriented S-Adv	41
1.4.4 Position of S-Adv	46
1.4.5 Polysemy	48
1.5 Summary	50
2. Discourse coherence, rhetorical relations and grammaticalization	52
2.1 Discourse coherence	52
2.1.1 Discourse-pragmatic approaches	53
2.1.2 Linguistic approaches	56
2.1.3 Social action approaches	58
2.2 Units of discourse analysis	59
2.2.1 Spoken vs written language	59
2.2.2 Segmenting spoken and written language	61
2.3 Rhetorical Structure Theory (RST)	64
2.3.1 Some examples of RST relations	64
2.3.2 Rhetorical relations and DCs	67
2.3.3 Nuclearity	70
2.3.4 RST and discourse planes	73
2.3.5 Applying RST to the analysis of modal adverbial contexts	75
2.3.6 Rhetorical patterns	78
2.4 DCs and grammaticalization	79
2.4.1 The development of DCs as instances of grammaticalization	79
2.4.2 Recent approaches to characterising grammaticalization	83
2.4.3 Sentence-based vs discourse-based grammaticalization	85
2.4.4 Semantic / pragmatic change	88
2.4.5 The grammaticalization cycle	91
2.4.6 Grammaticalization, discourse connectives and polysemy	94
2.5 Conclusion	96

3. Outline of method and data	99
3.1 Introduction	99
3.2 Corpus linguistic method	99
3.2.1 The approach	99
3.2.2 Corpus linguistics and lexical semantics	101
3.2.3 Use of statistical method in corpus linguistics	104
3.2.4 Corpus design and sampling issues	105
3.2.5 Analysis of the data	105
3.3 The data used in this study	105
3.3.1 Selection of data	106
3.3.2 The synchronic data	108
3.3.3 The diachronic data: full-text sources	109
3.3.4 The diachronic data: non-full-text sources	110
3.4 Summary	111
4. Modal adverbs and polysemy	112
4.1 Modal sentence adverbials of certainty	112
4.2 <i>Clearly</i>	115
4.2.1 Historical development	115
4.2.2 Contexts of use	117
4.3 <i>Surely</i>	118
4.3.1 Historical development	118
4.3.2 Contexts of use	120
4.4 Conclusion	121
5. <i>After all</i>	122
5.1 Introduction	122
5.2 The history of <i>after all</i>	124
5.2.1 The data	124
5.2.2 Rhetorical patterns of <i>after all</i>	124
5.2.3 Summary of historical development	128
5.3 <i>After all</i> in PDE	129
5.3.1 Distribution	129
5.3.2 Counter-expectation <i>after all</i>	130
5.3.3 Justificative <i>after all</i>	131
5.3 Summary	134
6. <i>In fact</i>	136
6.1 Introduction	136
6.2 The history of <i>in fact</i>	137
6.2.1 The data	137
6.2.2 Development of epistemic / evidential <i>in fact</i>	138
6.2.3 Contrastive contexts of <i>in fact</i>	140
6.2.4 Elaborative contexts of <i>in fact</i>	143
6.2.5 Summary	146
6.3 <i>In fact</i> in PDE	147
6.3.1 Types of <i>in fact</i>	147
6.3.2 Emphasizer <i>in fact</i>	148

6.3.3 Antithetical <i>in fact</i>	149
6.3.4 Elaborative <i>in fact</i>	150
6.3.5 Distribution of <i>in fact</i>	152
6.4 Summary	153
7. <i>At least</i>	154
7.1 Introduction	154
7.1.1 Current definitions	154
7.1.2 Scalar <i>at least</i>	157
7.1.3 Retreat <i>at least</i>	158
7.1.4 Evaluative <i>at least</i>	161
7.2 The history of <i>at least</i>	163
7.2.1 The data	163
7.2.2 VP-adv > S-Adv	164
7.2.3 The emergence of Epistemic-Retreat <i>at least</i> : the <i>or at least</i> pattern	166
7.2.4 Towards Evaluative <i>at least</i> : the <i>if (not) p, (then) at least q</i> pattern	170
7.2.5 Summary of historical development	173
7.3 <i>At least</i> in PDE	174
7.3.1 Distribution of <i>at least</i>	174
7.3.2. Rhetorical contexts of <i>at least</i>	176
7.3.3 Layering	180
7.4 Conclusion	183
8. <i>Of course</i>	187
8.1 Introduction	187
8.2 The history of <i>of course</i>	189
8.2.1 The data	189
8.2.2 The origins of adverbial <i>of course</i>	189
8.2.3 Reanalysis of <i>of course</i> in the C18th	193
8.2.4 Functional split of <i>of course</i>	194
8.2.5 Connective implicatures of <i>of course</i>	198
8.2.6 Summary	200
8.3 <i>Of course</i> in PDE	201
8.3.1 Hedge <i>of course</i>	202
8.3.1.1 Concession / contrast	204
8.3.1.2 Background	207
8.3.1.3 Topic transition	209
8.3.1.4 Topic introduction	211
8.3.1.5 Irony	216
8.3.2 Emphatic <i>of course</i>	217
8.3.3 Distribution of <i>of course</i>	218
8.4 Summary	220
9. From rhetorical pattern to polysemy	223
Appendix A. Definitions of some rhetorical relations	227
Appendix B. Data sources	231
References	234

Key to symbols and abbreviations used in examples

Symbols used in examples

	discourse segment boundary
<u> </u>	[underline]host of adverbial expression
^ ^	overlap of speakers
[...]	omission in Conversation
..	replaces comma in Conversation transcription
...	replaces full stop in Conversation transcription; omission in written examples
(***)	unclear words not transcribed
<p>	new paragraph
/	end of verse line
<i>italics</i>	target expression

Abbreviations

Conv	Conversation corpus (PDE)
Rev	Reviews corpus (PDE)
Per	Periodicals corpus (PDE)
Sp	Political speeches corpus (PDE)
Ceeecs	Corpus of Early English correspondence sampler
EMED	Early Modern English dictionaries
Helsinki	Helsinki corpus
Hist	Historical texts
Lampeter	Lampeter corpus
MED	Middle English Dictionary
MEMEM	Michigan Early Modern English materials
Newdigate	Newdigate corpus
OED	Oxford English Dictionary, 2 nd edition

INTRODUCTION

BESIDES Words, which are names of *Ideas* in the Mind, there are a great many others that are made use of, to signify the *connexion* that the Mind gives to *Ideas*, or *Propositions*, one with another. The Mind, in communicating its thoughts to others, does not only need signs of the *Ideas* it has then before it, but others also, to shew or intimate some particular action of its own, at that time, relating to those *Ideas*. ... the Mind does, in declaring its Sentiments to others, connect, not only the parts of Propositions, but whole Sentences one to another, with their several Relations and Dependencies, to make a coherent Discourse.

The Words, whereby it signifies what connection it gives to the several Affirmations and Negations, that it unites in one continued Reasoning or Narration, are generally call'd *Particles*... To think well, it is not enough, that a Man has *Ideas* clear and distinct in his Thoughts, nor that he observes the agreement, or disagreement of some of them; but he must think in train, and observe the dependence of his Thoughts and Reasonings, one upon another: And to express well such methodical and rational Thoughts, he must have words to *shew* what *Connexion*, *Restriction*, *Distinction*, *Opposition*, *Emphasis*, etc. he gives to each respective *part of his Discourse*.

This part of Grammar has been, perhaps, as much neglected, as some others over-diligently cultivated. 'Tis easy for Men to write, one after another, of *Cases* and *Genders*, *Moods* and *Tenses*, *Gerunds* and *Supines*: ... yet he who would shew the right use of Particles, and what significancy and force they have, must take a little more pains, enter into his own Thoughts, and observe nicely the several Postures of his Mind in discoursing. ... They are all *marks of some Action*, or *Intimation of the Mind*; and therefore to understand them rightly, the several views, postures, stands, turns, limitations, and exceptions, and several other Thoughts of the Mind, for which we have either none, or very deficient Names, are diligently to be studied. Of these, there are a great variety, much exceeding the number of Particles, that most Languages have, to express them by: and therefore it is not to be wondred, that most of these Particles have divers, and sometimes almost opposite significations."

John Locke (1692/1975: 471-2)

Lexical discourse connectives, that is, expressions which link two or more discourse segments, are a sub-set of discourse markers with metatextual, discoursal and rhetorical functions. They are almost always polysemous forms, the connective use often having developed out of other adverbial uses. This study examines the thesis that some epistemic modal adverbials in English have developed or are developing discourse connective functions as a result of their occurrence in particular rhetorical contexts. It tests the hypotheses that there is variation in the use of certain modal adverbial expressions across text types in present-day English, that these adverbial expressions have distinct discourse-level functions that arise from rhetorical patterns, and that present-day English usage of these expressions reflects a grammaticalization-like process of unidirectional change. The four modal adverbs chosen for detailed analysis are *after all*, *in fact*, *at least* and *of course*. Looking at the frequency and distribution of the senses of these forms suggests how best to account for their role in textual coherence. By

bringing together grammaticalization theory and rhetorical structure theory (RST), the study illustrates the process of context-induced reinterpretation at the level of discourse structure. The research is corpus-based and combines quantitative and qualitative analysis.

The choice of expressions

The four expressions have been chosen to exemplify non-*-ly* adverbs which can be used either in VP modification or as discourse-marking sentence adverbs. They have undergone meaning extension and/or reanalysis. Although they are not prototypical discourse connectives, they seem to play an important role in text structuring. The aim of the analysis is to examine their contexts of use in search of clues to how and why they have come to have this function. Each of the forms is used in present-day English (PDE) to express both propositional content (examples a. below) and metatextual comment (examples b. below), as illustrated by (1-4).

- (1) a. I'm having a little rest *after all the hard work* (Per, A5M)
b. The Emperor would have realised this .. He had, *after all*, used this very route ... (Per, EE1)
- (2) a. All that folklore ... would appear to have some basis *in fact*. (Per, ABP)
b. There are quite a few other fish too; *in fact* there are enough different species to make a show that would rival the daily display at Harrods. (Per, A5U)
- (3) a. So far, *at least* two cabinet ministers have confirmed acceptance ..(Per ,A5M)
b. No coach, no problem -- *at least* it isn't if you follow our training plan .. (Per,AR7)
- (4) a. those whose ideas are the object of censorship risk death and imprisonment as a matter *of course* (Per, BMH)]
b. You could, *of course*, have the cement .. delivered and hire a mixer, but you would still find it hard work (Per, CCX)

The majority of English adverbs are *-ly* derivatives of adjectives, and it is unsurprisingly the *-ly* adverbs that have been most studied (e.g. Swan 1988). Non-*ly* adverbs, especially phrasal ones, have been less studied. It is sometimes assumed that adverbs are a sub-class of adjectives (e.g. Radford 1997: 62) and that the non-derivational adverbial forms are quirks. However, in a study embracing 41 European languages, Ramat and Ricca (1998) report that while around only 20% of sentence adverbs in their sample are non-derivationally formed, epistemic modal adverbs display a

special such tendency (1998: 248 and 225). For instance, all but one of the 41 languages have a non-derivational form for ‘perhaps’ (1998: 234). Certainly many English modal adverbs are lexicalizations rather than *-ly* derivations: for example, *maybe*, *perhaps*, *indeed*, *no doubt*.

The types of expression in the b. examples above have also been studied under the rubric of discourse marking, although such studies have often been limited to identifying, on the basis of textual analysis, the pragmatic function served in certain contexts. Discourse marker studies have tended not to examine how the expression has come to be used for that purpose or whether it fits any pattern in common with any other expressions. Indeed, many discourse marker studies either ignore other uses of the same form or explicitly treat them as unrelated homonyms (e.g. Fraser 1990, 1998, Jucker 1993, 1997). However, if the diachronic perspective is kept in mind, it can be appreciated that uses of a form may cluster at points along a cline of ‘ad-hoc contextual variation — polysemy — homonymy’. This study aims to show how these clusters may arise out of recurrent discursive and rhetorical patterns.

The approach

What discourse pragmatic issues are relevant to the study of discourse connectivity? How should the study of discourse connectives be approached? According to Chafe:

As soon as one looks beyond sentences, one finds oneself forced to stop dealing with artificial data concocted to suit one’s purposes, and to look instead at *language in use*. For it is, to say the least, awkward to try to invent whole discourses ... My guess is that much of what passes for syntax today will be explained in functional-discourse terms tomorrow. Most of the rest is likely to yield to *historical explanations*: the shift from functional to grammatical roles that often befalls *patterns in frequent use*.

(Chafe 1990: 20-21, emphasis added)

This passage highlights three strategies that need to be adopted in order to explain discourse-level phenomena.

First, discourse phenomena can only be adequately understood by looking at *language in use*, to see how people manipulate the linguistic resources at their disposal. For the analysis of actual occurrences of the linguistic objects of study, corpus linguistics provides a method for acquiring the

necessary data.

Second, *historical explanations* can help to account for current use: synchronic analysis should be supplemented by diachronic analysis. If the target of analysis is the current linguistic competence of an individual speaker, as is the case in most generative linguistics, the diachronic aspect is irrelevant, since it is not a part of that competence. But to account for how a speaker's language changes, and how the linguistic practise of the speaker at t_1 might influence that at t_2 (and thereby explain, in part, how the current linguistic competence of the speaker came to be as it is) it is useful to look at historical data. As Bybee argues, "synchronic states must be understood in terms of the set of factors that create them" (1988: 351). Sweetser argues that synchronic polysemy, semantic change and pragmatic ambiguity can all be shown to be motivated by the nature of human cognition (1990: 11-13). Important generalizations about language may therefore be missed unless the synchronic and diachronic data are viewed together.

Third, the role of *patterns in frequent use* warrants investigation. Repetition of a linguistic sequence (e.g. *sort of*) or construction (e.g. *be about to*) favours increasing automatization; that is, it encourages increasingly holistic rather than compositional mental representations. The more holistic representations in turn lead to loss of semantic specificity for individual components of the sequence or construction (v. Tabor 1993, Bybee in press). Frames (Fillmore 1976, Fillmore & Atkins 1992) and schemas are conceptual entities, while constructions are their linguistic correlates. Individual expressions gain part of their meaning from the constructions they enter into. These constructions can emerge from repeated rhetorical patterns which are ultimately socially-motivated conventions. A construction is thus a conventionalization of a typical, frequent mapping from conceptual level to linguistic level. Pragmatic systems, according to Du Bois, rely on "recurrent, stable, structured, and extremely high-frequency language functions" (1994: 3259). It will be argued here that repetition of rhetorical schemata (e.g. concession + preferred argument) may have a similar effect at discourse level to that of constructions at clausal level.

Overview

Overall it is argued that conventional rhetorical patterns or schemata acquire typical linguistic means of expression. These become more or less stable discourse patterns, constructions or idioms, which in turn lead to the repeated generation of implicatures which stabilize into ‘utterance type meanings’ and may eventually semanticize. Hitherto, more attention has been paid to constructions and idioms than to multi-clausal discourse patterns. This study argues that at discourse level, very similar processes are at work.

‘Discourse marker’ (DM) and ‘discourse connective’ (DC) are argued here to be pragmatic categories. The distinction between sentence-meaning and utterance-meaning (v. Lyons 1995) has long been the basis for the distinction between semantics and pragmatics. The boundary, however, is fluid or uncertain. This uncertainty is demonstrated by the difficulties that both semantic and pragmatic theories have in accounting for phenomena such as discourse connectives. For both disciplines, they are inconvenient. From a semanticist’s viewpoint, they resist truth-conditional and compositional analysis, and are better left to pragmatics. From the Gricean pragmaticist’s perspective, since their implicatures are not defeasible, they require the setting up of an intermediate category of ‘conventional implicature’ — almost an oxymoron. While a great deal of research in pragmatics has focused on conversational implicature, conventional implicature has received less attention. Areas of language such as discourse connectives have thus tended to fall between two stools, as is reflected in the great profusion of terminology that surrounds them, and in the absence of any widely-accepted definitions.

To address this predicament, Levinson (1995, 2000) proposes a third level of meaning. Levinson argues that the sentence-meaning vs utterance-meaning distinction fails to account for “the regularity, recurrence and systematicity of many kinds of pragmatic inferences” (1995: 93). He proposes an additional ‘utterance-type meaning’ to bridge the gap between semantics and pragmatics and to deal with that “great body of language lore, .. beyond knowledge of grammar and semantics, extensively studied of course by both ethnographers of speaking and students of second-language learning” (1995: 94). Levinson’s main concern is stable but defeasible generalised conversational implicature. But relevant to our study of emergent connectives is his proposal to remove the semantic-

pragmatic boundary and replace it with a continuum along which utterance-type meaning “will shade into speaker-meaning at the one end and sentence-meaning at the other. This is in part because there is plenty of evidence that language use is the source for grammaticalised patterns, and that there is a diachronic path from speaker-meanings to utterance-type-meanings to sentence-meanings.” (1995: 95). Our aim here will be to show that default inferences or patterns of preferred interpretation are particularly important for understanding how discourse coherence relations are understood and how patterns of use create this utterance-type meaning at discourse level.

In PDE, forms used as DMs/DCs are nearly always polysemous, a fact which begs an explanation. The view is taken here that relatedness among tokens of a form is a matter of degree: from transparent through semi-transparent to opaque relations. This gives a cline of polysemy as follows: ‘contextually modulated interpretations — polysemy — homonymic senses’. It is related to Levinson’s cline of ‘speaker meaning — utterance type meaning — sentence meaning’ in so far as contextually-modulated distinctions are ad hoc, like speaker meanings, while homonymic relations are encoded, like sentence meanings.

The synchronic polysemy can be seen as the result of diachronic layering of meanings, new uses having come in alongside older ones. It can be shown that modal S-AdvS are often reanalyses of VP-AdvS, a development that is well attested for English and cross-linguistically. There is no consensus on whether the diachronic meaning shifts take place by metaphorical extension or by conventionalization of implicature (via context-induced reinterpretation). It will be argued that the latter mechanism is responsible; that discourse connective meanings are not metaphorical uses of a content-level expression, but rather develop gradually by crystallization of implicatures. New constructions embody meaning which arises out of conversational implicatures generated in context. Thus, constructions and polysemies often develop in parallel, and may be inseparable.

We can draw correspondences between rhetorical patterns and linguistic patterns. For example, ‘concession + preferred argument’ is often realized by ‘expression of certainty + p + expression of contrast + p’; ‘false claim + corrected claim’ is often realized by ‘evidential + p + expression of certainty + p’, where p is a proposition. The preferred information structure is where p

has predicate structure (e.g. is a main clause) and the comment is contained in an expression attached to it (e.g. an adjunct such as a DM). But for focus or emphasis reasons or language-change reasons, this can vary. P can be realized as a subordinate clause, for instance, with an evidential as main clause (a so-called ‘transparent predicate’). It will be argued that ultimately, discourse connective functions of expressions of certainty and evidentiality arise from information structuring patterns imposed by rhetorical purposes.

Levinson notes that previous attempts to investigate the influence of language function on linguistic form have failed to clearly identify a set of language functions (1983: 40-43). That is, the categories set up by pragmatic theories such as Gricean maxims, speech act theory, traditional lists of language functions such as those of Bühler (1934), Jakobson (1956/1990), or Halliday (1994) have dubious empirical motivation. One solution, according to Levinson, might be to investigate cross-linguistically what notions are grammaticalized in which languages, and posit those notions as key functions. It would then be possible to identify for a given language the different forms of expression of those functions. A most interesting step in this direction is Anderson’s (1986) attempt to motivate and map a universal semantic space of evidentiality. Motivating and mapping out a rhetorical relations space has yet to be successfully undertaken (see ch. 2) and is beyond the scope of this study. Relations will be identified using Rhetorical Structure Theory (Mann and Thompson 1987), which provides a guide to defining and categorizing relations in English and is based on the developers’ experience of analysing large numbers of texts.

Previous approaches to the study of discourse connectives are outlined in chapter 1. Previous work shows multiple ways of classifying lexical items with a discourse connective function. This is partly because discourse connectivity, or the coherence relation, can be realized by a range of grammatical categories and by expressions that also have non-connective functions. The term ‘discourse connective’ is argued to refer to a pragmatic category, typically realized by sentence adverbials. The idea that different parts of utterances reflect different conceptual ‘planes’ expressed through different linguistic levels is explored in relation to modal sentence adverbials.

A major aim is to discover what discourse purposes the expressions serve. The function of

discourse connectives is to signal the type of coherence relation that the speaker wishes to establish between two discourse segments. Analyses of connectives are therefore dependent on, and contribute to, a theory of discourse coherence. Chapter 2 looks at the nature of discourse coherence and at several proposals for taxonomies of coherence relations. The framework adopted for the description of rhetorical patterns, RST (Rhetorical Structure Theory), is described. Finally, the chapter looks at possible ways of characterizing lexical discourse connectives as resulting from grammaticalization.

Chapter 3 describes the corpus linguistic method, the sets of data used in the study, and the way the data were analysed. Corpus linguistics is a relatively new method for the exploration of traditional linguistic concerns. The quantitative information it provides has already proved useful in NLP and lexicography. Here, the corpus linguistic approach is applied to provide support for the claim that frequently-generated implicatures can imbue modal sentence adverbials with discourse connective functions.

Chapter 4 illustrates the tendency for English modal adverbials of certainty to show ‘layering’, as sentence adverbial uses branch off from verb phrase adverbs (VP-Adv). The new S-Adv are shown to co-occur with particular coherence relations. It illustrates this phenomenon with the *-ly* adverbs *clearly* and *surely*, which are claimed to be developing discourse coherence functions in PDE.

Four non-*ly* adverbials, *after all*, *in fact*, *at least* and *of course* are analysed in more detail in chapters 5-8. It is argued that they all evidence polysemy and that they have acquired or are acquiring connective functions. For each adverb, the current polysemies are described, some history is traced and the PDE situation in three text types is compared. In particular, the frequency and distribution of the expression across certain rhetorical contexts are examined for evidence of a rhetorical motivation for the development of connective functions.

Chapter 9 concludes that the development of the expressions investigated shares many characteristics with the types of change that have been grouped under the ‘grammaticalization’ label, in particular those of unidirectional semantic change.

CHAPTER 1

DISCOURSE CONNECTIVES

1.1 Defining ‘discourse marker’ and ‘discourse connective’	9
1.2 Previous approaches to discourse markers	15
1.2.1 Functional approaches	17
1.2.2 Formal approaches	22
1.3 Discourse markers and discourse planes	27
1.4 Syntax and Semantics of DCs	34
1.4.1 Syntactic category	34
1.4.2 Semantic classification of S-Advs	39
1.4.3 Sub-categorization of speaker-oriented S-Adv	41
1.4.4 Position of S-Adv	46
1.4.5 Polysemy	48
1.5 Summary	50

1.1 Defining ‘discourse marker’ and ‘discourse connective’

Some expressions¹ seem to have a metatextual² or commenting function. Examples are *initial Well, ...*, *Indeed, ..* or *Even so, ...*, followed by a comma or comma intonation and a clause. The following extracts from the British National Corpus (BNC) exemplify these expressions:

- (1) yeah³ .. we allow dogs in here ... *well* .. you’ve managed to get one in anyway (Conv)⁴
- (2) even where Darwin does mention natural selection .. he also mentions inheritance of acquired characteristics .. erm .. paralleling it ... *indeed* .. he goes on to say .. that some physical changes produced in the nerve cells .. or nerves which are habitually used .. can hardly be doubted. (Conv)
- (3) I knew it would be a grand affair. *Even so*, I don’t believe I was prepared for the facts. (Per)

¹ The term ‘expression’ is used in Crystal’s sense of an element or string of elements “treated as a unit for the purposes of analysis and discussion” (Crystal 1991: 130).

² The term ‘metalinguistic’ is often used to refer to the commenting function of a discourse marker on its ‘referential’ host. However, the term ‘metatextual’ will be preferred, on the grounds that ‘metalanguage’ is best reserved for its use, in opposition to ‘object-language’, to mean language used to describe language (Lyons 1977: 10).

³ *Yeah* is sometimes treated as a discourse marker (e.g. Jucker and Smith 1998: 179-182).

⁴ The code in brackets after each example identifies the source (see Appendix B).

Well in (1) links its host clause to the previous segment⁵ of the text (which is about whether dogs are welcome in shops). *Well* introduces a comment on or modulation of the previous segment. Here, it introduces a weakening of commitment (the implicature is that there is no hard and fast rule about the admission of dogs, or that not all dogs in all circumstances are necessarily allowed in) and so represents a ‘rhetorical retreat’. The retreat is reinforced by the final *anyway*. At the same time, the *well*-clause introduces partial evidence for the allowing of dogs. Finally, *well* acts as a ‘relevance hedge’ for the speaker, it being clear to both speaker and hearer from the context that the hearer’s dog is already inside.

Indeed in (2) also introduces a modulation of the previous segment, but, by contrast with *well* above, it signals a strengthening, or reinforcement, presenting the next idea as evidence for what has just been asserted.

In (3), *even so* makes explicit the contrast between two segments, and so explains the apparent contradiction between knowing what it (a ceremony) was going to be like and not being prepared for it. Again, it can be described as a hedge to neutralize the apparent non-coherence. Each of the expressions in the above examples both expresses a comment on its host clause and signals a particular type of link with the previous segment.

Most scholars distinguish different sets and sub-sets of ‘comment’ or metatextual expression by using two or more terms. Broadly, the items discussed in the literature serve to express one or more of the following: epistemic modality (including evidentiality), attitude or evaluation, hedge, relevance of one idea to another in the discourse or (sometimes) to the context. Relatively well-established category labels for comment expressions include discourse markers (e.g. Schiffrin 1987), evidentials (e.g. Chafe & Nichols 1986), focus particles (e.g. König 1991), hedges (e.g. Brown & Levinson 1987: 145-172), intensifiers (e.g. Bolinger 1972, Labov 1984), miratives (e.g. Delancey 1997), modal particles (e.g. Abraham 1981, 1991) and stance markers (e.g. Biber & Finegan 1988, 1989).

These categories overlap. Moreover, an expression may serve more than one function at once.

⁵ ‘Discourse segments’ are discussed in chapter 2. The term ‘segment’ refers to a unit of discourse. It corresponds roughly to an information unit and is frequently a clause.

Comment expressions are often attributed an additional stylistic function. What they have in common is that some sort of judgment -- of likelihood, desirability, relevance⁶, etc. -- is passed on an idea or on a speech act. That is, they express subjectivity.

In the examples in (1)-(3) above, *well*, *indeed* and *even so* all indicate the relevance of the upcoming idea to the previous one. *Well* reduces commitment to the previous assertion. *Indeed* does just the opposite: it indicates that the upcoming idea enhances the previous one. *Even so* acknowledges the contrast between knowing and not being prepared.

The divergence in previous work stems partly from the fact that different scholars have been interested in accounting for different types of data and have imposed differing and overlapping categories. And it is partly because comments are realized by a range of grammatical categories and by expressions that also have other functions, as Locke observed, and are hard to match cross-linguistically⁷. In English linguistics, the category of discourse marker (DM) is perhaps the best established. Yet this category is ill-defined and has been addressed under a variety of labels. While the “fashionable” term (Östman 1995: 95) seems to be ‘discourse marker’ (e.g. Schiffrin 1987, Fraser 1988, 1990, Brinton 1990, Traugott 1995a, 1995b, Aijmer 1996a), other terms include: ‘discourse particles’ (e.g. Levinson 1983: 162), ‘discourse operators’ (e.g. Redeker 1990, 1991; Polanyi 1995), ‘discourse deictics’ (Levinson 1983), ‘pragmatic particles’ (e.g. Östman 1995), ‘pragmatic operators’ (e.g. Ariel 1994), ‘pragmatic markers’ (e.g. Brinton 1996), ‘rhetorical markers’ (Scott and de Souza 1990) and ‘rhetorical adverbs’ (e.g. Aarts 1996). Where the emphasis has been on the function of signalling rhetorical relations between discourse segments, they have often been termed ‘connectives’ of some sort: ‘pragmatic connectives’ (e.g. van Dijk 1979, Moeschler 1989), ‘inference connectives’ (Blakemore 1987), ‘discourse connectives’ (e.g. Warner 1985; Blakemore 1992), ‘phatic connectives’

⁶ ‘Relevance’ in its everyday sense.

⁷ Using a Swedish-English translation corpus, Aijmer (1996b) found that the Swedish discourse marker *nog* (which can express either emphasis or probability) was not translated at all in more than one third of cases. Where it was, it was rendered into English by twenty-three different means (excluding obvious mistranslations). While the term ‘discourse marker’ is commonly used in English linguistics, expressions such as Fr. ‘particules énonciatives’ or Ger. ‘Modalpartikeln’ are not quite co-terminous. Cross-linguistic comparisons are therefore difficult. Moreover, English is atypical in that markers tend to ‘float free’, against the cross-linguistic tendency for them to be “prosodically subordinate to another word” (Östman 1995: 99). See Spanoghe (1996) for a discussion of the problems of translation equivalence of both markers and linguistic terminology.

(Bazzanella 1990) or just ‘connectives’ (Mann, Matthiessen, & Thompson 1992). In natural language generation research, the terms ‘cue phrase’ (e.g. Grosz and Sidner 1986) and ‘discourse cue’ (e.g. Di Eugenio et al. 1997) are used, because such expressions “cue the hearer to some change in the discourse structure” (Grosz and Sidner 1986: 178). Sometimes it is the vaguer expressions that are frequent in informal conversation that are referred to as ‘particles’, without it always being clear whether this category is grammatical or discourse-pragmatic⁸. There may be a case for dividing the vaguer expressions (such as *well* and *so*) from the more specific (such as *frankly* and *allegedly*). But it will be assumed for the time being that all such expressions, when used with a discourse comment function, and whether in spoken or written language, can be described as DMs.

Interest in grammatical theory tends to be associated with a focus on speakers’ linguistic competence, or internalized language, and use of data provided by native speaker introspection. The apparently meaningless or vague expressions - semantically elusive and syntactically redundant - that are so frequent in spoken language, in literature and in highly context-bound texts tend not to feature prominently in introspected data. As ‘semantic leftovers’ (Du Bois 1994:3259), they have often been removed to the domain of pragmatics. The functions they have been shown to fulfil (e.g. Schiffrin 1987; Brinton 1996: 30-31), however, fall under the umbrella term ‘metatextual comment’, for which formal language and written language exhibit equivalent types of linguistic device, such as *unfortunately* or *clearly*.

The difficulties of providing a structural definition, together with issues arising from particular data sets, mean that DM is a controversial and fuzzy category. Brinton (1996: 30-31) reviews more than a dozen different definitions which she groups into five types on the basis of what they highlight as the key function of DMs. These five functions can be glossed as: (i) to express the relation between

⁸ It is not always clear whether ‘particle’ is to be taken as a syntactic category or a pragmatic category. Traditionally, a ‘particle’ is a word which, like prepositions, is a member of a closed class and has a largely grammatical function, whereas comment expressions are more open and lexical. However, as Crystal points out, the word ‘particle’ is often used to refer to an invariable item “which does not readily fit into a standard classification of parts of speech” (1991: 251-2). Indeed, Trask defines ‘particle’ as “a label typically applied to some more-or-less well-defined class of uninflected words in the grammar of some particular language when no more obvious label presents itself” (1993: 201). Hartmann (1998) emphasizes the lack of a standardized terminology for dealing with uninflecting expressions, citing widely divergent definitions of ‘particle’. His own characterization of the class of particles, using a mixture of syntactic and semantic criteria, is valid for German but not for English. We shall largely avoid the term ‘particle’.

two utterances, or between an utterance and the context; (ii) to structure or demarcate chunks of discourse; (iii) to signal a response to prior discourse; (iv) to achieve conversational continuity (acting as ‘fillers’); (v) to carry attitudinal, interpersonal meanings. Brinton goes on (1996:36-38) to propose an inventory of nine different functions, some ‘textual’, some ‘interpersonal’. Jucker and Ziv (1998) list twelve different functions of discourse markers.

DMs are sometimes characterized negatively as having no propositional content and no grammatical function. (Fraser 1990, Brinton 1996, Schiffrin 1987, Aijmer 1996a, Blakemore 1992: 146). The emerging view of discourse markers is of expressions in limbo between grammatical and lexical, between sentence and discourse, and between semantics and pragmatics. There is especial confusion over whether ‘discourse marker’ is a separate syntactic category, a sub-category of the category Adverb, or a pragmatic category typically realised by adverbials (v 1.4.1 below). Most studies of discourse markers do not attempt to provide an unambiguous criterial definition, but rather refer to typical functions realized by DMs and treat it in effect as a pre-theoretical category. Östman suggests pragmatic particles be defined as a prototype concept (1982: 150) and that they be identified as ‘core/prototypical’ or ‘peripheral’ (1982: 153, 1995: 99). In similar vein, Jucker and Ziv propose a “scalar conception of the membership in the class of discourse markers” (1998:3).

The view taken here is that ‘discourse marker’ refers not to a member of a closed-class set of linguistic forms, nor to a syntactic category, but to any lexicalised or semi-lexicalised expression when used as a comment on its host unit. In so far as it is a category, it is a pragmatic one. It is the function fulfilled in a particular instance that warrants the DM label, not the linguistic item used, nor its syntactic status. DMs act as comments on discourse segments, or as links between discourse segments, or both. Forms may range from the vaguer expressions typical of informal spoken language (*oh, you know, so, etc.*) to more specific words and phrases associated with either formal or written language (*nevertheless, in a nutshell, metaphorically speaking, etc.*).

For the present purposes (investigating indicators of coherence relations in English), three features from among those highlighted in the recent literature will be deemed central: (i) they are

metatextual (Aijmer 1996a: 206; Traugott 1999: 181); (ii) they defy analysis on the clausal level and are best explained with reference to the discourse (Brinton 1990: 45; see also Schiffrin 1987: 35); (iii) they may have scope over not only sentential but sub-sentential and supra-sentential units (Hansen 1998b: 236).

A discourse marker is defined here as:

an element that, in a particular utterance, conveys speaker comment on its host discourse segment.

Following Redeker (1991), this definition incorporates a structural constraint, via the expression ‘discourse segment’ (v. chapter 2), which by definition is no larger than one clause complex, i.e. one syntactically autonomous unit. A discourse marker is parasitic on its host discourse segment and is usually realised as an adverbial or a parenthetical.⁹

Syntactically autonomous expressions will therefore be ineligible for discourse marker status. (For example, *That’s clear* in *It’s rained. That’s clear* is an autonomous expression, whereas *clearly* in *It’s rained, clearly* is not¹⁰.) Clause-length ‘operators’ are not discourse markers by this definition. It will be argued that discourse marking expressions are a product of information structuring: that both their metatextual function and their syntactic status are the result of cognitive information-packaging processes that apply cross-linguistically.

The term ‘discourse connective’ (DC) will be used to refer to a comment expression that

⁹ Parentheticals are clausal or sub-clausal expressions (fragments) that are interpolated into or juxtaposed to a clausal or sub-clausal host with which they have no apparent syntactic link. A parenthetical differs from an appositive in that the latter is referentially equivalent to its host. Parentheticals typically express a comment on the host, and are sometimes regarded as a functionally-defined rather than syntactically-defined class (e.g. Ziv 1985: 191). They are typically marked off in English by comma intonation (Espinal 1991: 734-5, Peterson 1998: 231) but there are many exceptions to this (Reinhart 1983). “Parentheticals present formidable difficulties of analysis” according to Trask (1993: 199). As Cinque puts it, a “potential source of unexpected orders is the ‘parenthetical’ usage of (typically ‘higher’) adverbs, something for which I have no interesting account to propose” (Cinque 1999: 32). Some analyses have host and parenthetical belonging to a superordinate S-level syntactic unit (e.g. McCawley 1982: 103). Others claim that there is no structural relationship between host and parenthetical. Parenthetical adverbials do not display the same syntactic properties as adjunct adverbials (Espinal 1991: 729-32). Yet there are constraints on the linear positions a parenthetical can occur in with regard to the host. For Peterson (1998), there is discourse unity only, and these constraints are merely a processing phenomenon. For Espinal (1991), parentheticals are likewise independent of the host at every syntactic level, do not modify their host nor enter into any scope relation with their host. She proposes a three-dimensional syntactic model to allow for an independent root node for the parenthetical within the host, and a new relation of COMMENT to account for utterance interpretation. An interesting feature of the adverbial expressions discussed in chapters 4-8 is that they are sometimes parenthetical and sometimes not.

¹⁰ The notion of syntactic autonomy is less clear-cut than this example suggests: *Clearly* may constitute a whole utterance, for instance, and thereby be arguably autonomous. And the boundary between the so-called ‘transparent predicates’ (v. note 20 below) and syntactically dependent adverbials is sometimes blurred.

signals a rhetorical relationship between two or more discourse segments (at whatever level). Roughly, if the comment expression and its host cannot begin a discourse without obvious incoherence, the comment expression is connective. In Rhetorical Structure Theory terms (v. chapter 2), a connective expresses a rhetorical relation between two text spans. Discourse connectives are thus a sub-set of discourse markers.¹¹

Discourse markers express speaker viewpoint and are an inalienable part of all linguistic communication, written or spoken. Bréal calls attention to the importance of

*particules ... qui servent à nuancer les impressions ou les intentions des interlocuteurs. On peut les comparer à des gestes faits en passant ou à des regards d'intelligence jetés du côté de l'auditeur. ... La trame du langage est continuellement brodée de ces mots.*¹² (1897: 255-258)

But they have received scant attention from linguists. They have been described as “linguistic Cinderellas” (Enkvist 1972: 95). Perhaps this is because they appear to be “grammatically optional and semantically or functionally unmotivated” (Brinton 1996: 1) and therefore have an uncertain status in approaches that focus on the mappings between linguistic form and propositional content.

The next section reviews previous approaches to discourse markers. Section 1.3 examines the notion that comment expressions operate on a different discourse plane from their host. Section 1.4 looks briefly at the syntax and semantics of discourse markers, and considers their polysemy.

1.2 Previous approaches to discourse markers

The approaches to discourse markers outlined here share a broad view of a category of discourse-level expressions that act as comments on or as joints between discourse segments. They differ in their theoretical framework, choice of linguistic model and treatment of data. Previous studies

¹¹ There has been a recent tendency to separate ‘interpersonal’ markers of speaker stance from markers of relationships between idea units. Traugott (1999: 180-181) terms the former ‘pragmatic markers’ and the latter ‘discourse markers’. However, there is much multifunctionality and overlap, so that it seems preferable to include all expressions of speaker comment under the umbrella term ‘discourse marker’ which, as pointed out above, is now the best-established term.

¹² ‘particles ... which serve to express the impressions or the intentions of the interlocutors. They may be compared to accompanying gestures or meaningful glances directed at the hearer. ... The fabric of language is embroidered all over with these words’.

can be divided roughly into (i) semasiological approaches, which take lexico-grammatical elements as the starting point and analyse their role in building discourse structure and (ii) onomasiological approaches, which take coherence relations as the starting point and see how they are indicated (or not) by linguistic means.¹³ This study draws mainly on the first approach: it examines some lexical items which may be beginning to be used connectively, but uses a coherence relations framework (chapter 2) designed under the second approach to analyse the contexts in which the forms occur.

There have been many small-scale studies of individual expressions used as discourse markers in English¹⁴. Among the more colloquial expressions studied are *well* (Owen 1981, Schourup 1983, Watts 1989, Jucker 1993, 1997), *like* (Schourup 1983, Romaine and Lange 1991, Miller and Weinert 1998: 306-334, Jucker and Smith 1998: 182-191), *because/ 'cos* (Stenström and Andersen 1996, Stenström 1998), *actually* (Aijmer 1986, Watts 1988), *really* (Watts 1988), *anyway* (Altenberg 1986, Lenk 1995, Ferrara 1997, Takahara 1998), *even* (Kay 1990, Kalokerinos 1995), *you know* (Schourup 1983, Jucker and Smith 1998: 191-197), *I think* (Thompson and Mulac 1991, Stenström 1995, Aijmer 1997). Expressions not so much associated with conversation have been less well studied. They include: *in fact* (Traugott 1999, Schwenter and Traugott 2000, Lewis 1999: 365-367), *after all* (Blakemore 1987: 81-84), *then* (Schiffrin 1992).

This section looks first at some major studies of DMs within the functionalist tradition, illustrating the points of view of conversation analysis, discourse pragmatics, register variation and historical linguistics. Second, work in more formal pragmatics, including Relevance theory, is described.

¹³ On the semasiological/onomasiological distinction, v. Anttila 1972/1989; Geeraerts 1983; Traugott 1999. 'Semasiology' was until recently synonymous with 'semantics' (Ullman 1962: 142); on the origins of semasiology in the 1820s, see Ullmann (1964: 30). For an overview of NLP approaches to discourse markers, many of which combine the semasiological and onomasiological by, alternately, using extant markers to delimit and segment the discourse marking space, and then categorizing the markers accordingly, see Oates (1999).

¹⁴ Only studies of English discourse markers are described, for reasons of space, although cross-linguistic and comparative studies have also contributed to our understanding of English markers, e.g. Anderson (1986), Willet (1988), Aijmer (1996b), Spanoghe (1996). Analyses of similar phenomena in other languages, where the data may have suggested a different approach to categorization, are also relevant. Major studies using discourse-pragmatic approaches similar to those adopted for English discourse markers include Kroon (1996) on Latin, Hansen (1995, 1997, 1998a) and Nølke (1989) on French. There is also a tradition of study of French DMs based on argumentation theory (Anscombre & Ducrot 1983) which focuses on the rhetorical purposes of discourse. A semantico-syntactic approach is taken to German modal particles by Abraham (1991). Smaller studies or analyses of particular expressions include Risselada (1998) and Langslow (1998) on Latin. Relevance theory is used for several studies of Japanese markers, e.g. Tanaka (1997) and for studies of markers in several languages in Jucker and Ziv (eds.) (1998).

1.2.1 Functional approaches

Two features of functional studies, perhaps motivated by practical considerations, are: first, they tend to be semasiological rather than onomasiological (that is, they look at particular forms to see what functions they serve, rather than at particular functions to see how these are realized linguistically¹⁵); second, they tend to focus on spoken language.

The focus on spoken language stems from interest in discourse function often coinciding with interest in conversation data, so that the term ‘discourse marker’ has tended to be associated with those expressions which are characteristic of spontaneous spoken language (such as *Well,...* or *then* or *you know*). Indeed, markers are sometimes described as an exclusively oral phenomenon. Chafe describes ‘emphatic particles’ expressing speaker ‘involvement’ as being diagnostic of spoken language (1982: 47). Markers are best seen as “a feature of oral discourse” according to Brinton (1990: 49). Middle English markers are “oral residue”, while those found in present-day written English are “oral devices ... transferred into written texts” (Brinton 1996: 4). Discourse markers are said to be “peculiar to spoken language” (Stenström 1990); they “occur exclusively in spoken language” (Hartmann 1998: 662). Many studies (e.g. Schiffrin 1987, Aijmer 1986, 1996a) are based entirely on spoken conversational English data.

The functional approach implies no commitment to truth-conditional semantics. It tends to define a DM as a member of a set (in principle a closed set) of expressions which have one or more of a particular range of pragmatic functions. However, it is “difficult to differentiate pragmatically motivated from non-pragmatically motivated occurrences” (Brinton 1996: 35) of these expressions, and the criteria an expression must meet to gain membership of the set are often left vague.

Schiffrin’s (1987) study, inspired by conversation analysis, is perhaps the best known treatment of English discourse markers and the group of markers that she chose for analysis has come for many to epitomize ‘discourse marking’ in English conversation.

¹⁵ The semasiological approach is more practical in that, given current computational tools, quantitative studies can reasonably be made of forms but not functions, and in that current taxonomies of discourse marking functions are in any case based on sets of features derived from analysis of markers (e.g. Knott & Mellish 1996, Fraser 1996).

Schiffrin takes the view that discourse structure cannot be approached in the same way as sentence structure. Her corpus-based study of a range of colloquial DMs is highly inductive and atheoretical, a limited theoretical framework resulting from the study itself. DMs are “sequentially dependent elements which bracket units of talk” (1987: 31). They signal relations between units of talk; that is, they ‘bind’ discourse units. They mark structural transitions between different parts of a discourse, for instance, and they convey speakers’ attitudes towards the discourse (1987: 13).

However, the criteria for delimiting the set of DMs are left vague. Their function is to “propose the contextual coordinates within which an utterance is produced and designed to be interpreted” (1987: 315): markers “select a meaning relation from whatever potential meanings are provided through the content of talk, and then **display** that relation [original emphasis]”, but they do not create the relation (1987: 318-9). Not only are DMs themselves non-propositional, but they often serve as links between other non-propositional entities: Schiffrin cites the use of *because* to signal the reason for a speech act rather than the reason for the proposition expressed in that speech act. (Her example is: ‘Is that your newspaper? *because* I haven’t seen the headlines yet today’.)

Perhaps the main limitation of Schiffrin’s study is that it focuses on a small yet disparate collection of high frequency expressions that occur at discourse unit boundaries in the data. Schiffrin convincingly argues against a structural definition of markers (1987: 32) while acknowledging the dangers of her own ‘operational’ definition (1987: 31-7). In fact, the markers are too random and varied to allow interesting generalizations to be made and too few to allow large-scale patterns of use to be detected. Secondly, the semantics of DMs are inadequately explored. Schiffrin posits two types of DM: those that have semantic content (such as conjunctions) and those that do not (such as *oh* or *well*) (1987: 314, 319). Her insistence on a very minimal, core semantics for each expression, or on no semantics at all, has been seen as unable to explain the non-syntactic distributional differences observed for the expressions (Redeker 1991: 1164). A third weakness is that although the approach is exclusively semasiological, little attention is paid to the possibility that the expressions may be polysemous: the work of distinguishing among different functions of an expression is left entirely to an

account of discourse coherence. Overall, the theoretical framework is too insubstantial to prevent the study from being a collection of studies of individual items rather than the analysis of a category that the title, *Discourse markers*, anticipates. However, the study is valuable not only in its tackling an area of pragmatics that had received scant previous attention, but also for its use of a carefully designed and transcribed corpus of spoken language and for (albeit limited) quantitative as well as qualitative analyses of the uses of the expressions.

Most discourse marker studies adopt what can be called the discourse coherence approach. This term covers a range of discourse models which share the view that DMs signal discourse coherence relations (v. chapter 2). The approach is exemplified by Redeker (1990) and Aijmer (1986, 1996a chapter 5).

Redeker's approach echoes Schiffrin's concerns, but advocates a firmer theoretical basis in the shape of a model of discourse coherence that clearly delimits the class of discourse markers. Redeker's 'discourse operators' are "linguistic signals of textual-coherence links" (1991: 1139). They are defined as a "word or phrase ... that is uttered with the primary function of bringing to the listener's attention a particular kind of linkage of the upcoming utterance with the immediate discourse context" (1991: 1168). She invokes structural criteria in delimiting the class of discourse operators: as signals of discourse-segment boundaries, they are parentheticals, not syntactically or intonationally integrated into the utterance (1991: 1166). Conjunctions and adverbials are admitted, but "clausal indicators of discourse structure" are excluded (1991: 1169) on formal grounds. Moreover, Redeker excludes any expression whose scope does not exhaust the utterance (1991: 1169). The kinds of linkages available are ideational (linking the ideas expressed by the utterances), rhetorical (indicating the discourse structure) and sequential (temporal).

For Aijmer, discourse markers are one type of 'conversational routine' in English. These routines are formulaic expressions of one of three types: (1) formulaic speech acts of thanking, apologizing, etc., (2) discourse-organizing routines, and (3) attitudinal routines (Aijmer 1996a:2). Discourse markers are "routines with a discourse-organizing function" (Aijmer 1996a:20). Aijmer's

analysis of spoken English corpus data is based on the notion of ‘lexicalized sentence stems’:

“cognitive units which provide an economical way of storing linguistic variants which show some similarity, and .. can be used ready-made in speech production” (Aijmer 1996a:22).

Interest in register variation in English prompted Biber & Finegan (1988, 1989) to analyse the cross-register distribution of ‘stance adverbials’. Their (1988) cluster analysis of the distribution in two register-diversified corpora (the Lancaster-Oslo-Bergen corpus (LOB) and the London-Lund corpus (LLC)) of several dozen such adverbials finds that texts tend to conform to one of eight types with regard to stance marking. They find, further, that the discourse functions of stance marking adverbials often differ from their ‘literal’ meanings. Their 1989 study is of a wider range of stance markers, including verbs, adjectives, hedges, emphatics, as well as adverbs, and identifies six stance styles. The characterization of text types by feature clustering, pioneered by Biber (v. Biber 1988) is an extremely valuable tool made possible by large machine-readable corpora. At present it is a rather coarse implement, limited by the arbitrary choice of and categorization of input expressions¹⁶. Despite these limitations, the method outputs what are intuitively very plausible categories, and with some refining on the input side it has enormous potential for transforming language variation studies and stylistics.

A number of diachronic analyses of English discourse markers have been made in the functionalist tradition (e.g. König 1985, Brinton 1990, 1996, Nevalainen 1991, Powell 1992, Schwenter and Traugott 2000, Traugott 1995a, 1995b, 1997, 1998, in press, Stenström & Andersen 1996, Tabor & Traugott 1998). Brinton (1996) is a book-length treatment of the historical development in English of what she calls ‘pragmatic markers’. Brinton’s study is carried out within a Hallidayan functionalist framework that views language as operating on ideational, interpersonal and textual levels. She suggests that the functions of markers fall within the interpersonal and the textual functions of language. The criteria by which expressions can be classed as pragmatic markers are

¹⁶ The choice of expressions is partly determined by the need for them to be automatically identified by simple string-matching. The method is therefore not suitable for dealing with polysemous expressions. With more sophisticated syntactic and semantic tagging software some of these limitations will be overcome.

vague: markers are sometimes phonologically reduced, they are difficult to specify lexically, they have no clear grammatical function, they are often described as ‘optional’ and as having relatively low semantic or propositional content (Brinton 1996: 33-38). Broadly, however, they are seen as expressions operating at discourse level which tend to derive diachronically from the propositional component via the same semantic and syntactic principles that operate in grammaticalization (Brinton 1996: 26). Brinton establishes that initial position is overwhelmingly the most common for OE and ME markers. All are possible in this position and some are restricted to it. Brinton is especially alert to the rhetorical functions of the markers, and her data “calls into question the usual correlation between old or given information and backgrounding and new information and foregrounding” (1996: 270). She also finds support in her data for the idea that diachronic development of markers in English is unidirectional: textual and attitudinal meanings develop from content ones. She further finds that these meaning changes can be accounted for by conventionalization of conversational implicatures (1996: 275). Finally, she notes that the same form can develop differently in different discourse types, e.g. in narrative vs in dialogue. Brinton’s method of identifying pragmatic markers results in a motley collection, as is the group Brinton selects for her study, which range from the OE ‘introductory particle’ of indeterminate meaning *hwaet* to the ME verb *bifel*, which syntactically appears as a main clause verb.

The range of linguistic expressions embraced by the label ‘discourse marker’ (and comparable terms) thus varies greatly with each approach. In fact, as discussed above, the semasiological approach leads to some confusion over whether what is at issue is a set of expressions or a set of functions. In some studies discourse markers are described as having certain pragmatic functions, but the category is then treated as a closed or semi-closed set of linguistic forms. Östman refers to “a pseudo-open class” of pragmatic particles (1982: 154). Schiffrin views DMs as ‘a list of items’. She wonders whether they belong to a single word class, and asks: “is it possible to define so disparate a list of items in a way which will let us identify other elements as members of the same class?” (1987:40), a question which receives no answer. Aijmer sidesteps the issue of defining what is meant

by ‘discourse marker’, but also assumes that the term refers to a set of linguistic expressions: “it is difficult to make a list of discourse markers” (Aijmer 1996a:204). Aijmer notes the close connection between discourse markers and metalinguistic function, although “not all metalinguistic expressions are discourse markers” (1996a:206). However, she assumes that metalinguistic expressions are also particular linguistic items, so that “the problem remains of delimiting the set of ‘metalinguistic items’” (1996a:207). For Brinton, the object of study is “a diverse set of forms” (Brinton 1990: 45). Hansen is one of the few to emphasize the functional-pragmatic nature of discourse markers. As Hansen puts it, “when I speak of certain linguistic items as *being* discourse markers, it should be understood as not more than a convenient shorthand: ... I find it more appropriate to speak of items which in certain contexts, and in certain syntactic positions, may fulfil a discourse marking *function*” (1998b: 242). As mentioned above, DM being a pragmatic category, this latter view is preferred here. It allows clear distinctions to be maintained between a form, its polysemies, and its occurrences.

1.2.2 Formal approaches

For formalists, DMs are expressions which do not contribute to the truth conditions of the utterance to which they belong, but rather indicate how the utterance is to be construed. DMs are thus defined in relation to propositions. Their semantic contribution is that they act as a form of operator on the proposition. The lexemes acting as DMs are nonetheless seen as conventional signs with conventional meanings. Few formalists have focused on DMs, but rule-based NLP (natural language processing) research has naturally expanded into supra-sentential analyses of the type illustrated by Polanyi (1988, 1995). Formal pragmaticists such as Fraser (1988, 1990) believe that formal methods can profitably be incorporated into the analysis of discourse coherence where appropriate. Another, rather different, neo-Gricean formal approach is that of Relevance theory (Sperber and Wilson 1995), which has inspired very many recent analyses of individual English DMs. Blakemore (1988a, 1988b, 1992) illustrates this viewpoint.

Polanyi (1988, 1995) is concerned with the development of formal models of discourse that

can be used in natural language processing tasks and therefore needs to deal with naturally-occurring texts. Polanyi's aim is to set up a formal model of discourse structure analogous to models of sentence structure. Like Redeker, she adopts the term 'discourse operators', but as a wider category than Redeker's. The essential characteristic of discourse operators (DOs) is that they are non-propositional, in contrast to 'discourse constituent units' (dcu's), which carry propositional material. DOs either "give information about the relationship (both structural and semantic) obtaining among the dcu's as well as pragmatic information linking the discourse to its utterance situation" (1995: 42), or they are "attitudinal evaluations" by the speaker (1995: 15). In either case, they are to be "treated as clitics attached to the propositional hosts" (1995: 16). The relations among dcu's that are indicated by DOs may be logical, rhetorical or interactional (examples of interactional relations are question-answer pairs and error-repair sequences). In the category of 'discourse operators' Polanyi therefore includes connectors (such as *therefore* and logical connectives such as *or*), attitude markers (such as *actually* or *unfortunately*), assigners (eg. vocatives) and discourse 'push/pop' markers for embeddings (such as *By the way, ..* to introduce a digression, or *Anyway, ..* to resume a main topic). Metacommunicative propositional utterances are also treated as discourse operators. As the name implies, these are propositional in addition to being discourse operators. Thus, whole clauses, such as *So where were we?*, can function as discourse operators (Polanyi 1988: 605). This seems to be an attempt to distinguish between what is 'propositional' and what is merely 'propositionalized'. Expressions such as *yes*, *OK* and *hello* are also classified as discourse operators. This classification has arisen from a 'negative' (non-propositional) definition of DOs, and the resulting set of expressions is heterogeneous.

Fraser draws a sharp distinction between what he calls 'content meaning' (i.e. propositional content) and 'pragmatic meaning', the latter being conveyed through 'basic', 'commentary' and 'parallel' pragmatic markers which respectively signal the communicative force of the sentence (e.g. declarative, interrogative mood), speaker comment on the sentence (e.g. '*Frankly, ...*'), and separate but concomitant messages relating to the sentence (e.g. forms of address) (Fraser 1988: 19-20).

Discourse markers are defined as one type of commentary pragmatic marker. Fraser challenges the

idea that discourse markers have any ‘semantic content’ at all (1990: 394). He argues instead that DMs have a ‘core pragmatic meaning’, although “what constitutes this ‘core’ is elusive” (1988: 23). The core pragmatic meaning is separate from the ‘content meaning’ of any homophonous form (1988: 24). The DM class is restricted to those expressions that signal how the discourse segment they introduce relates to the foregoing discourse. Fraser is largely concerned with definitions and categories. He proposes an elaborate sub-categorization scheme for discourse markers within a classification scheme for pragmatic markers. He is not much concerned with data, so that his is apparently an intuitive categorization. However, the fine-grained classification he proposes has the advantage of clearly separating out expressions in a way that allows a single form to be entered under more than one type of marker and thereby highlights some overlaps and lexical relations among markers. Fraser finally suggests that “we may ultimately conclude that such an attempt at typing is not a useful way to understand discourse” (1988: 32).

Recently several studies of discourse markers have been carried out in the Relevance Theory framework. Examples for English are Blakemore (1987, 1992), dealing with discourse connectives such as *so*, *therefore* and *after all*; Jucker (1993, 1997), which look at *well* in its discourse marking uses; Watts (1988), comparing *actually*, *really* and *basically*; Takahara (1998), which analyses *anyway*; Jucker and Smith (1998), on particular uses of conversational expressions such as *yeah*, *like* and *you know*. Rouchota (1998) is a general defence of the Relevance-theoretic approach to adverbial discourse connectives such as *moreover* and *after all*.

According to Relevance Theory, hearers process input until they retrieve sufficient contextual effects to achieve a Relevant interpretation. For this, they need to identify the appropriate context, which may not be the current context. A distinction is drawn between conceptual meaning (roughly, communicated content) and procedural meaning (processing instructions). Discourse connectives, just

like focusing devices such as clefting, stress¹⁷, and intonation, are said to encode procedural meaning.

Blakemore is the Relevance theorist most associated with the study of discourse markers (Blakemore 1987, 1988b, 1992: 137ff, 1996). Blakemore first argues (1987) that because ‘inferential connectives’ such as *therefore* or *after all* “do not map onto concepts that are constituents of propositions” (1987: 143), they cannot be conceptual; they do not contribute to the logical form (LF) of their host. She therefore proposes a non-unitary semantics (1987:144): a conceptual semantics for the constituents of propositional representations and a procedural semantics for the rest. Wilson and Sperber (1993), however, argue that evaluative sentence adverbs such as *unfortunately*, although non-truth-conditional, are conceptual. This results in a four-way distinction of meaning type, and discourse markers are now held to be both non-truth-conditional and to encode procedural meaning, unlike evaluative sentence adverbs, which are non-propositional but conceptual. Rouchota (1998) cites data showing distributional and structural differences between *unfortunately*-type sentence adverbials and adverbial discourse markers as evidence of this conceptual/procedural distinction.

Discourse markers, according to Relevance Theory, are “linguistically specified devices whose contribution to the interpretation of the utterances that contain them ... must be analysed in terms of constraints on the relevance of the proposition” (Blakemore 1987: 141). They limit the hearer’s choice of contextual assumptions for the interpretation of the utterance (Blakemore 1988a: 186-8). This conclusion is therefore not very different from Schiffrin’s position, according to which DMs “index an utterance to the local contexts in which utterances are produced and in which they are to be interpreted” (1987: 326). Jucker’s analysis of *well* “differs from earlier analyses in that it uses Sperber and Wilson’s relevance theory as a descriptive framework” (1993: 437). Nonetheless, his conclusion that “the discourse marker *well* indicates that the addressee has to reconstruct the background against which he can process the upcoming utterance. ... [It] is used to indicate a shift in the relevant context”

¹⁷Blakemore notes that discourse marking expressions may often “be replaced by intonation for the same effect” and illustrates this point with the following example of *after all* (1987:142):

You will have to invite him. After all, he’s your brother
 You will have to invite him. He /is your /~brother
 (where / = foot boundary and ~ = fall-rise)

But it is also the case that the speaker-oriented meanings associated with discourse markers may be propositionalized for the same effect.

(1993: 438, 451) is again not very different from Schiffrin's, according to which *well* indicates that “an upcoming contribution is not fully consonant with prior coherence options” (1987: 102-3). So it is not clear that a Relevance Theory approach sheds any more light on discourse markers than the highly inductive, data-driven Conversation Analysis approach of Schiffrin (1987).

In the Relevance Theory approach, data on DMs do not serve to refine the theory, since the conceptual/procedural distinction is framed in such a way that DMs are procedural by definition. Indeed, in Relevance Theory analyses of DMs, the Relevance Theoretic framework is treated as given and the various uses of the DM suggested by the data are described in its terms. The conceptual/procedural distinction is said to be cognitive (Relevance Theory is a cognitive theory) so neurological or psycholinguistic evidence is needed to test it. In the absence of such evidence, ‘procedural’ is just another label. It corresponds quite closely to what Bellert (1977) on distributional grounds terms ‘conjunctive’, Parsons (1990) terms ‘speech-act modifying’ (v. section 1.4 below) and Sweetser (1990) ‘speech-act domain’ (v. section 1.3 below). Another difficulty is that, especially in English, virtually any kind of meaning can be propositionalized (v. Lyons 1995: 274), so that what appear to be equivalent meanings may presumably be conceptual when propositionalized and procedural when expressed by a DM. Yet Relevance Theory does not suggest why language might be so designed. Moreover, Relevance Theory provides no account for the pervasive polysemy of discourse marking expressions; propositional (content) uses of the same form are simply treated as homonyms. Nor does it provide an account for the diachronic development of DMs; that is, why it is that DMs are recruited from particular types of lexical item and not others. Nonetheless, Relevance Theory research has recently produced numerous studies of DMs, providing a valuable collection of data on some ubiquitous discourse marking expressions and doing much towards building taxonomies of senses.

Despite their very different approaches, sufficient overlap can be appreciated in the ranges and functions of the expressions discussed in the studies mentioned in this section to claim that the studies are focusing on broadly the same phenomenon. In most of the studies, what seems to warrant the treatment of these expressions as a group is their metatextual function. There may be some evidence

from code-switching studies that DMs are perceived by speakers to be a natural category (Maschler n.d., cited by Ariel 1994:3251), but much more empirical research is needed.

1.3 Discourse markers and discourse planes

Discourse markers/connectives are described above as metatextual, as well as non-analysable at clausal level and having scope over variable-sized units. This characterization depends crucially on the notion that there are different discourse planes. If an expression is used primarily or wholly to comment on, rather than contribute to, propositions referring to the external world, it can be said to operate on a different discourse plane. Discourse markers are such expressions: parasitic elements conveying speaker comment on the content of their host units. This section looks at a number of frameworks that have been proposed for describing these different planes of discourse.

The idea that there is more than one level of meaning is found across a range of very different approaches to language and reflects a long-standing distinction between attitude or modality on one hand and proposition on the other. In the traditional view, propositional modality “may be likened to a shell that encases [the clause] but does not tamper with the kernel inside” (Givón 1995: 112). The idea is present in Jespersen’s characterization of moods as expressing “certain attitudes of mind of the speaker towards the contents of the sentence, though in some cases the choice of a mood is determined not by the attitude of the actual speaker, but by the character of the clause itself and its relation to the main nexus on which it is dependent” (Jespersen 1924: 313). It is an idea that stems from a more general multifunctional view of language. This section briefly surveys functionalist and formalist proposals for formulating the discourse plane notion¹⁸.

Truth-conditional semantics has long faced the problem of what to do about expressions (including discourse markers) that resist truth-conditional analysis and so do not directly contribute to

¹⁸ Schiffrin’s (1987) model of discourse coherence comprising five ‘planes of talk’ is omitted from this discussion, because it was developed solely as an outcome of analysing a narrow range of expressions in a narrow corpus (Schiffrin 1987: 313). It is therefore very different from the kind of model discussed here. Schiffrin claims that different discourse markers function on different combinations of five planes. There are a number of problems and inconsistencies with this framework, however. In particular, the planes do not jointly account for any clearly defined conceptual space, but overlap and mix linguistic, conceptual and social aspects of conversation. And contra Schiffrin’s claim, the five-plane framework does not adequately distinguish among discourse markers (Redeker 1991: 1159).

propositional content (as expressed in logical form). Modal adverbial discourse markers are usually analysed as operators on propositions (Parsons 1990, Cann 1993: 272)¹⁹. They are said to contain predominantly or only presuppositional meaning (Chierchia & McConnell-Ginet 1992: 386) in contrast with denotational or representational meaning. These two types of meaning can be considered to be different ‘planes’. Discourse markers are therefore “words whose semantic contribution is chiefly presuppositional” and whose “lexical entries will involve not only meaning postulates but in many cases also something like presupposition schemata” (Chierchia & McConnell-Ginet 1992: 388). Haegeman (1985: 4) distinguishes ‘content’ and ‘epistemic’ uses of adverbial subordinators such as *while* and *since*, which can be used to express a real-world relation or a speaker-viewpoint one. Although recent work in formal syntax has addressed the problems of various types of modality (v. section 1.4.3 below), there seems not yet to be a fine-grained formal semantic analysis of discourse marking.

For formal pragmatists, the boundary between the propositional and the non-propositional is the limit of what Grice would term ‘saying in a favored sense’ (Grice 1989: 25). Grice, in discussing how a contrast between two speech acts is signalled, suggests that an expression such as *on the other hand* is a “higher-order speech act of commenting in a certain way on the lower-order speech acts” (Grice 1989: 362). The higher-order speech act relates lower-order speech acts about the external world to each other. Grice aptly terms these lower-order ones ‘ground-floor statements’. This same two-way distinction is reflected in Davison’s (1975) work on indirect speech acts. There are elements which are “in addition to what one might call the ‘core’ of the speech act, or the expression of the action of proposition involved” and which “do not contribute in any obvious way to the meaning of the sentence” (Davison 1975: 144). Davison notes similarities between indirect speech acts, speech-act

¹⁹ But there is also a ‘dual assertion’ analysis in which they are themselves mini-propositions. Bellert (1977) puts forward evidence that “sentences with modal adverbs express two propositions ... the occurrence of a modal adverb gives rise to an additional metalinguistic proposition” (1977: 345-6). V. also Davidson (1982) on the dual proposition analysis of modality.

modifying adverbial clauses (1975: 154n) and transparent predicates²⁰ such as *I'm afraid* or *I suppose* (1975: 157n): they express speaker attitude or signal some psychological state, and they seem to be “mysteriously absent from the structure at some levels of derivation” (1975: 158n). As has already been seen (1.2.2), Relevance theory posits a conceptual-procedural distinction to account for these levels. The distinction is intended to account for broadly the same idea/comment divide, ‘procedural’ referring to the speaker’s processing instructions to the hearer. By contrast with some other analyses, conceptual and procedural are defined not as two concurrent aspects of discourse, but as conveyed separately by distinct linguistic expression. All these approaches tend to take the ‘misfit’ view: they group together elements that will not fit into the ‘basic’ level of meaning, rather than formulate a unified framework for all meaning, and the result is an emphasis on what metatextual expression is not, rather than what it is.

In Functional Grammar (Dik et al 1990, Siewierska 1991) there is no qualitative distinction between the propositional and the non-propositional. Rather, discourse can be analysed in terms of a hierarchy of planes from simple predicate to illocution, each embedded in the higher one. Lexical means of modulating linguistic expressions are known as ‘satellites’ and non-lexical means (particles and inflections) are known as ‘operators’. It was “the behaviour of operators and satellites, more specifically the evident differences in the type of semantic contributions effected by different operators and satellites, that provided the initial motivation for developing the multi-layered model of the clause” (Siewierska 1991: 38). The four-layer hierarchy of Functional Grammar is summarized in table 1.1.

²⁰ ‘Transparent predicates’ are predicates which appear to be syntactically main clauses but which serve to introduce another clause that is pragmatically more important. Examples are *I suppose*, *I'm afraid*, *I think*, *I gather*, etc. Davison (1975) discusses this type of structure. Bolinger (1972) discusses this phenomenon in terms of ‘syntactic shift’, with the example ‘It’s a nuisance, I tell you’, where ‘I tell you’ is considered an intensifier. Bolinger claims that elements that become intensifiers are for the most part “from ‘outside’ the *dictum* .. they are terms that originally expressed some relationship between what is said and the declarativeness of saying it, or the certainty of emphasis or truth attached to it” (1972: 93). See also Hooper and Thompson (1973), who show that certain verbs are used ‘parenthetically’ with a first person subject, so that it is the complement that expresses the main assertion of the utterance. Thompson and Mulac (1991) analyse *I think* with *that*-deletion as a case of the grammaticalization of main subject and verb into epistemic adverb. Cross-linguistically there is evidence that comparable notions of modality and speaker attitude are expressed by sentence adverbials in some languages and by transparent predicate type constructions in others (Bisang 1998).

Level	Host layer	Example satellite types
4	illocution	modification of illocutionary force
3	proposition	speaker evaluation of proposition
2	predication	location / setting of state of affairs
1	predicate	additional properties of state of affairs

Table 1.1. Satellite qualifications in the layered clause structure of Functional Grammar
(based on Siewierska 1991 and Dik, Hengeveld, Vester & Vet 1990)

A satellite expressing a qualification attaches to one or other of these layers according to its semantic scope (Nuyts 1993b: 283). Discourse markers were described above (1.1) as conveying the speaker's evaluation of or comment on the host discourse segment. Satellites with this function have wide scope, and thus come towards the 'top' of the hierarchy: those that indicate speaker evaluation of the content of the host proposition (such as *allegedly*) attach at the level of propositions; those (such as *frankly*) that modify the illocutionary force of an utterance attach at the illocutionary level, i.e. the level of the speech act (Dik, Hengeveld, Vester, & Vet 1990: 28-9). Role and Reference grammar adopts a similar layered clause structure (van Valin and LaPolla 1997: 40-45). This layered representation originates in a view of modification according to which expressions can be ranked: one expression can modify another "only if the modifying expression is lower than, or equal to, the modified expression in terms of its rank" (Lyons 1977:439). In this view, the higher-ranking, or 'first-order' entities are more basic in that they refer to physical objects or things which 'exist'. 'Second-order' entities are events and processes, etc., or things which 'occur', while 'third-order' entities are abstract things such as propositions, which can be the object of propositional attitudes of belief and judgement, etc. (Lyons 1977: 442-445). It is this third type of entity that is modified by attitudinal discourse markers (cf. Lyons 1977: 452).

Some scholars suggest some form of bipartite analysis of text to deal with the idea/comment distinction. An example is Chafe's distinction between 'ideas' and 'non-idea information'. The latter "has such functions as the specification of the relations between ideas, as with conjunctions or prepositions, or the inflectional or quantificational modification of an idea, as with modals, negators,

evidentials, intensifiers, articles, numerals, and so on” (1994: 80). Elsewhere Chafe identifies linguistic evidence for ‘point of view’. This evidence includes reference to the emotions and evaluations of the person from whose point of view a discourse topic is developed (1994: 133, 136). The relationship between expressions of point of view and ‘non-idea information’ is not explored, but in both the emphasis is on the relationship between the speaker and discourse.

From a grammaticalization perspective, Heine, Claudi and Hünemeyer (1991) and Heine (1993) locate expressions in one of two types of cognitive domain: the “domain of concrete, ‘real-world’ phenomena” and the “more abstract world of discourse functions ... relating to the relative time, the temporal contours, and the truth value of events presented in linguistic discourse” (Heine 1993: 96).

These bipartite analyses tend to involve overlapping categories and lack detailed analysis of the multifunctional character of the specific linguistic features that express these interactive, procedural or evaluative notions.

More common are tripartite classifications such as those adopted by several functionally oriented approaches. These are broadly modelled on Bühler (1934), whose view was that language has three main functions -- representational (‘Darstellung’), expressive (‘Ausdruck’) and vocative (‘Appell’). Systemic functional linguistics (v. Halliday 1994) classifies meanings as ideational (roughly equivalent to representational meaning), interpersonal (for referring to speaker-hearer-discourse relations) and textual (for indicating inter-discourse relations). Discourse-marking expressions such as evaluatives therefore operate on the interpersonal plane and connectives on the textual. The Hallidayan approach is adopted directly by Hovy (1993: 362) and Hovy and Maier (1995: 13) to create a tripartite taxonomy of coherence relations. The systemic-functional approach is also the source of Traugott’s three levels of meanings: “meanings situated in the external described situation”, “meanings situated in the textual situation” and “meanings situated in the internal (evaluative/ perceptual/ cognitive) situation” in increasing order of abstraction (Traugott 1988: 409). Expressions on the more abstract planes were said often to be the result of a process of gradual

grammaticalization of items originally used on a less abstract one. In more recent work, however, Traugott has abandoned the three-plane approach, on the grounds that certain phenomena that the systemic-functional model divided between the interpersonal and textual planes are now treated together as ‘metalinguistic’ or ‘metatextual’ (Traugott 1995a: 47, 1995b: 28n6).

Grosz and Sidner (1986), based partly on Grimes (1975), also posit three ‘components’ of discourse: a linguistic structure, which consists of a string of utterances making up the discourse; an intentional structure, that is, the overall discourse purpose and sub-purposes, and an attentional state representing the focus of attention of interlocutors - the relative salience of entities, properties and relations - as the discourse unfolds. What Grosz and Sidner refer to as ‘cue phrases’ (discourse markers) are multifunctional: they can signal discourse boundaries, can indicate relations between intentions, and can convey propositional content. Moreover they can do these things simultaneously (Moore and Pollack 1992).

From a cognitive linguistic perspective, Sweetser, too, views utterances as “multi-leveled objects” (1990: 75) bearing meanings on three levels or planes. These are termed ‘content’, ‘epistemic’ and ‘speech-act’, corresponding to the socio-physical world, the world of reasoning and the world of speech exchange. Sweetser cites as evidence for the existence of these three planes the “large number of lexical items which show regular, parallel ambiguities” (1990: 81). That is, the similarity in the nature of the polysemy²¹ exhibited by diverse items can be partly accounted for by positing these planes. One of Sweetser’s examples is causal conjunction: each of the items *because*, *since*, *so*, or *although*, for instance, can function on each of the three planes (1990: 77-81). Causal conjunction on the content plane marks real-world causality of an event; on the epistemic plane it marks the cause of a belief or conclusion; on the speech-act plane it indicates the cause of the performance of the speech-act. Difference of meaning can be seen to parallel difference of plane. Haspelmath and König (1998: 568-570), following Sweetser, stress the importance of these three “levels of linking” (1998: 568) for

²¹ Sweetser distinguishes in principle between polysemy and pragmatic ambiguity: “in polysemy, a morpheme has several related semantic values; in pragmatic ambiguity ..., a single semantics is pragmatically applied in different ways according to pragmatic context” (1990: 76). However, these are not clearly distinguished in practice. Rather, different uses of items with highly abstract meanings are analysed as pragmatic ambiguity while different uses of items with firmer lexical identities are analysed as polysemous.

understanding the different grammatical constraints that can apply at each level.

There is of course no one-to-one correspondence between the expression of meanings on a particular plane (in Sweetser's sense) and the use of a particular linguistic structure. Evaluation of likelihood, for instance, can be expressed in English by more or less grammaticalized means: modal auxiliaries (e.g. *may*), epistemic adverbials (e.g. *perhaps*), predicates with adjectives (e.g. *is possible*), mental state verbs (e.g. *I expect*) and so on. Conversely, a single conjunction such as *because* can express content, epistemic or speech-act causality. Different languages grammaticalize the various evaluative and speech-act types of meaning in different ways (v. Palmer 1986: 7-8). Nuyts (1993a) points out that it would be good to have an account that recognized a plane such as epistemic modality as "one coherent qualificational category [underlying] all the different expression forms" (1993a: 963). Such an account would require a more fully-specified non-linguistic hierarchy of levels than has been proposed so far. However, there is evidence that different languages grammaticalize similar types of meaning. There are focal concepts which crosslinguistically tend to be grammaticalized and which Bybee and Dahl (1989) term 'gram-types'. Epistemic modality is one such type. There therefore seems to be a partial correspondence between a hierarchy of levels in the linguistic domain and a hierarchy of levels in some conceptual domain.

The notion of planes of meaning along the lines suggested by Sweetser (1990) will be useful to the discussion of DCs. These will be related later to rhetorical preferences and information structure (section 2.3.4). The terms 'content', 'attitudinal' (equivalent to speaker viewpoint) and 'speech-act' will be used, where 'content' refers to the propositional content²² of the host discourse segment, 'attitudinal' refers to speaker judgment concerning epistemic commitment and valuation (Givón 1995: 112) as well as the relations between discourse segments or between discourse segments and context of utterance, and 'speech act' refers to attributes of the discourse. To posit three planes in this way is no doubt an oversimplification, and the various types of modality, for instance, which often co-occur,

²² 'Propositional content' is in the sense of Lyons (1995:44), who distinguishes propositional from non-descriptive meaning. Lyons also emphasizes that "it is not propositions that sentences purport to express, but propositional content" (1995: 144).

should better be separated out. Until there is a principled way of achieving this, however, it seems most useful to adopt Sweetser's planes and regard them as bundles of planes. We shall distinguish, however, among attitudinal-epistemic (speaker commitment to the veracity of an idea), attitudinal-evaluative (speaker evaluation of the benefit of an idea) and attitudinal-rhetorical (speaker attribution of relations such as cause, contrast, concession among ideas).

It will be seen that, for the expressions under study here, the content and attitudinal planes are distinct, in so far as vague readings²³ that would combine interpretations from different levels are not usually possible. For the discourse markers to be discussed here, the content/metatextual distinction is often at least partially visible in the syntax where these polysemous expressions are concerned. As Givón comments, "one may, to quite an extent, identify within sentential syntax the code elements that bear specific responsibility for coding propositional-semantic information, and those that are more specifically responsible for coding discourse-pragmatic function." (Givón 1989: 84). However, there are cases of ambiguity, i.e. alternative readings rather than a merged reading, and of possible multiple analysis. A relation on one plane may co-occur with a similar relation on another. In actual discourse, therefore, the planes are closely intertwined.

The next section looks at the syntactic category of discourse connective expressions and at the syntactic relations between them and their hosts.

1.4 Syntax and semantics of discourse connectives

1.4.1 Syntactic category

Abraham (1991a: 250) notes that "in the literature on discourse particles it has always been taken for granted that there is such a separate category". But the status of that category and the criteria for determining its members are often unclear. According to one view, DMs constitute a syntactic category, and this is the position adopted in some studies of English DMs (e.g. Fraser 1988, 1990; Traugott 1995b, 1999, in press), although no tests are proposed by these authors that might serve as

²³ Following Kortmann (1997: 93), the term 'reading' is construed as neutral between 'meaning' and 'implicature'.

criteria for membership of the category. An alternative view is that DM is a pragmatic category and discourse-marking expressions are to be analysed as various types of sentence adverbial²⁴ (e.g. Jackendoff 1972, Higginbotham 1995), including modal adverbials, disjuncts and conjuncts (Greenbaum 1969).²⁵ However, many studies of discourse-marking expressions espouse neither view explicitly, and tend to remain vague about their syntactic status. This section argues that DMs and DCs are a pragmatic category and that syntactically they are usually sentence adverbs.

There does seem to be agreement that, if DM is not a word class, it is unclear what word class or syntactic category²⁶ discourse marking expressions belong to (Schiffrin 1987: 40; Brinton 1996: 35). A common analysis is that outlined by Ariel (1994: 3252): “It seems plausible to assume that the syntax will assign them [DMs] some syntactically motivated category on the sentence level (based on their distribution), with no strong expectation that they will be of the same syntactic category, while their interpretation at the discourse level may assign them a larger scope perhaps (e.g. a string of utterances).” Traugott seems to echo this approach: although discourse markers “have scope over far more than the sentence, in constituent structure terms they are part of the structure of the sentence” (Traugott in press). There is no obvious reason why they need be integrated into sentence structure. Elsewhere, as noted above, Traugott appears to treat DMs as a syntactic category; thus the shifts from VP-Adv to IP-Adv and from IP-Adv to DM are presented as the same type of category change (e.g. 1999: 177). Hansen proposes a category of ‘discourse items’: “such elements would function exclusively on the non-referential level of messages, and would be structurally optional, although semantically closer to the category of function words than to that of content words” (1998a: 225). But it is not clear what structural analysis they need where they do appear.

²⁴ The term ‘adverb’ usually denotes the part-of-speech or syntactic category, while ‘adverbial’ refers to an element of clause structure (Crystal 1991: 11), which may be a single word, a phrase or a clause. But the terms are not always used consistently; as Matthiessen and Thompson point out, “*adverb* is sometimes used to refer to a class of word, sometimes to a function in the clause” (1988: 319n).

²⁵ Occasionally they are described as belonging to the syntactic category of ‘particle’ (e.g. Polanyi 1995). See note 8 on problems of identifying particles.

²⁶ The terms ‘part of speech’, ‘word class’ and ‘syntactic category’ are used by different scholars. In generative approaches, the notion of word class has given way to a theory of categories as composites of syntactic features (Radford 1988: 154), and the term ‘syntactic category’ is now more widely used than ‘word class’. ‘Word class’ is used in many functional approaches. ‘Part of speech’ is a traditional term.

The problem lies in the apparently peripheral nature of DMs: they seem to be neither syntactically integrated into the sentence nor semantically integrated into the proposition. This poses a problem for syntactic theories which are dominated by the notion of the sentence. The traditional distributional methods of identifying categories are hard to apply; the range of sentential positions in which the DMs can occur is hard to account for (see section 1.4.4 below).

In practice, membership of the putative class ‘discourse marker’ is based more on perceived pragmatic function than on syntactic constraints: ‘discourse marker’ is a pre-theoretical category. This means that not all DMs need be of the same category and the question becomes which category any particular discourse marking expression belongs to. Yet there is a very close relationship between discourse marking and adverbial behaviour which needs to be accounted for. Firstly, there are constraints on the kinds of expressions that can be harnessed for discourse marking: expressions traditionally classed as manner adverbs are particularly likely to be recruited as discourse markers. Secondly, even those discourse-marking expressions that have no formal similarity with anything adverbial, such as transparent predicates, behave remarkably like sentence adverbials when used as discourse markers.

Conventional dictionaries label their entries with part of speech names, as though the category were an inherent quality of the word. But the notion of part of speech is itself problematic and especially that of adverb. Recent definitions are typically framed in negative terms: “modifiers of constituents other than nouns” (Schachter 1983: 20), “all modifiers other than adjectives, ... with considerable looseness in the application of the notion ‘modifier’” (McCawley 1983: 263). Such generous definitions have allowed the ‘adverb’ category to be variously described as “nebulous and puzzling” (Quirk et al 1985: 438), a “catch-all term” (Jackendoff 1972: 47), a “wastebasket” (McCawley 1983: 263; Brugman 1986: 52n) and a “grammatical dustbin” (Trask 1993: 10) for words that resist classification. Lyons (1977: 425ff) draws attention to the problems of part of speech definition. Allocation has traditionally been done on the basis of a mixture of the morphological, the syntactic and the semantic characteristics of a lexeme. But this method has been criticized for its risk

of circularity, and the most important criteria in deciding the category of an expression are syntactic (Lyons 1977: 425, Radford 1988: 57). Words arguably acquire their category, over time, through their use in syntactic structures (cf. Hopper & Thompson 1984). It is therefore possible to think of expressions as re-acquiring category or sub-category with each use, rather than as inherently belonging to a particular category. Much of the fuzziness and uncertainty in the allocation of expressions to syntactic categories is due to ongoing processes of language change, wherein expressions are gradually reanalysed into new roles and so acquire new categories. Many of the common discourse marking expressions of present-day English are a result of this type of reanalysis, and are arguably becoming grammaticalized (v. chapter 2.4). Henry Sweet summed up this situation by viewing expressions like *still* and *nevertheless* (DCs) as half-way between S-Adv and conjunction. He coined the term ‘half-conjunction’: “half-conjunctions connect logically only, not formally also ... half-conjunctions are necessarily sentence-modifying adverbs” (1892: 143-4). The discourse connective expressions to be discussed here will likewise be treated as sentence adverbials, as discussed by Jackendoff (1972), Bellert (1977) and Parsons (1990).

How to handle the borderline between predicate adverbs and sentence adverbs, as well as how to discriminate among various types of sentence adverb, is an old linguistic problem (McCawley 1988 chapter 19). Both the syntax of sentence adverbs and that of parentheticals pose serious problems for most current theories of syntax.²⁷ Previous semantic approaches to sentence adverbs and their classification are surveyed briefly below (1.4.3).

Cross-linguistically, most adverbs are derived from nouns, verbs or adjectives and are said to be more grammaticalized than those categories. Some of the commonest adverbs in English are used for discourse marking purposes, and as noted above, there is a close relationship between discourse marking and adverbial meanings. Yet adverbs are a difficult category and appear not to be well understood. Indeed, Cresswell suggests they are “semantically, and also syntactically ..., the least understood large class of words in natural language” (Cresswell 1985: 173). They are “messy”

²⁷ Recent proposals for dealing with sentence adverbs within generative syntax include Rizzi (1997), Cinque (1999). Nakamura (1997) proposes a cognitive linguistic account and Dik et al (1990) a Functional Grammar account.

(McConnell-Ginet 1982: 182) and have received less attention than other categories. “In the history of generative grammar”, writes Jackendoff, “perhaps the least studied and most maligned part of speech has been the adverb ... adverbs have been maltreated beyond the call of duty,... neglected in favor of more tractable constructions” (1972: 47). The very label seems to be more controversial than ‘noun’ or ‘verb’ or ‘adjective’. There are a number of grounds for calling into doubt the categorial status of ‘adverb’.

First, there seems to be little cross-linguistic evidence of a firm relationship between an adverbial meaning type and an identifiable syntactic category. Second, the desirability of a theory that can provide the most economical account, or capture the most generalizations, has led to what McCawley calls “the common prejudice that the number of ‘parts of speech’ should be small” (1983: 262). One target of winnowers has been the adverb. The idea that adverbs and adjectives, as modifiers of non-nominals and nominals respectively, are best treated as belonging to a single category is widespread among formal linguists (e.g. Radford 1988: 137-141; Roberts 1997: 15). Third, by contrast, it has been suggested that adverbs are too diverse a group and should be sub-divided into two or more categories²⁸. Cresswell notes that while for some categories it is enough to consider a small range of examples to come up with an adequate semantics, with adverbs it seems necessary to consider many examples: “maybe this shows that adverbs do not form a genuine semantic class” (Cresswell 1985: 173). Finally, there is evidence that in any case word classes should be treated as fuzzy categories (e.g. Comrie 1989: 109, Croft 1991, Taylor 1995, Haspelmath 2000).

Adverbs will therefore be treated as a fuzzy category having an important sub-category of sentence adverb. As noted above, different languages grammaticalize ‘speaker comments’ in different ways. English is unusual in having a very large number of sentence adverbs compared with other languages. In a cross-linguistic study of sentence adverbs in forty European languages, Ramat and

²⁸ This idea is not new - Dionysius Thrax (c. 100 B.C.) apparently listed twenty-eight different sub-types of adverb (Michael 1970: 101). And the ninth-century grammarian Smaragdus wrote of the adverb:

Est huius partis nimium latissimus ordo,
Multifer est pectus et violenta manus.

(‘The range of this part of speech is altogether too vast; its heart is abundant, its hand powerful’) (cited and translated by Michael 1970: 73)

Ricca (1998: 218) found that English lexicalized more of the 31 sentence-adverbial concepts they identified than any other language. One possible reason for this is that English lacks grammaticalized means of expressing certain types of modality and attitudinal meaning and recruits lexical expressions, especially adverbs, for the task.

In summary, we wish to distinguish clearly between the syntactic category (such as S-Adv) of an expression and its discourse category (such as DM or DC). On the syntactic level, sentence adverbs can connect clauses or clause complexes (or fragments), while on the discourse-pragmatic level, DCs connect ideas.

The next section looks at some semantic classifications of S-Adv and how these relate to DCs; the following section discusses the positions of S-Adv functioning as DMs including DCs.

1.4.2 Semantic classification of S-Adv

Semantic studies that have dealt with comment expressions have usually done so as part of wider analyses of various types of sentence adverbial. Schreiber (1971: 84), Jackendoff (1972), Bellert (1977), McConnell-Ginet (1982) and Parsons (1990) classify sentence adverbs into a number of semantic categories and some are described below.

Adverbs are usually classified by semanticists according to what categories they modify (Thomason & Stalnaker 1973; McConnell-Ginet 1982). They can be distinguished by their scope, and in the first instance by whether they modify a verb or verb phrase (V or VP) or a sentence (S). These are termed ‘predicate modifiers’ and ‘sentence modifiers’ respectively by Thomason and Stalnaker (1973: 201), who suggest a set of semantic criteria for distinguishing the two classes²⁹. Examples of sentence modifiers are modal adverbs like *necessarily* and *probably* and adverbs of attitude like *unfortunately* or *happily* in (4).

²⁹ The criteria are: (1) only sentence modifiers give rise to opacity; (2) only sentence modifiers give rise to quantifier scope ambiguities; (3) only a sentence modifier can include a sentence modifier within its scope; (4) only if an adverb is a sentence modifier can the sentence be paraphrased by ‘It is ADV true that ..’. Criterion (4), suggest Thomason and Stalnaker, “comes close to being a necessary and sufficient condition” (1973: 206). The four criteria are consistent with all but the vaguest of discourse markers being sentence adverbs.

- (5) a. a *clearly* ruffled professor [S-Adv] (cf. ‘a badly ruffled professor’)
 b. a *clearly* written essay [VP-Adv] (cf. ‘a badly written essay’)

Clearly in (5a) is S-Adv while *clearly* in (5b) is VP-Adv, reflecting the respective paraphrases of ‘it is clear that the professor is ruffled’ and ‘the essay is written in a clear manner’. The difference between S-Adv *clearly* and VP-Adv *clearly* then derives not from the size of the host unit, but rather from the discourse plane on which it operates.

A number of sub-categorization schemes for these adverbs will now be discussed.

1.4.3 Sub-categorization of speaker-oriented S-Advs

S-Advs, in McConnell-Ginet’s words, are “a heterogeneous bunch” (1982: 174). Table 1.2 shows Bellert’s (1977) refinement of Jackendoff’s (1972) categorization. It further sub-divides Jackendoff’s speaker-oriented sentence adverbs, arguing that there are distributional features other than just position which correlate with the semantic differences observed among them. Bellert’s tests include: whether the adverb can be independently negated (as in ‘dishonestly’, etc), whether corresponding questions are possible, and whether occurrence of the adverb is possible in performatives. Semantic differences among the speaker-oriented adverbs are correlated with these distributional differences. Parsons (1990), largely following Jackendoff (1972) and Bellert (1977), also proposes five types of modifiers³¹: speech-act modifiers (similar to Jackendoff’s speaker-oriented adverbs), sentence modifiers³², subject-oriented modifiers, VP modifiers and other modifiers³³. Like McConnell-Ginet, Parsons regards as homonyms sentence adverbs and VP adverbs that have the same form (1990: 66).

³¹ Parsons’ ‘adverblike modifiers’ include single-word adverbs, prepositional phrases and subordinate clauses (Parsons 1990: 40).

³² Sentence modifiers comprise alethic modality (‘necessarily’, etc.) and prepositional phrases such as ‘according to Agatha’ (Parsons 1990: 63).

³³ Other modifiers include focus adverbs such as *merely*, *only*, etc. Parsons also identifies a temporal category of modifiers, which cuts across these five categories.

	Distributional properties	Semantic properties	Examples
evaluative	-do not occur in Qs; -can be independently negated (UNfortunately) -outside scope of negation; -do not occur in performatives	-factive -adv + host is 2 propositions; -predicate over event referred to;	<i>fortunately</i> <i>surprisingly</i>
modal	-do not occur in Qs; -cannot be independently negated; -outside scope of negation; -do not occur in performatives; -can occur in if..then conditionals;	-adv + host is 2 propositions; -predicate over truth of event referred to;	<i>probably</i> <i>certainly</i> <i>evidently</i>
domain	-can occur in Qs; -outside scope of neg.; -do not occur in performatives	-constitute single proposition with host; -limit truth to one domain	<i>logically</i> <i>mathematically</i>
conjunctive	-can occur in Qs; -can occur in performatives;	-adv is additional proposition in that it expresses truth of particular relation bet 2 Ss;	<i>nevertheless</i> <i>therefore</i>
pragmatic (i) and (ii)	-can occur in performatives; -can occur followed by 'speaking' -cannot be independently negated;	-adv + host is 2 propositions; -are predicates with two arguments: (i) speaker and proposition, (ii) speaker and form of sentence	(i) <i>frankly</i> , <i>sincerely</i> (ii) <i>briefly</i> , <i>roughly</i>

Table 1.2. Summary of Bellert's scheme of speaker-oriented sentence adverbs (based on Bellert 1977)

Table 1.3 shows how some other proposals for the sub-categorization of sentence adverbs compare with Bellert's scheme, which remains an influential and enlightening way of looking at this varied bunch of adverbials.

In terms of discourse planes, expressions falling into Bellert's 'evaluative', 'modal' and 'conjunctive' categories tend to occur in the attitudinal plane, while expressions in her 'pragmatic' category tend to occur in the speech-act plane (see section 1.3 above)³⁴. It is the former three categories that are most relevant for the expressions that are the subject of this research. They will be

³⁴ The 'domain' category falls outside the scope of this study and will not be discussed.

	← - - - - -	speaker-oriented or neutral P _{speaker}	- - - - -	domain	frequency	← - -	subject-oriented P _{subject}
Jackendoff 1972	conjunctive	pragmatic	modal	evaluative	domain	frequency	subject-oriented
Bellert 1977							
Schreiber 1971		'consequently' 'frankly'	modal	evaluative	'economically'		
Allerton & Cruttenden 1974	contingency	presentation interpretation	message-oriented				interpretation subject-oriented
Huang 1975		performative	epistemic	attitudinal			
Quirk et al 1985	contingency		modality	respect	time		
Koktova 1986	conjunctive listing appositional focusing	style	likelihood	factive-attitudinal non-factive-attitudinal emotional	domain		
Parsons 1990	conjunctive	speech act modifiers	pragmatic epistemic/modal	evaluative	sentence modifiers		subject-oriented modifiers
Swan 1988, 1991		speech act adverbials	modal adverbs	evaluative adverbs			subject disjuncts
Kato 1993		performative	modal	evaluative			
Pin-Ngern Conlon & Evens 1994	transition	speaker manner	epistemic	speaker attitude			subject-oriented
Ramat & Ricca 1998	connecting	speech act	interpersonal level ← - propositional modal	event-oriented evaluatives	domain	representational event	interpersonal level propositional participant-oriented evaluatives

Table 1.3. Some semantic categorizations of sentence adverbials: comparisons with Jackendoff (1972) and Bellert (1977)

termed the ‘evaluative’ (following Bellert 1977, Parsons 1990, Swan 1988, 1991, Kato 1993), ‘epistemic [modal]’ (idem) and ‘rhetorical [connective]’³⁵. These three categories are illustrated in table 1.4.

Attitudinal speaker - oriented sentence adverbials	
rhetorical connectivity	Discourse connective (<i>However, ...</i>) (<i>In fact</i>) (<i>After all, ...</i>)
modality	Epistemic/evidential (<i>Perhaps, ...</i>) (<i>Apparently, ...</i>)
evaluation	Evaluative (<i>Unfortunately, ...</i>)

Table 1.4. Categorization of some attitudinal speaker-oriented sentence adverbials

Notions of modality are more or less grammaticalized. They are variously expressed (intra- and cross-linguistically) by modals, affixes, particles and modal adverbials. Grammaticalized modality (mood), as Palmer warns, may drift away from the semantics of modality (1986: 22). Equally, it is not easy to constrain notional modality: as Jespersen points out, “there are many ‘moods’ if once one leaves the safe ground of verbal forms actually found in a language” (1924: 321). Nonetheless, important semantic correspondences across the different realizations need to be taken into account. For van Valin and LaPolla (1997), for instance, adverbs are related to operators (e.g. ‘yesterday’ is related to Tense, etc.) and “adverbs related to more outer operators occur outside of adverbs related to more inner operators” (1997: 165). Cinque’s cross-linguistic study of adverbs adopts a stronger position still: he claims a perfect matching between particular adverbials and particular Mood projections. He argues, on the evidence of ‘partial orderings’ in different languages of various modal affixes, inflections and particles, especially stacked suffixes, that Mood projections follow a universal order of Mood_{speech act} > Mood_{evaluative} > Mood_{evidential} > Mood_{epistemic} (1999: 76) and that this order matches a universal order of S-Advs as follows: speech act (e.g. *frankly*) > evaluative (e.g.

³⁵ The term ‘conjunctive’ is prone to confusion with Quirk et al’s (1985) grammatical category of ‘conjunct’. ‘Rhetorical connective’ picks up Mann et al’s (1992) notion of ‘rhetorical structure’.

surprisingly) > evidential (e.g. *allegedly*) > epistemic (e.g. *probably*) (1999: 33-43, 106).³⁶ In “language after language, the same (restricted) functional notions appear to be expressible either via head morphology or via adverbs” (1999: 132)³⁷.

To sum up, many of the semantic notions expressed in English by S-Adv (such as evidentiality, evaluation, epistemic status or speaker commitment, grounding) are expressed in other languages by affixes, inflections or particles; what in English tends to be lexically expressed, in many languages is more grammaticalized.

Bellert’s categorization of S-Adv seems to be the most consistent and comprehensive: her five categories cover many, though by no means all, of the expressions commonly considered to function as DMs. However, as Ramat and Ricca (1998: 247) warn, the diagnostic tests of the sort proposed by Greenbaum, Bellert and others need rethinking.

A deficiency of Bellert’s and other classifications of S-Adv is that scant if any attention has been paid to the occurrence of epistemic, evaluative, etc. adverbials as modifiers of sub-sentential units. But the term ‘sentence adverb’ is well established, and the sub-category of speaker-oriented S-Adv as discussed by Bellert, Jackendoff and Parsons among others is the most adequate syntactic description of the type of modal ‘comment expressions’ that are the object of this study. The term S-Adv will therefore be used regardless. In fact, as explained in chapter 2, the unit ‘sentence’ will not be made use of at all; rather, text will be discussed semantically in terms of discourse segments or text spans, and syntactically in terms of phrases, clauses and clause complexes. A further problem is that, although much attention has been paid to forms that can serve as both S-Adv and VP-Adv, less attention has been paid to how some forms can also serve as more than one type of S-Adv.

As noted by Bellert, these S-Adv types at least partially correlate with syntactic position. The next section looks at the positions in sentential and sub-sentential units of S-Adv functioning as DMs/DCs.

³⁶ Cinque equates *perhaps* with Irrealis Mood, which comes below other Mood heads in the hierarchy he proposes and below some Tense heads (1999: 88).

³⁷ Although Cinque largely follows Bellert (1977), he ignores Bellert’s ‘conjunctive’ adverbs.

1.4.4. Position of S-Adv

Fraser claims that discourse connectives (his ‘discourse markers (DMs)’) occur in complementary distribution with the same form in the propositional plane: “there is never a doubt when an expression is functioning as a DM” (1998: 257). This seems to correspond to the idea that, for full clauses, it is possible to distinguish S-Adv from VP-Adv by their syntactic position.

Sentence adverbs occur in clause-initial and pre-verbal positions. A speaker-oriented adverb can also occur in clause-final position with appropriate ‘comma intonation’. Greenbaum (1969: 183) records how this distinction is put to effect in an old film, where an anxious bridegroom says to his friend, ‘I expect my wife to behave, *naturally*’, to which the friend retorts ‘I expect my wife to behave *naturally*’.

Clearly in (6) is ambiguous.

(6) They *clearly* understand the needs of the public (Per)

VP-Adv: in a clear manner, distinctly.

S-Adv (evidential): Speaker infers (e.g. from their actions) that they understand.

This ambiguity is explained by plotting the possible positions for VP-Adv *clearly* and for S-Adv *clearly*:

[▼	▼	▼]	VP-Adv
they	understand	the needs of the public	
▲,	▲[], ▲	S-Adv

As a manner adverb, *clearly* is positioned inside the VP; as a speaker comment, it is outside the VP.

These different syntactic structures are usually evident when more than one auxiliary element is present:

(7) They should be clearly heard by the audience (Per)

▼	▼	▼	VP-Adv
they	should	be heard	by the audience
▲,	▲	▲	S-Adv

In PDE, however, VP-adv *clearly* is occasionally found in S-adv position, e.g.:

- (8) ?The remains of the engine and parts of the wings could *clearly* be seen but most of the plane had been recovered ... (Per)

The question of whether there is a (set of) position(s) specific to DMs is, of course, linked to the question of the status of DM as a syntactic category. Aijmer claims there is “a special discourse marker slot”, which turns out to be that of speech-act adverbials (Aijmer 1996a: 207, 212-5, 233). Brinton agrees but points out that initial position may be the preferential slot (1996: 33; see also Levinson 1983: 87).³⁸ According to Schiffrin, on the other hand, discourse markers “can occur quite freely within a sentence at locations which are very difficult to define syntactically” (1987: 32). This may refer to the fact that discourse marking adverbials frequently occur as parentheticals. Much of the work on the syntax of S-Advs (e.g. Greenbaum 1969, Allerton & Cruttenden 1974, Bartsch 1976, Ernst 1984, McCawley 1982, 1983, Espinal 1991, Kato 1993; Cinque 1999) addresses the problems posed by the positional flexibility of sentence adverbs, especially their occurrence as parentheticals. Most of the studies attempt to provide an analysis within a sentence-based framework, but so far there is little consensus on an adequate analysis.

There have been attempts recently within the generativist framework to account for the syntactic behaviour of sentence adverbs by having them generated in high-level functional projections on the ‘left periphery’. Rizzi (1997) has them generated in a Topic projection, while for Cinque (1999) they are generated in one of a variety of high-level Mood projections. Cinque attempts to integrate the semantic categories identified by Jackendoff and Bellert with a hierarchy of Mood projections.

It has been observed that whereas subordinating conjunctions can occur at the beginning of either segment in a relation, parenthetical DCs always occur in the second segment (Power et al (1999) cited by Oates (1999: 38)). In the subordinating case (e.g. *although*), the segments can occur in either order, whereas in the parenthetical case (e.g. *nevertheless*) the order of ideas and the choice of DC are interdependent. This observation is borne out by the data in the following chapters. It emphasizes the importance of information structure for the understanding of DCs. For example, a DC in a

³⁸ It has been claimed that native speakers tend not to notice initial position DMs (lefthand discourse brackets) but “are sensitive to those that occur as righthand discourse brackets, so much so that they can use discourse markers occurring in these positions as socially stigmatizing features of an individual’s speech style” (Watts 1988: 244).

patterns.” (1992: 101).

Sweetser (1990: chapter 1 and p. 76) argues for a distinction between polysemy (different senses) and pragmatic ambiguity (different planes). She argues that an account of both “can be given in the same framework independently needed to account for certain major trends in semantic change” (1990: 11-12). The polysemy continuum can therefore be seen in both a synchronic and a diachronic perspective. If the sense of a form diverges over time, it moves rightwards towards homonymy, the relatedness of meaning of the divergent senses becoming less and less transparent. A synchronic analysis will reveal forms at different points along the continuum, i.e. at different stages of sense divergence.

Maintaining the traditional view of one form - one meaning forces one to adopt either a pragmatic analysis or a homonymy analysis. A second approach therefore is to suppose that there is a core semantics for the form and the polysemies are contextually-determined, i.e. pragmatic. Several analyses of discourse markers derive from this view, distinguishing the ‘literal’ or ‘semantic’ meaning of a form from its discourse-marking or ‘pragmatic’ meaning (e.g. Biber & Finegan 1988, Fraser 1988, Ariel 1994). This unitary analysis equates semantic sense with propositional use and the metatextual or discourse marking use is then treated as a pragmatic function. That is, discourse meaning is not regarded as semantic. Or alternatively DMs belong on a continuum of transparency from primarily semantic ones (transparent markers, close to the ‘literal’ meaning of the expression) to those that are devoid of semantic meaning (opaque markers) (Ariel 1994)³⁹.

But to claim that the discourse-marking expression has no semantics is to imply both that the two meanings are related obviously (which often they are not) and that one is consciously metaphorical (which it clearly is not). Perhaps ‘literal’ meaning is confused with ‘older’ meaning, in cases where the semantic relationship between the propositional and the discourse-marking uses is neither transparent nor completely opaque. The ‘semantically void’ analysis is rejected here, then, on the grounds that the

³⁹ It is helpful to distinguish clearly between the semantic-pragmatic cline and the transparent-opaque one. Idioms too are commonly viewed as being on a continuum of transparency from those whose sense is almost computable from their parts to those that are unfathomable till learned as wholes. This does not mean that opaque idioms are somehow ‘more pragmatic’ than transparent ones, but only that their senses bear less resemblance to those of their parts. The same applies to strings of morphemes in lexicalizations and, here, to expressions used as discourse markers.

distribution of discourse marking expressions is far from random: they clearly have a conventional semantic meaning and are not wholly context-dependent.

The third approach views the different uses as homonymic. This analysis treats the senses as separate, identity of form being synchronically immaterial. The drawback of this analysis is that it implies that the formal correspondences are purely random. It cannot account for speakers' intuitions that the senses are related, nor for the regularities observable in the pairings of senses.

It is more useful, then, to discuss the different meanings in terms of polysemy, resulting from diachronic development. A polysemy analysis tries to have the best of both the unitary and homonym approaches by accounting for the relatedness of form and meaning while allowing for the distinctness and opacity of the different uses.

1.5 Summary

The literature on DMs is heterogeneous, using categories that cross-cut one another and claiming generalizations based on very different data sets. Nonetheless, what emerges is a relatively stable consensus on the existence in English of a set of discourse marking expressions with certain characteristics, namely:

- they behave as sentence-adverbial adjuncts or parentheticals (whether their supra-sentential semantic scope and syntactic independence allows them to be treated as sentence constituents is a general problem of parentheticals)
- they are predominantly subjective and speaker-oriented
- as 'comment expressions', they express ideas on a different discourse plane from that of their host
- they contribute to discourse coherence (or Relevance in the case of Relevance theorists who reject coherence)

These points all concur with Locke's (1692) observations on the nature of particles, quoted at the beginning of this chapter.

There is some disagreement on the following points:

- the status of the discourse marker category (discourse-pragmatic or grammatical)
- the criteria for deciding whether an expression is a discourse marker
- how markers can be subcategorized
- whether forms used as markers are homonymic or polysemous
- whether the development of markers can be considered grammaticalization.

The multiple ways of defining and classifying discourse markers and discourse connectives that have been proposed in the literature stem partly from the fact that discourse connectivity, or the ‘coherence relation’, can be realized by a range of grammatical categories and by expressions that also have non-connective functions. We have defined discourse connectives as discourse markers (comment expressions) that signal a rhetorical relation between two or more discourse segments. They therefore contribute to discourse coherence, which is dealt with in the next chapter.

CHAPTER 2

DISCOURSE COHERENCE, RHETORICAL RELATIONS AND GRAMMATICALIZATION

2.1 Discourse coherence	52
2.1.1 Discourse-pragmatic approaches	53
2.1.2 Linguistic approaches	56
2.1.3 Social action approaches	58
2.2 Units of discourse analysis	59
2.2.1 Spoken vs written language	59
2.2.2 Segmenting spoken and written language	61
2.3 Rhetorical Structure Theory (RST)	64
2.3.1 Some examples of RST relations	64
2.3.2 Rhetorical relations and DCs	67
2.3.3 Nuclearity	70
2.3.4 RST and discourse planes	73
2.3.5 Applying RST to the analysis of modal adverbial contexts	75
2.3.6 Rhetorical patterns	78
2.4 DCs and grammaticalization	79
2.4.1 The development of DCs as instances of grammaticalization	79
2.4.2 Recent approaches to characterising grammaticalization	83
2.4.3 Sentence-based vs discourse-based grammaticalization	85
2.4.4 Semantic / pragmatic change	88
2.4.5 The grammaticalization cycle	91
2.4.6 Grammaticalization, discourse connectives and polysemy	94
2.5 Conclusion	96

2.1 Discourse coherence

Discourse connectives help hearers/readers to make sense of a discourse by contributing to discourse coherence. They are lexical markers of coherence relations (v. ch. 1.1), also known as rhetorical relations. Discourse, or text, is understood to cover both monologue and dialogue, written or spoken.

Chapters 5-8 will describe co-occurrences between four English modal adverbials (*at least, in fact, after all, of course*) and rhetorical contexts. It will be argued that these co-occurrences have fostered connective readings and grammaticalization of the adverbials. This chapter describes the concepts of coherence, rhetorical context and grammaticalization as they are used in the later analyses. First it outlines an approach to discourse coherence. It then describes Rhetorical Structure Theory (RST), the tool used to identify rhetorical relations and to examine some of the functions of the English modal adverbials discussed in chapters 5-8. The final section of the chapter describes differing

characterizations of grammaticalization and discusses their relevance to the development of DCs.

An account of coherence is central to discourse studies, although there is little consensus on the use of terms such as coherence, connectedness, connexity, cohesion or textuality (Conte et al. 1989). We shall assume that coherence is the property of being pragmatically meaningful; it is what distinguishes a text from a string of clauses. We suggest that coherence is a matter of degree. It is not a property of the language, but refers to the fit between the rhetorical goals of the discourse and the interpretation of the discourse.

The myriad approaches to discourse coherence include attempts to determine formal text structure rules (e.g. Werlich 1976, van Dijk 1977, Kamp and Reyle 1993, Polanyi 1995), analysis of lexico-grammatical means of indicating relations among sentences (e.g. cohesion studies such as Halliday and Hasan 1976, Martin 1992), and building of taxonomies of the possible relations among the parts of a text (e.g. coherence studies such as Hobbs 1985, Knott and Dale 1992). Approaches to coherence can be broadly classified as cognitive/discourse-pragmatic, linguistic, and sociolinguistic.

2.1.1 Discourse-pragmatic approaches

This study is based on a discourse-pragmatic view of coherence as anchored in speaker/writer purpose, along the lines of Tomlin et al's (1997) framework. For Tomlin et al, building coherent discourse involves four types of information management — rhetorical, referential, thematic (topical) and focus management — each of which is applied at global, episodic and local levels (1997: 66). We shall be concerned with rhetorical management and, to a lesser extent, topical management. Rhetorical management is constrained by rhetorical purpose, which (in a trivial, uncontroversial sense) motivates discourse coherence¹. The better the interpretation matches the intention, the more coherent the discourse.

A topic is what the discourse is 'about' at a given point and is a pre-theoretical notion² (v.

¹ The rhetorical purpose being logically prior, it is unconstrained by the topic, but not vice versa.

² The term 'topic' is sometimes defined in relation to 'focus', implying that topics cannot be in focus. Topics as outlined here can be in focus. Following Chafe's advice (1994: 78), we shall largely avoid the term 'focus'.

Lambrecht 1994: 118, Tomlin et al 1997: 83-91). Topics are a property of the speaker, not of the discourse and are identified intuitively by addressees and analysts. A distinction is often made between 'sentence topic' and 'discourse topic'. But topics are viewed here as hierarchically ordered — super-topics, topics, sub-topics — regardless of sentence boundaries, with pragmatic constraints on the conjunction of ideas at different topic levels (cf. Chafe 1994: 121, van Dijk 1977: 136). At any given point in the progression from topic to topic, a 'presupposition pool' (Vennemann 1975: 304) or 'presuppositional situation' (Lambrecht 1994: 65) is available to speaker/writer and addressee. Coherence is seen as belonging to the communication rather than to the language or to the social context.

Cognitive and discourse-pragmatic approaches view coherence as a cognitive rather than a linguistic or social category. As de Beaugrande notes on language reception, "the question of how people know what is going on in a text is a special case of the question of how people know what is going on in the world at all" (1980:30). As mentioned above, we adopt the view that purpose or intention is central to coherence. Several otherwise diverse approaches to discourse analysis share this notion that coherence is based on communicative purpose.

Grice notes that discourse has "a common purpose or set of purposes, or at least a mutually accepted direction" (1989:26) and goes on to suggest that it is this purpose or direction that constrains the possible conversational moves. The conversational maxims that constitute Grice's Cooperative Principle are therefore themselves dependent on purpose. Brown and Yule emphasize that the speaker or writer's 'intended meaning' describes the coherence of a text (1983: 224-5).

Communicative purpose is also at the centre of Grosz and Sidner's 'intentional structure': the purposes of discourses and discourse segments "provide the means of ... distinguishing discourses that are coherent from those that are not" (1986: 178). Moore and Paris (1992) emphasize the role of 'intended effect' in any model of discourse. Likewise, purpose is central to Mann, Matthiessen and Thompson's model, in that every relation between two ideas carries an 'intended effect' on the reader: the source of coherence is the contribution of every part of the text "to a single purpose of the writer,

... to achieve a single effect” (1992: 43). This echoes Johnson-Laird’s dictum that “a necessary and sufficient condition for discourse to be *coherent* ... is that it is possible to construct a single mental model from it. The use of natural language to communicate depends on an *intentional* correspondence between a symbolic expression and a state of affairs”(1983: 370, 405).

Purpose is often revealed via the modality patterns of the text, according to Lundqvist. The “global modality structure”, particularly of argumentative texts, leads the reader towards the “global communicative goal of the sender” (1989: 111). Or as van Dijk puts it, “it is not the ordering of the facts themselves, but the ordering of the perceptions and the knowledge about them [that] determines the structure of the discourse” (1977: 97).

And a focus on purpose provides a link with traditional rhetorical analysis via what van Eemeren calls ‘normative pragmatics’ (van Eemeren 1990). This seeks to bridge the gap between prescriptive and descriptive approaches to language. Traditional rhetorical analysis is sometimes overlooked by linguists, but the Aristotelian ‘topoi’ approach to rhetoric has been revived by some scholars (e.g. Perelman and Olbrechts-Tyteca 1969, Anscombe and Ducrot 1983). The topoi, as well as generating arguments, could ensure transitions from arguments to conclusions. These relied on relationships of identity, similarity, difference, contrast, subsumption (e.g. part/whole relations), causality, and so on (Kienpointner 1995). The role of analytical research, according to van Eemeren, is to reconstruct discourse as “an attempt to counter doubt regarding the acceptability of a standpoint” (1990: 41), and so to provide insight into the attitudinal and rhetorical aspects of the discourse, that is, the persuasive aspects. The role of empirical research includes identifying “patterns in interactional argument” (1990: 41). The link with traditional rhetoric has also been taken up recently by scholars developing a ‘historical pragmatics’ (Jucker 1996, Arnovick 1999), particularly the idea that many pragmatic effects are the result of rhetorical/persuasive strategies. An account of meaning change in terms of ‘rhetorical strength’, for instance, “demands an examination of the relation of rhetorical meaning to pragmatic meaning” (Fitzmaurice 2000: 2).

Moreover, evidence from language change studies suggests that “the speaker/writer does most

of the work of innovation, not the hearer/reader ... rarely does the act of interpretation itself lead directly to innovation” (Traugott 1999: 95). This supports the idea that it is speaker/writer intentions, rather than hearer/reader interpretations, that build discourse coherence. As Keller puts it, “understanding means recognizing the speaker’s open intentions³ ... Misunderstanding occurs if the addressee ... has ‘understood’ something the speaker did not mean” (1994: 96).

The ‘intentionalist’ or ‘personalist’ view of meaning as rooted in speaker intentions has been attacked on the grounds that it is more productive to study meaning as revealed in hearer/reader interpretation. It has been, claimed, for example, that in metonymic semantic change it is the hearer who “triggers the metonymic innovation which induces a metonymic change later on” (Koch 1999: 156). Others argue that it is social interaction that constructs meaning (e.g. Duranti 1993) or that much language, particularly ritual language, is intention-less (e.g. Du Bois 1993). Moreover, it is particularly likely that the observable frequent use by individuals of ‘favourite’ discourse markers, such as *actually*, *basically*, *essentially*, *in fact* and so on, is highly automated and may even conflict with other, more conscious communicative goals. Nonetheless, the distribution of these items is far from random. At some level, however unconscious or inaccessible to introspection, we must assume their use is motivated. Examination of their distribution may give us clues about why speakers select them how and when they do.

2.1.2 Linguistic approaches

Linguistic approaches to coherence look for structure above sentence level: for discourse rules, analogous to grammatical rules, which would account for ‘texture’ or textual well-formedness. They have focused on formal means of achieving connectivity among sentences. What van Dijk termed ‘discourse coherence’⁴ was to be accounted for by discovery of rules of structural well-formedness: “Those utterances which can be assigned textual structure are ... acceptable discourses of the language

³ Keller distinguishes between open and hidden speaker intentions (1994: 97).

⁴ van Dijk uses ‘coherence’ to include what is usually called ‘cohesion’ in the sense of Halliday and Hasan (1976) (van Dijk 1977: 126n).

... for any serious linguistic theory, RULES must be formulated systematically relating the semantic representation of sentences to that, at the macro-level, of the sequence [original emphasis]" (van Dijk 1977: 3-6). However, discourse is too open-ended and unpredictable to be rule-based in the linguistic sense. Ideas are linked by a mixture of lexico-grammatical and non-linguistic means.

Lyons has claimed that, to be deemed to be 'text', a stretch of language "must exhibit the related, but distinguishable, properties of cohesion and coherence [original emphasis]" (1995:263). Cohesion refers to the lexical and syntactic means, such as reference, conjunction and lexical collocation. Coherence refers to the content. Discourse well-formedness is thus a mixture of the linguistic (cohesion is largely subject to grammatical rules) and the non-linguistic. This is the approach of Halliday and Hasan. For them, coherence is achieved through the dual means of cohesion, or text-internal coherence, and register, or coherence with the context of situation (1976: 23). Cohesion is said to be a semantic relation that may or may not be realised by linguistic features, and which involves a relation of presupposition between one textual element and another (1976: 4-9). At the same time, cohesion is described as "the means whereby elements that are structurally unrelated to one another are linked together" (1976: 27). The disadvantage of Halliday and Hasan's concept of cohesion is that it is a mixture of different types of feature: some lexico-grammatical, some semantic; some propositional, some non-propositional. They have been brought together under one label because they signal relations that hold across sentence boundaries. The concept of cohesion is thus subordinated to the idea that the sentence marks a natural boundary between syntax and discourse, ie. the boundary of 'structural relatedness'. But the very data they adduce suggest that this is not the case. Halliday and Hasan's arguments for ignoring intra-sentential cohesion and for cohesion being "above considerations of structure" (1976: 9) are weak. Their three types of 'cohesive relation' (1976: 304) can be better accounted for separately: relatedness of form by linguistic presupposition, relatedness of reference by topic continuity or thematic progression and semantic connection by coherence relations.

2.1.3 Social action approaches

In the sociolinguistic approach to coherence, utterances realize social actions.⁵ For some, this approach still involves rules. For Labov, for instance, "the fundamental problem of discourse analysis is to show how one utterance follows another in a rational, rule-governed manner" (1972: 299). But they are not linguistic rules: "sequencing rules do not operate between utterances, but between the notions performed with those utterances. In fact, there is usually no connection between successive utterances at all." (Labov 1972: 301). Labov's discourse 'rules', then, contain social constructs. In similar vein, conversation analysis emphasizes the social act, but it rejects the notion of 'rules' of discourse and adopts an inductive theory-neutral approach to uncover the properties of discourse sequencing (v. Levinson 1983: 286). Speech Act theory also focuses on the utterance as social action, especially on the relationship between particular sets of verbs and the context of situation (v. Searle 1969). The drawback of the social action approaches is that "rather little attention is paid to the linguistic aspects of the realisations of these actions" (Brown and Yule 1983: 228). These approaches have difficulty relating particular linguistic features to the discourse structures they set up. As a result, they shed little light on the nature of linguistic entities or on the relationship of these to discourse meaning.

Coherence, then, on the discourse-pragmatic view adopted here, resides in intentions and their recognition. It is an informal, pretheoretical category; it is not intended to explain anything. Intentions can be signalled by DCs among other mechanisms. DCs connect ideas and are themselves ideas (Mann and Thompson 1986: 59). Some way is therefore needed of identifying ideas in text and of dividing discourse into units that correspond to ideas. The identification of these idea units, although in practice often obvious, is not easy to formulate. The presence of a connective is often taken as evidence of a boundary (v. Brown and Yule 1983: 94-100). But to discover the extent to which particular expressions are connective, we need to have an independent way of dividing discourse into units, so

⁵ This approach is partly inspired by ideas from the social sciences, particularly ethnography. Ethnographers such as Geertz (1973) saw cultures as built up of social actions that were conventional symbols to be interpreted. Yet these ethnographers were in turn 'borrowing' from linguistics in their attempts to 'read' cultural actions as texts (v. Ricoeur 1971).

that we can then see what happens at the unit boundaries.

2.2 Units of discourse analysis

It is uncontroversial that "a theory of text structure describes what sort of parts texts have, and the principles of combination of parts into entire texts" (Mann, Matthiessen and Thomson 1992:41). Discourse segmentation methods have long been a matter of debate among discourse analysts (e.g. van Dijk 1977, Longacre 1996, Polanyi 1988, 1995, Petöfi (ed.) 1988, Werth 1984, Rickheit and Habel (eds.) 1995). The tendency has been to segment written language along syntactic lines and spoken language along prosodic lines.

2.2.1 *Spoken vs written language*

A great deal has been written about the differences between spontaneous spoken language and written language (e.g. Akinnaso 1982, Chafe 1982, Tannen 1982, Brown and Yule 1983, Biber 1988 ch 3, Miller and Weinert 1998 ch.1). One of the most cited differences is the mechanisms used to convey information structure.

First, written language is said to be denser (Chafe 1992: 288). In written narratives, for instance, Beaman (1984) found greater information density, evidenced by structures such as nominalizations and non-defining relative clauses. Spoken language she found more repetitious and redundant. Beaman attributes much of this observable difference in the relative proportions of different types of coordination and subordination to information-structuring requirements.

Second, written narratives are said to contain more explicit markers of connectivity and of information status than do spoken ones (Tannen 1982). To signal relationships between ideas in spoken language, particles and intonation and paralinguistic features are said to be preferred over conjunctions or subordination (Miller and Weinert 1998: 22-3).

These types of observation give rise to the generalization that written language favours hypotaxis while spoken language favours parataxis. Croft, for instance, suggests that a main difference

may be degree of embedding (2000, 83).

Both Beaman (1984) and Biber (1988), however, point out that much of the research is contradictory. While some studies find more subordination and more complexity in written language, other studies find the opposite. Chafe (1988), comparing punctuation units with intonation units, finds that some writings are much more 'prosodically spokenlike' than others (1988: 424). As Beaman points out, there is not always a clear distinction made between the effect of the medium (spoken or written) and the effect of level of formality or of rhetorical purpose (1984: 79). Indeed, it has also been suggested that more variation can be found among written texts and among spoken texts than between written and spoken (Biber 1988). In particular, discourse marking seems to be genre-specific rather than medium-specific. For example, it is likely to involve temporal relations in narrative, and logical relations in persuasive discourse.

There is evidence that, although superficially different, written monologue inherits from spoken dialogue many of the most natural patterns of conversation, and that equivalences can be found. Some researchers have suggested that written monologue contains parallel organizational structures to those found in conversation. According to Schiffrin, "many features of rhetorical argument can be understood only if we define arguments as talk which is basically dialogic in intent" (1988: 253). There may be a monologic equivalent, for instance, of the 'adjacency pair' notion of conversation analysis, that is, "organized patterns of stable, recurrent actions that provide for, and reflect, order within conversation" (Schiffrin 1994: 236). This is especially so for argumentative discourse, where patterns of recurrent contrast (argument, counter-argument) are reminiscent of dialogue: "the vestiges of indisputably dialogical interactions remain palpable in meanings/functions that otherwise appear monological in viewpoint", as Schwenter puts it (2000: 260-261). In sum, "argument seems to be a mode of discourse which is neither purely monologic nor dialogic" (Schiffrin 1987: 17).

Since this study uses both written and spoken data, we should like to be able to perform comparable analyses on the two mediums. We shall consider two proposals for the segmentation of

discourse: that of Chafe (1982, 1988, 1992, 1994), developed for spoken language, and that of RST (Mann and Thompson 1987), developed for written language, and try to find a compromise method applicable across spoken and written language.

2.2.2 *Segmenting spoken and written language*

For Chafe “idea units typically have a coherent intonation contour, they are typically bounded by pauses, and they usually exhibit one of a small set of syntactic structures. They are a striking, probably universal property of spoken language” (1982: 37). Many analysts have investigated these likely correspondences between idea units, intonation units and grammatical units (e.g. Chafe 1980, 1988, 1992, 1994, Polanyi 1988, 1995, Croft 1995). Chafe uses the term ‘ideas’ in a technical sense to refer to “items of information” that are referents, events, states or discourse topics (1994: 80). In particular, Chafe’s (1980) observations support a ‘one new idea constraint’, whereby an intonation unit can contain several ideas, but only one new idea. As for the relation between intonation and grammar, intonation groups have been observed to correspond with clauses more than with any other syntactic unit (Cruttenden 1986: 76, Chafe 1994: 65-6, Croft 1995: 844-5). None of the three types of unit — idea unit, intonation unit, syntactic unit — is necessarily basic, and identifying one in terms of the others becomes circular (Cruttenden 1986: 36)⁶. Nevertheless, together they provide a practical if informal means of segmenting spoken discourse.

Chafe’s above definition of ‘idea’ is equally applicable to written language. But what in written language plays the role of the intonation unit?⁷ Arguably, there is intonation in written

⁶ Although intonation units seem to be the most empirically reliable way of segmenting discourse, Cruttenden acknowledges that “when regular correspondences between intonation and grammar/semantics have been established in cases where boundary assignment is clear, we may lean heavily on such correspondences when assigning boundaries in the difficult cases” (1986: 36).

⁷ Chafe (1980, 1994: 139-144) compares the sentence in written language with the group of intonation units that finishes with a falling intonation in spoken language (cf Brown’s ‘paratones’ (Brown 1977: 86), which she terms ‘extended sentence’). Both, she suggests, can best be described as multi-foci information chunks that constitute ‘centres of interest’. The motivation for the cohesiveness of the ‘sentence’ in each medium surely derives from its information structure. Croft (1995: 841-2), on the other hand, suggests that the sentence in written language resembles the intonation unit in spoken language on the grounds that each is the structural ‘upper limit’ of its respective medium. However, it is not obvious that there is a clear-cut upper structural limit for spoken or written language. Rather, there seems to be a cline of bondedness reflecting information structure needs (v. section 2.3.3). Moreover, the distinction between written and spoken language is not so clear-cut either. In discourse written to be spoken, it is certainly not the case that sentences correspond to intonation units. In order to understand the rhetorical structure of discourse, and thereby the role of lexical discourse connectives, it is important to segment along information boundaries. As discussed here, these

language too, as acknowledged by many analysts. Petöfi reminds us that for written language “the prosodic structure as perceived by the analyser ... plays a relevant role, since it constitutes an essential component of the meaning-carrier structure of the text” (1983: 8). Chafe also claims that the reader’s projection of an intonation pattern on to writing reveals the “covert prosody of written language, one of whose principal qualities is the presence of segments akin to accents and intonation units” (1992: 273). Chafe segments a written English text into these ‘intonation units’ and finds that, as in spoken language, an intonation unit often contains more than one idea. However, written language does not conform to the ‘one new idea per intonation unit’ constraint that Chafe found for spoken language: in around 28% of ‘written intonation units’ in the English text Chafe found more than one new idea (1992: 288).

Mann and Thompson (1987) segment texts into ‘text spans’, the smallest of which are known as ‘units’. These spans and the relations between them combine in a hierarchical structure to form a complete text having an overall purpose. A text span is defined rather vaguely as “an uninterrupted linear interval of text” (1987: 4) and as having “some relatively theory-neutral functional integrity” (Mann, Matthiessen and Thompson 1992: 51). Spans are “typically located at the boundary region shared by relational structure and syntactic structure” (Mann, Matthiessen and Thompson 1992: 51). Thus, “units are essentially clauses” (Mann and Thompson 1987: 6). Despite the poor definition of ‘unit’, in practice the syntactic means of identifying units, supplemented by the ‘functional integrity’ requirement, works well for written language. In fact it is the ‘combining clause’ that is the default minimal unit of analysis. A sharp distinction is drawn between **clause combining** or **hypotaxis** (adverbial clauses of time, cause, condition, etc) and **embedding** (defining relative clauses, clausal complements etc). Hypotactic clauses enter into relations with other clauses, while embedded clauses do not (Matthiessen and Thompson 1988: 279; see also Halliday and Hasan 1976: 196). So although the identification of spans is based on interpretation, there is a clear structural constraint.

Chafe’s notion of idea is too fine-grained for present purposes: discourse coherence relations

usually though not always correspond, for both spoken and written language, to constituent boundaries on one hand and to intonation boundaries on the other.

hold between ideas corresponding to events or states but not referents. On the other hand, Mann and Thompson's reliance on fully-fleshed clauses and clause complexes is impractical for dealing with relations among phrases and fragments.⁸ We shall therefore, partially following Chafe, take a unit containing a new idea in the sense of 'an assertion' (Lambrecht 1994: 54-5) as the unit of discourse segmentation. That is, the interpreted information structure is taken as basic, since it is the interpretation that enables an intonation contour to be mapped to written language. The discourse segmentation is thus "determined by semantic criteria and guided by syntax and intonation" (Polanyi 1995: 15). This approach is exemplified in (1). The material in square brackets disambiguates. The underlining identifies the host of *at least*. The accents mark primary stress. Segment boundaries are marked by |.

- (1) a. | *At least* the ventilátion is improved | [and maybe some other features too]
 b. [Passenger comfort is much enhanced.] | *At léast*, | the ventilátion is improved
 c. [There are still many drawbacks, but] | *at least* | the ventilátion is improved |

The sequence *at least the ventilation is improved* is analysed as follows. In (1a), it is a single discourse segment and a single intonation unit. In (1b) the sequence is two segments — *at least* (attitudinal-rhetorical) and *the ventilation is improved* (content) — and two intonation units (cf section 1.4.4). In (1c) it is two segments: *at least* (attitudinal-evaluative) and *the ventilation is improved* (content). But it is a single intonation unit. (These three uses of *at least* are discussed further in chapter 7.)⁹

It will be argued in the following chapters that typical patterns of information structure at discourse level have influenced the semantic development of each of the English modal adverbials of certainty in this study in similar directions. Various classes of adverbial have been said to have an

⁸ As pointed out by Scott and de Souza (1990), rhetorical relations can hold between two phrases and between a phrase and a clause, as well as between clauses. In principle, RST allows for relations "between portions of text below the clause level, or between clause-level units and units below the clause level" (Mann and Thompson 1986: 59n). In practice, work in RST has concentrated almost exclusively on clause-length and larger units. It has ignored fragments. Yet in English there are many 'fragment-taking conjunctions' (Fillmore, Kay and O'Connor 1988: 517) where the fragment that is the complement of the conjunction has a clausal function. Examples of fragments will be seen in chapter 7.

⁹ The context of the source clause shows that the original *at least* is Evaluative: 'The dash also looks uninviting. Many cars costing a third as much have a more upmarket fascia, better quality trimmings and more comfortable seats. *At least* the ventilation is improved.' (Per).

information structuring role in that they occur at discourse boundaries (Brown and Yule 1983: 97-100, Thompson and Longacre, 1985: 208-210, Givón, 1990: 847, Tannen 1993: 48, Miller and Weinert 1998: 196, 253-355). To understand this role, we need to examine (i) the relation holding between the discourse segments, or groups of segments, where the adverbial occurs and (ii) the relation holding between the adverbial and its host. The next section outlines a framework for describing those relations.

2.3 Rhetorical Structure Theory (RST)

Rhetorical Structure Theory (RST) (Mann and Thompson 1986, Fox 1987, Mann and Thompson 1987, Mann and Thompson 1988, Matthiessen and Thompson 1988, Mann, Matthiessen and Thompson 1992, Mann 1999) was developed partly in response to the needs of NLP text generation software. In its original or modified form, it has probably been more extensively tested in NLG systems than any other coherence relations framework (v. Oates' (1999) survey). It was devised inductively, to provide a means of describing text structure and in particular the coherence relations (known as rhetorical relations) among discourse segments (known as text spans). At the same time it provides "a functional basis for studying the discourse-relevant specific forms in texts, 'discourse markers' and other formal correlates of discourse structure" (Mann 1999).

RST views text as having three principal kinds of structure, with no sharp boundary between them: **holistic structure** (properties of genre or variety), **relational structure** (the organization of ideas into coherent contiguous text) and **syntactic structure**. RST is a theory of **relational structure**, the level of discourse coherence (Mann, Matthiessen and Thompson 1992: 41).

2.3.1 Some examples of RST relations

Example (2) shows an instance of the relation of CIRCUMSTANCE.

- (2) Probably the most extreme case of Visitors Fever I have ever witnessed was a few summers ago | *when I visited relatives in the Midwest*. (Mann and Thompson 1987: 48-49)

The second text span, *when I visited relatives in the Midwest*, is related to the first in that it expresses the circumstance of the first. This relation is expressed in RST notation as in Fig. 2.1. The direction of the arrow indicates that span 2 is the circumstance for span 1.

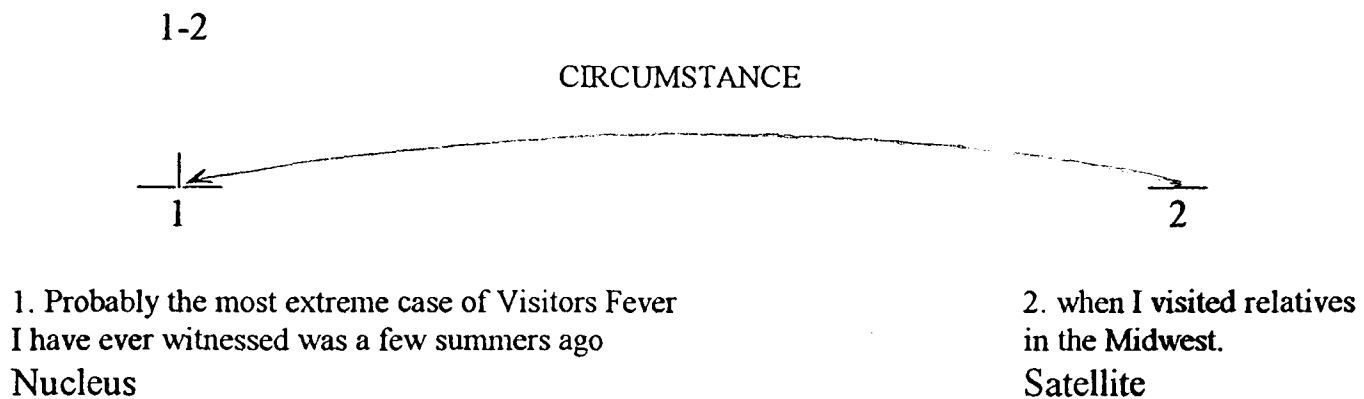


Fig. 2.1. RST diagram for example (2): the CIRCUMSTANCE relation

A distinction is made, in the CIRCUMSTANCE and most other relations, between the Nucleus (N) text span and the Satellite (S) text span. In example (2), the first span is the Nucleus, and the second, *when I visited relatives in the Midwest* is the Satellite. Relations are identified between two spans, but these spans are more often than not asymmetrical in that, for instance, if the second span is the circumstance of the first, the first is not the circumstance of the second. The Nucleus is indicated in the notation by a vertical line.

The CIRCUMSTANCE relation, like many others, is defined in terms of the constraints on Nucleus and Satellite, and the effect of the relation, as in table 2.1.

Relation name:	CIRCUMSTANCE
Constraints on Nucleus (N):	none
Constraints on Satellite (S):	S presents a situation (not unrealized)
Constraints on N + S combination:	S sets a framework in the subject matter within which the reader is intended to interpret the situation presented in N.
The effect:	The reader recognizes that the situation presented in S provides the framework for interpreting N.
The locus of the effect:	N and S

Table. 2.1. Definition of the CIRCUMSTANCE relation (Mann and Thompson 1987)

Text structure is described in terms of 'schemas'. The CIRCUMSTANCE relation above exemplifies the most common type of schema, the mononuclear schema:



Example (3) below, repeated in chapter 4, illustrates the EVIDENCE relation. The corresponding RST diagram is given in Fig. 2.2.

- (3) They are clearly raring to go — it was difficult to get them to sit down long enough to drink their tea (Sp)

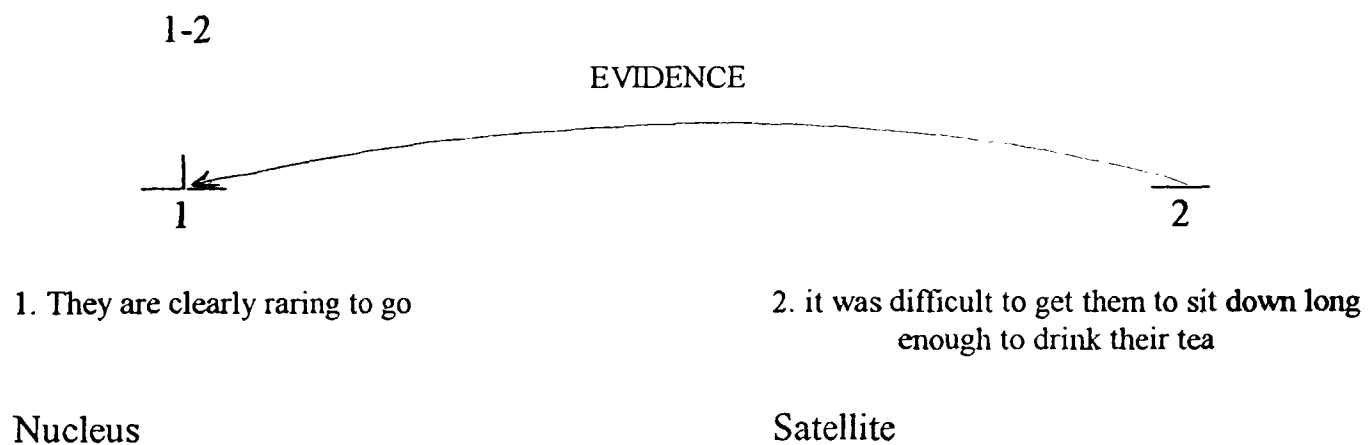
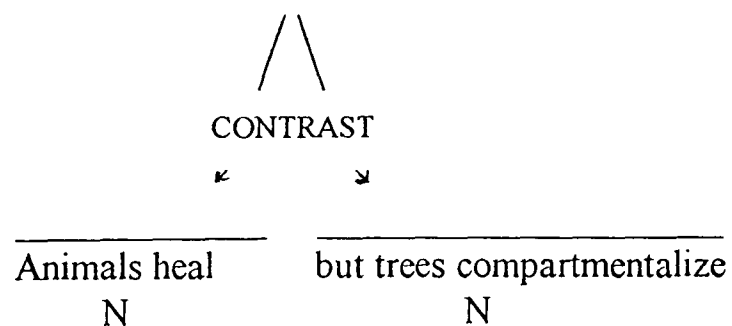


Fig. 2.2. RST diagram for example (3): the EVIDENCE relation.

The idea in span 2, the Satellite, is presented as evidence for the assertion in span 1, the Nucleus. It explains the grounds the speaker has for making the assertion.

Multinuclear relations are those that hold between two or more segments of equal import, where the relation is symmetrical (v. 2.3.3 below). An example is the CONTRAST relation (4).

- (4) Animals heal, but trees compartmentalize. (Mann and Thompson 1987)



Other common relations are: ELABORATION, RESULT, CONCESSION, BACKGROUND, ANTITHESIS, JUSTIFY, SUMMARY, INTERPRETATION, RESTATEMENT. Appendix A gives definitions for these and further rhetorical relations.

Relations can be grouped in a number of ways. One way is by schema type. In particular, it will be convenient for our discussion of the history of certain expressions to group together several relations that share the N-S schema (exemplified above for the CIRCUMSTANCE relation). These are

basically the elaborative-type relations and include ELABORATION, SUMMARY, INTERPRETATION, EVALUATION and RESTATEMENT. These all involve more about the current topic, and, after Halliday (1994: 219ff), will be referred to as Expansion relations.

RST is designed to be able to describe a whole text by a single rhetorical structure tree. An important tenet is that the same relations hold between single ideas as hold between complexes of ideas, so that “the same sorts of relations that characterize clause combining operate at higher levels of text structure” (Thompson and Mann 1987: 359).

Where a topic is announced and a number of points made about it, the following simple schema of ELABORATION, the commonest relation, is often adopted:

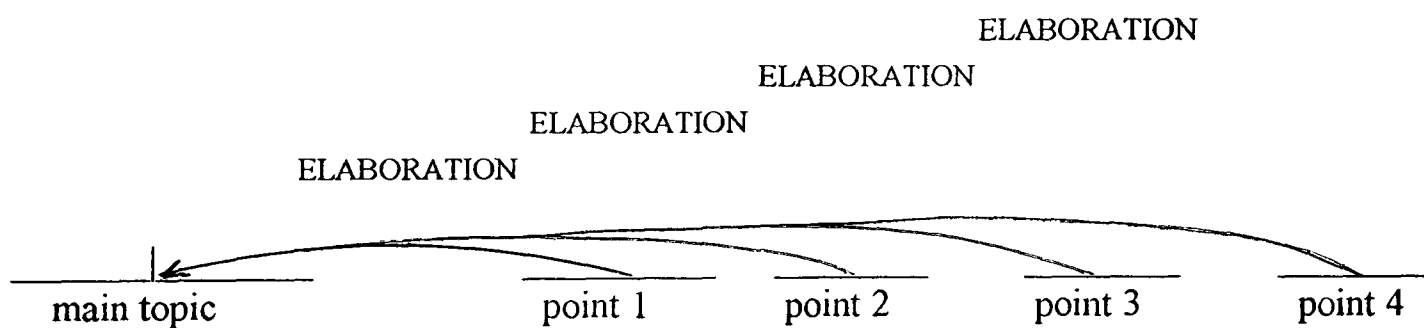


Fig. 2.3. RST schema for multiple ELABORATION

As a tool for the description of discourse coherence, RST relies on the analyst being able to identify not only rhetorical purpose and discourse segments, but, above all, rhetorical relations.

2.3.2 Rhetorical relations and DCs

Rhetorical relations, also known as coherence relations or logical relations, are relations of time, causality, contrast, etc. between ideas. In RST, these relations are viewed as propositions. This follows Grimes' notion of 'rhetorical predicates' whose function is to organize the content of the discourse (1975: 207). Mann and Thompson too argue that the relations between discourse units are themselves propositional, but that they “convey essential subject matter” rather than being limited to textual organization (1986: 76).

Relations hold between parts of a text¹⁰, i.e. between discourse segments or groups of segments (section 2.2 above). Intrasentential relations, e.g. of hypotaxis, are of the same type as intersentential relations and inter-episode relations. Evidence from subordination in a wide range of European languages suggests that there is a set of cognitively basic interclausal relations, the ‘hard core’ of which consists of the relations of Condition, Cause, Similarity, Concession, Place and Simultaneity-overlap (=‘when’) (Kortmann 1997: 207-8)¹¹. RST claims that these same relations provide coherence among multi-clausal stretches of language too.

Natural language generation researchers have been interested in developing taxonomies of relations and mapping these to connective devices (e.g. Sanders et al 1992, 1993, Hovy 1993, Hovy and Maier 1995, Sanders and Spooren 1999, Bateman and Rondhuis 1997, Corston-Oliver 1998, Knott and Sanders 1998). In several research programmes, collections of markers, or ‘cue phrases’, are used as the basis for trying to identify a comprehensive set of relations (e.g., Knott 1995, Knott and Mellish 1996, Harabagiu 1999).

RST, by contrast, identifies relations independently of connective markers. RST emphasizes that relational propositions are largely inferred. There is no necessary correspondence between particular structural elements and rhetorical relations (Mann and Thompson 1987: 19). More often than not, relations have no linguistic expression. Instead, “it is the conjunctive *relationships* which are pervasive, with explicit conjunctions as occasional manifestations” (Mann and Thompson 1986: 89). This gives RST an advantage in diachronic and cross-linguistic studies. It does not obscure the many-to-many relationship between discourse connectives and rhetorical relations. A relatively independent set of relations can act as a *tertium comparationis* for contrastive analysis (e.g. Eng. *but* vs. Ger. *aber* or *sondern*) and for lexical semantic change (e.g. *of course* at time₁ versus *of course* at time₂).

While some relations are almost always linguistically marked (e.g. CONCESSION, CONTRAST),

¹⁰ This is a shorthand manner of expression. Strictly speaking (v. 2.1) relations hold between the ideas expressed by parts of a text: RST describes relations between meanings and intentions, not between chunks of language (Mann, Mattheissen and Thompson 1992: 45).

¹¹ Kortmann’s list of ‘core relations’ is: Cause, Condition, Concession, Result, Purpose, Simultaneity overlap (‘when’), Simultaneity duration (‘while’), Anteriority, Terminus ad quem (‘until’), Place and Similarity (1997: 207).

others are seldom marked (e.g. EVALUATION, ELABORATION, BACKGROUND). Estimates are that on average around 30%-40% of relation tokens are linguistically marked¹². In the examples above, the CIRCUMSTANCE and CONTRAST relations are signalled by *when* and *but* respectively, while the EVIDENCE relation is not obviously marked.

Why are some relations signalled more than others? A possible answer is that only relations involving counter-expectation need to be linguistically expressed. It is often observed that DMs and DCs signal discontinuity, i.e. they are used where the discourse deviates in some way from the expected (Hellman 1995: 199-200, Traugott 1999). They indicate discourse boundaries, topic shifts, focus, degree of communicative salience, newness, grounding, and so on. Their function overlaps with that of other information-structuring devices such as prosody, syntactic constructions like clefting and subordination, and word order. Dorgeloh points out, in the context of inversion, that while unmarked language implies topic continuity and foregrounded material, “relative markedness points to topic shifts, backgrounded material, digressions or turns within the structure of the text” as well as “subjective viewpoint, or change in focus of attention” (1997: 117). Like inversion, DMs and DCs are markers in this sense.

Addressees’ two default assumptions are continuity, i.e. congruence with the previous discourse, and normality, i.e. congruence with experience of the world. According to Ducrot, “le sens d’un énoncé contient une allusion à son éventuelle continuation: il lui est essentiel d’appeler tel ou tel type de suite, de prétendre orienter le discours ultérieur dans telle ou telle direction”¹³ (1980: 11). Expectations are set up in discourse about its continuation in terms of speaker/writer attitude and purpose and thematic development. This is what Lundquist calls the ‘pragmatic continuity of the text’ (1989: 104) and Tannen (1993) calls ‘structures of expectation’ (see also Keller 1994: 92). Types of evidence that reveal the existence of expectations (properly, the speaker’s notion of the addressee’s expectations) include: repetition, backtrack, hedges, negatives, contrastive connectives, modals, inexact

¹² Mann (8.2.00). Message to RST-List. Archived at <http://listserv.linguistlist.org/archives/rstlist.html>.

¹³ ‘the meaning of an utterance contains an allusion to its later continuation: it naturally summons up this or that follow-up, it seeks to orient the discourse in one direction or another’.

statements, generalization, evaluative language. Discourse connectives belong to a “set of cues marking denial of expectation” (Tannen 1993: 39). This would explain why relations such as CONCESSION are much more often marked than, for instance, ELABORATION. Traugott and König (1991: 211) refer to the ‘surprise factor’ in the development of concessives: a marker of simultaneity would not be used if there were not something surprising about that simultaneity, thus the marker of simultaneity comes to index the surprise factor and thus the concessive. The fact that markers are used in counter-expectation contexts more than in ‘consonant’ contexts also suggests that there may be greater qualitative differences among rhetorical relations than are recognized by RST.

The relationship between coherence relation and linguistic form is a many-to-many one. That is, one relation can be indicated by more than one linguistic form, and one linguistic form can serve to indicate more than one coherence relation¹⁴. RST does not seek to account for the nature of the relationship between connective devices and rhetorical relations; it merely provides a framework for identifying and describing relations.

2.3.3 Nuclearity

As seen above (2.3.1), most RST relations hold between a Nucleus and a Satellite. Nuclearity (Mann and Thompson 1987: 31-38) refers to this asymmetry (Grimes’ ‘asymmetric predicates’ (1975: 210). For each pair of spans, the Nucleus (N) is more central to the writer’s purpose, while the Satellite (S) is more peripheral and usually contains supporting material.

As well as by centrality to rhetorical purpose, Nucleus and Satellite are distinguished in two other ways: (1) the removal of the Nucleus may lead to a non-sequitur, while the removal of the Satellite leaves the text coherent; (2) the Satellite, but not the Nucleus, may be replaced by a different proposition without damaging the overall purport of the text. Thus, if only Nuclear material were retained from a text, there would still be a text expressing something akin to the writer’s purpose.

¹⁴ Of course, DCs can encode more than just the type of rhetorical relation that holds between two ideas; they can also encode other types of modality and attitude. As Chafe notes, for instance, “contrastiveness is often accompanied by affect, since a speaker who is expressing a contrast is likely to be emotionally involved in the assertion that it was X rather than Y” (1994: 77).

Good summaries of texts can be produced by selecting Nuclear material (Marcu 1999). If only Satellite material is retained, an incoherent passage results, giving little clue of the writer's purpose. The Nucleus-Satellite distinction of RST is a semantic one: "the identification of nuclei is .. generally a byproduct of recognition that a particular relation holds" (Mann and Thompson 1987: 31). Nuclearity itself is an information structure concept. RST thus incorporates information structure into its definitions of relations.

Information structure is usually described in terms of categories such as focus, theme-rheme, topic-comment, given-new information, foregrounding-backgrounding, etc., though salience is also seen as scalar rather than binary (see especially Givón 1987: 177-9)¹⁵. These concepts tend to overlap. As Tomlin et al. point out, "it has been extremely difficult to develop definitions which are both theoretically satisfying and empirically manageable for basic notions in information management" (1997: 103). Nuclearity invites comparison with grounding, which describes the way some ideas are presented as central (foregrounding) and others as peripheral or supportive (backgrounding). Following Hopper (1979), grounding analysis has mostly been applied to narratives, which can be described in terms of mainline events vs supporting material. But it is equally applicable to other text types. In Tomlin et al's framework (v. 2.1), foregrounding is related to topic (theme) instead of storyline: "the centrality of any given proposition in discourse arises from the intersection of the theme of the discourse at that point ... and the rhetorical goal of the discourse" (1997: 92). What is important is that grounding, like RST schemas, is hierarchical; it is not 'flat' as some analyses of information structure have implied, but involves complex embedding. Texts can therefore be seen as having a single pyramid of information structure, described in terms of nuclearity (or semantically-identified grounding). This is more consistent with a view of coherence as pertaining to every level of the discourse. Such a continuum of coherence matches the notion of a syntactic "continuum of

¹⁵ Some scholars adopt a narrower definition of information structure. For example, Jones and Jones (1979) contrast information structure with thematic structure, associating the former with propositional/referential content and the latter with rhetorical purpose and speaker-hearer relationship. Jones and Jones suggest that one or other of these structures is likely to be dominant in a particular text type: thematic structure in expository prose and information structure in narrative (Jones and Jones 1979: 5). Oberlander and Moore (1999) limit the term 'information structure' to the given/new distinction. Here, by contrast, 'information structure' is used in its broad sense.

bondedness” (Traugott and König 1991: 189). Nuclearity does not account for all text structure, however. As Mann and Thompson point out, it cannot deal with texts where “parallelism is the dominant organizing pattern” (1987: 36).

Some languages have dedicated means of labelling degrees of supportive material (Jones and Jones 1979); but not English¹⁶. Analysis of grounding in English is therefore problematic, because it relies on both syntactic criteria (main clause vs. subordination, participles, gerunds, etc.) and semantic criteria (accomplishment verbs vs statives; temporally-sequential events vs non-sequential, etc.) and these are not always compatible (v. Brinton 1996: 44ff). For example, foregrounding has sometimes been equated with new information, but although they often co-occur, new information can be ‘backgrounded’ by a subordinating device. Conversely, subordinate clauses can be foregrounded if they meet certain semantic criteria (Dry 1983: 32). Oberlander and Moore (1999) show that the given/new distinction is quite separate from the background/foreground distinction, the latter resulting from the ‘intentional structure’ of the text. So where syntactic and semantic criteria clash, it is difficult to decide if any one discourse unit is foregrounded or backgrounded (Hopper and Thompson 1980). Having nuclearity fall out from rhetorical relation analysis alleviates this problem, though in relations such as BACKGROUND there is some overlap¹⁷.

Certain types of information are typically backgrounded. Generalizations, norms, stable states and frequent events tend to be backgrounded, so that expressions indicative of normality or stability become associated with reduced salience. Modal information is typically backgrounded: the adverbial expression of epistemic modality (e.g. *probably*) has been found to be far commoner than the adjectival expression (e.g. *it is probable that*) in which the modality is in focus (Nuyts 1993a). It has been suggested that perhaps modal adverbs cannot be in focus (Nuyts 1993a: 955). Sentence adverbs such as *undoubtedly* function as condensed clauses. Speaker attitude and rhetorical relations are also

¹⁶ Stein (1985), however, suggests that in Early Modern English the alternations (1) *s* vs *th* in 3ps present indicative verbs and (2) *do* vs finite form in declaratives are motivated by discourse structuring requirements and are therefore discourse markers. Stein emphasizes the “historically ephemeral character” of discourse meanings (1985: 300).

¹⁷ It has been claimed that the BACKGROUND relation differs from other relations in that it is never marked and is possibly better described in terms of focus than as a rhetorical relation (Knott (12.1.00 and 13.1.00): messages to RST-list, archived at <http://listserv.linguistlist.org/archives/rstlist.html>).

typically backgrounded. According to Longacre's 7-band 'salience scheme for English narrative', 'evaluation' and 'cohesion' (which include attitudinal and rhetorical meanings) are typically the two least salient types of information (Longacre 1996: 21-9). Thus DCs themselves are more often than not backgrounded ideas – they are sometimes said to operate on an 'implicit' level (Östman 1995:106).

The position of DCs is linked to nuclearity. Oberlander and Moore (1999) report on a corpus study which found that where the order was Nucleus-Satellite ('core-contributor' in their terms), the DC ('cue phrase') was never on the Nucleus ('core'):

	DC on Nucleus	DC on Satellite
N - S order	×	✓
S - N order	✓	✓

Power et al (reported in Oates 1999: 38) find the same.

Moreover, a DC was more likely to occur where there was Satellite-Nucleus order. Oberlander and Moore (1999) suggest this is because Satellite-Nucleus order is harder to process, so a DC is helpful. When the Nucleus comes first, a DC to indicate that the coming Satellite supports it is superfluous. Why, then, is the Satellite-Nucleus order so frequent? According to Oberlander and Moore (1999), it is because the preference for given-new order often conflicts with the preference for Nucleus-Satellite order (whenever given=S and new=N), and the former preference is stronger. Another possibility is the principle of end-focus, whereby, especially in persuasive discourse, more important points come last where they have more impact.

A single linguistic device such as a DC can signal both a rhetorical relation and a grounding relation. We have seen that RST combines these two types of relation such that each rhetorical relation entails a grounding relation. This isomorphism is controversial, as discussed below in 2.3.5.

2.3.4 RST and discourse planes

Rhetorical relations perform not only an informative function but also presentational and social functions; "it is in this sense that an RST structure is 'rhetorical'" (Mann, Matthiessen and Thompson 1992: 45). RST therefore recognizes two types of relation: 'subject-matter relations', which

hold in the world referred to by the text, and ‘presentational relations’, which hold in the minds of the interlocutors (Mann and Thompson 1987: 17-18). The CIRCUMSTANCE relation in example (2) above is a subject-matter relation, while the EVIDENCE relation is presentational. The purpose of a presentational relation is to “increase some inclination in the reader”, rather than to inform the reader of some attribute of the subject matter (Mann and Thompson 1987: 18). The presentational Satellite increases the reader’s willingness to believe the Nucleus. It increases the reader’s ‘positive regard’ for the Nucleus idea. This notion of ‘positive regard’ is essential to the definition of relations such as CONCESSION and ANTITHESIS (Appendix A).

This subject-matter/presentational distinction is another way of looking at discourse planes. RST is in principle, therefore, compatible with Sweetser’s (1990) model. Following on from the discussion in ch. 1.3, it is suggested that relations can exist on the content, attitudinal and speech-act planes. While subject-matter relations function on the content plane, presentational ones function on the attitudinal or speech-act planes, i.e. on a different plane from the discourse segments they relate. Similar distinctions are made by other discourse theories. For Relational Discourse Analysis the subject-matter/presentational distinction is expressed as ‘informational relations’ and ‘intentional relations’ (Moore and Pollack 1992, Moser, Moore and Glendening 1996: 5ff). Some scholars use the terms ‘semantic’ and ‘pragmatic’ (e.g. Sanders et al. 1992: 7-10, Sanders et al. 1993: 99-100, Knott and Mellish 1996: 155).

In principle, rhetorical relations can be mapped to discourse planes: some subject-matter relations to the content plane, most to the attitudinal-rhetorical plane, and presentational relations to the attitudinal-rhetorical or speech-act planes. However, RST has some problems coping with different discourse planes. It cannot represent the fact that the ‘same’ relation can operate at different levels, i.e. on different planes¹⁸. The drawback of this is some loss of parsimony. A more serious problem is that

¹⁸ To reflect the fact that a relation such as CAUSE can hold on different planes, one could use a single set of labels and indicate the plane as appropriate. But what RST does is to give different labels to subject-matter and presentational relations. For instance, subject-matter contrast is called CONTRAST, while presentational contrast is called ANTITHESIS. *But* can signal either. *Because* can indicate the subject-matter relation CAUSE, as in ‘the dog ate the meat *because* it was hungry’, the presentational (attitudinal-evidential) relation EVIDENCE, as in ‘they must be at home *because* the light’s on’, or the presentational (speech-act) relation JUSTIFY, as in ‘is that your newspaper? *because* I haven’t seen the news today yet’. This method obscures the relatedness of the three relations. We shall adhere to the RST nomenclature as far as possible, however, not to multiply schemes.

of overlapping relations. Where more than one relation on the same plane holds between two spans, RST recognizes this as ‘overlap’ (Mann and Thompson 1987: 28) and allows for two labels to be applied. However, overlap where the two relations are on different planes leads to difficulties. Because nuclearity is tied to the relation, it is possible to have overlap of a subject-matter and a presentational relation, whose Nucleus-Satellite structures are incompatible. This is because while the mapping between speaker/writer intention (Nucleus-Satellite structure) and presentational rhetorical relation is one-to-one, the mapping between speaker/writer intention and subject-matter rhetorical relation is many-to-many¹⁹.

Mann and Thompson’s solution to these cases of multiple analysis is that there will always be a ‘preferred analysis’; one relation will be stronger than the other (1987: 30). Others have proposed that separate, parallel structures of subject-matter (informational) and presentational (intentional) relations are needed (Moore and Pollack 1992, Moore and Paris 1992). Moser and Moore (1996) suggest that nuclearity be removed from RST subject-matter relations altogether and incorporated with presentational relations into an intentional structure²⁰. The solution adopted here, where the issue arises, will be to treat cases of conflicting intentionality as ambiguous. This is important when considering both the polysemy of connectives and their diachronic semantic shifts, since they may express more than one relation in certain contexts or at certain times.

2.3.5 Applying RST to the analysis of modal adverbial contexts

The aim of RST is to describe a given text as a single rhetorical structure in which each discourse segment enters into a rhetorical relation with its adjacent segment(s). One advantage of RST is therefore that it provides a way of describing the relations among all the discourse segments of a text, whether or not the relations are linguistically signalled. This facilitates a better understanding of the role of discourse connectives where they occur. RST is “a useful framework for relating the

¹⁹ This explains why RST has to have separate relations of CAUSE and RESULT for what is logically a single relation.

²⁰ Grosz and Sidner (1986) use the term ‘Intentional Linguistic Structure’ (ILS) to refer to the way speakers’ intentions determine the structure of the discourse. This structure is not addressed explicitly by RST, although it is implicit in RST’s definition of nuclearity.

meanings of conjunctions, the grammar of clause combining, and non-signalled parataxis" (Mann and Thompson 1987: 2).

A further advantage is that it is discourse-oriented: it takes account of the global and local rhetorical purposes of the writer/speaker and these are reflected in the tree-structure that results from an RST analysis. Being recursive, RST is comprehensive in its coverage of the information structure hierarchy: every segment is linked to one or more other segments. Lower level structures can be seen to be a part of, or motivated by, higher-level information design requirements. Long-distance relations are clearly brought out.

RST poses two problems of application for the analysis of the discourse contexts of modal adverbials. First, it was developed for the analysis of written, not spoken language. Second, it is ambivalent about what type of relationship exists between rhetorical relations and discourse connectives.

RST was developed specifically for written monologue and has never been extended by its creators to spoken dialogue, although Ford (2000) uses RST to analyse recorded spontaneous conversation. There is no reason in principle why it should be unsuitable. Indeed, according to Miller and Weinert, it should be apt, because it "offers a way of handling relationships between clauses without the clauses being arbitrarily grouped into sentences by the analyst" (1998: 99). But its constraints, especially that of completeness, need to be relaxed to accommodate the spontaneity of conversation. Conversation is open-ended and unlikely to have an overarching purpose, or a 'comprehensive locus of effect' (Mann, Matthiessen and Thompson 1992: 61). In spoken language, several discourse threads may be spun concurrently, so that interweaving analyses may be needed. Moreover, there are frequent ellipses, interpolations, false starts and interruptions, so that in RST terms, there will be many broken and incomplete rhetorical structures.

Fox argues that conversation analysis is the best model for analysing spoken data, while RST is the best model for written data, on the grounds that spoken and written language have fundamentally different structuring units: primarily interactional ones in spoken language and informational ones in

written language (1987: 3-4). However, to focus solely on informational units in written language and on interactional units in spoken language makes it impossible to compare the use of linguistic expressions across spoken and written language. Both types of language have informational and interactional aspects (see 2.2.1). Chafe also argues that RST is unsuitable for spoken language because spoken language is produced “one idea at a time” and is not susceptible to hierarchical modelling (1996: 54). But spoken conversation does display local level hierarchical rhetorical relations, and broken and incomplete rhetorical structures can be allowed for.

The second problem is that the relationship between rhetorical relations and the linguistic expression of them (‘signals’) is not specified by RST. As noted above (2.3.2), there is a many-to-many relationship between linguistic signalling devices and rhetorical relations. However, it is not clear what counts as a ‘signal’. In practice, it has been assumed that signals were lexical: co-ordinating conjunctions, subordinating adverbs, etc. But prosody, syntactic structure, word order, tense, mood and other means may signal relations too. All these ‘signals’ naturally influence the analyst’s identification of relations, but this issue is not addressed by RST. It is one of the aims of this study to clarify the relationship between particular adverbs and rhetorical relations.

RST is not a fully-fledged theory of discourse structure or of discourse coherence. It has little to say about topic development and topic progression for instance, and does not deal with focus or reference. But it is a useful tool for the identification of rhetorical relations. Its advantages are that it is not a closed taxonomy but an open set (Mann, Matthiessen and Thompson 1992: 46), and is therefore flexible. It does not have too many relations, though being open-ended, more can be added. It has been tested by its developers and others on hundreds of real texts, and has had successful trials in NLG.

There are very many other schemes for describing coherence relations. Bateman and Rondhuis (1997) contains an overview and comparison of several of them. Some, especially the feature-based approaches, involve very large numbers of theoretically-possible relations; others, such as RST, are more coarse-grained (see Hovy and Maier 1995). There is broad compatibility between classical RST (Mann and Thompson 1987) and Relational Discourse Analysis, based on Grosz and

Sidner's (1986) framework. A hybrid of the two is emerging from current research.

2.3.6 Rhetorical patterns

Chapters 5-8 examine the co-occurrence of modal adverbs *at least*, *in fact*, *after all* and *of course* with rhetorical patterns. A rhetorical pattern is a typical linguistic sequence comprising one or more rhetorical relations and often one or more of a small set of linguistic markers associated with the relation(s). It may be compared to the 'constructions' of construction grammar (Kay 1997, Fillmore et al 1988, Goldberg 1995) or to Barlow and Kemmer's (1994) 'schemas'. Schemas are "abstractions over actual instances of language use" (1994: 25) and they "come to represent collocational links" (1994: 21). RST-type relations hold at sub-clausal, clause-complex and discourse levels. Rhetorical patterns are a kind of discourse-level collocation. Just as "a major schema is the formal correlate of a prototype semantic structure" (1994: 29), we suggest that, with the caveat that structures at discourse level are looser, rhetorical patterns are the correlates of prototypical discourse structures. That is, from a discourse perspective, it is possible to view language constituency as forming a hierarchy, from lexical item to rhetorical unit, in which the units/components are more loosely associated and less predictable towards the discourse pole, but in which collocation plays a role at every level. As Chafe puts it,

understanding depends on the ability to place ordinary, particular experiences, whether they are derived through the senses or through events within ourselves, within some larger picture where they 'make sense'; that is, where they are recognized as at least partly familiar, and where expectations already partly established can provide ways of reacting to them and interacting with them. (1992: 81)

Rhetorical patterns are abstractions from repeated instances in use. New instances that can be mapped to these patterns are therefore 'partly familiar', so that they can be represented as instances of a type, and thereby acquire a type meaning. That is, such new instances are not straight compositions of the meanings of their parts, but carry the additional meaning associated with the quasi-conventional pattern.

Lexical semantic change caused by regular collocation of one item with another — what

Bréal aptly described as ‘contagion’ (1897, ch.XXI) — has long been recognised. But discourse level collocation has received less attention in studies of language change. It will be argued that the semantic and syntactic changes shown by some S-AdvS are conditioned by their recurrence in particular constructions and/or rhetorical patterns.

To summarize, information structure at discourse level is hierarchical. It is tied more to syntactic structure at ‘lower’ levels and more to discourse structure at ‘higher’ levels. As Levinson points out, much syntax seems to be concerned with “meshing sentence-construction with pragmatic principles: for example, for ‘foregrounding’ and ‘backgrounding’ informational content” (1983: 41). At ‘higher’ levels, where syntactic dependencies are weak or non-existent, connective devices (lexical items, phrases and clauses) are one means of meshing discourse-construction with rhetorical principles.

It has been seen that the same form may occur on different discourse planes (note 18). Diachronically, content plane uses tend to be the earliest, attitudinal and speech-act uses deriving from them in a manner akin to grammaticalization. The next section considers whether grammaticalization theory can shed light on discourse connectives. It is often argued that “grammatical structures arise diachronically under pressure from information-structure constraints” (Lambrecht 1994: 29). This is a claim that information structure is a force for grammaticalization. Matthiessen and Thompson, for instance, argue that “hypotactic clause combining is best understood as a grammaticalization of the Nucleus-Satellite relations” (1988: 317)²¹. Perhaps connective development can usefully be seen as grammaticalization resulting from rhetorical pressures too.

2.4 DCs and grammaticalization

2.4.1 *The development of DCs as instances of grammaticalization*

It is often claimed that discourse markers are created by a process of grammaticalization (e.g. Ariel 1994: 3252, Traugott 1995b, Brinton 1996: 272). Several studies of the development of

²¹ But see Harris and Campbell (1995) chapter 10 for arguments that hypotaxis did not develop from parataxis.

particular English DCs have been carried out in the grammaticalization framework. Finnell (1989) traces the history of *well* in terms of a move from content-plane to attitudinal meaning. Thompson and Mulac (1991) and Aijmer (1997) claim that *I think* is becoming an epistemic parenthetical that functions very like a discourse marking adverbial by a process of grammaticalization. Traugott (1995b), Schwenter and Traugott (2000) and Traugott (in press) argue that the developments of clause connectives and discourse markers such as *indeed*, *in fact*, *besides*, *instead of* are cases of grammaticalization, being “consistent with prototypical grammaticalization in its early stages” (1995b: 21). Brinton (1996) concludes that the historical developments of a range of pragmatic markers in English, such as *gan*, *anon*, *I gesse*, follow the recognized process of grammaticalization, most saliently that of decategorialization (Brinton 1996: 273). Pallander-Collin (1996, 1997) argues that *methinks* / *methought* grammaticalized into an evidential sentence adverb. This section considers whether indeed the phenomena said to be diagnostic features of grammaticalization also characterize the development of DCs out of VP adverbials.

Grammaticalization theory is relevant to the above expressions, and to the emergent English DCs analysed in this study, in that these expressions have undergone syntactic, semantic/pragmatic and prosodic changes of a type which are often associated with grammaticalization. Before considering the nature of these changes, however, it needs to be emphasized that the attribution of grammaticalization to the development of English DCs is controversial, not least because current definitions and characterizations of grammaticalization are so disparate. There follows, therefore, an outline of the significant changes that have occurred recently in grammaticalization theory and research and of some current views on what constitutes ‘grammaticalization’. Then the types of change are considered, followed by a look at the striking polysemy of DMs in the light of the grammaticalization cycle.

‘Grammaticalization’ traditionally refers to the development of grammatical material out of lexical material. The term²² is attributed to Meillet, who defines it as “l’attribution du caractère

²² The term ‘grammaticization’ is also in current use. The two terms are used interchangeably.

grammaticale à un mot jadis autonome” (1912/1958: 131). A definition that is sometimes regarded as classic is that of Kurylowicz, who described it as follows: “Grammaticalization consists in the increase of the range of a morpheme advancing from a lexical to a grammatical or from a less grammatical to a more grammatical status, e.g. from a derivative formant to an inflectional one” (Kurylowicz 1965: 69)²³. An example is the development of French *je chanterai* from Classical Latin [*cantare*] *habeo* via a Late Latin reanalysis²⁴ as [*cantare habeo*] (Benveniste 1966/1974: 131-33; Hopper and Traugott 1993: 42-44). Grammaticalization theory developed from attempts to elucidate and explain regular types of language change, including those observable across languages²⁵. The roots of grammaticalization theory can be traced back to the work of ancient Greek and Arabic linguists (Harris and Campbell 1995: 16-20). The upsurge of interest among nineteenth century scholars such as Schlegel, Bopp, von Humboldt and Gabelenz in comparative Indo-European philology produced a new exploration of pathways of language evolution. The findings of these scholars, based on historical data from the Indo-European languages, suggested that some lexical items weakened and turned into grammatical affixes. This process was seen as continuous and was likened to a spiral²⁶.

Grammaticalization has received increased attention in the last two decades (Lehmann 1982/1995, Heine and Reh 1984, Traugott and Heine (eds.) 1991, Heine et al. 1991, Hopper and Traugott 1993, Bybee et al. 1994, Pagliuca (ed.) 1994, Ramat and Hopper (eds.) 1998). It has become associated with functionalist and cognitive views of language change and with typological studies. Rather than being a type of language change, it is more often said to consist of a cluster of distinct changes: “grammaticalization is a subset of phenomena occurring in change” (Traugott and Heine

²³ One of Kurylowicz’s examples is the development of the English construction ‘have + past participle + direct object’ (as in ‘to have written the book’), which initially was ambiguous between ‘result of a previous action’ (‘have [written the book]’) and ‘previous action with present result’ (‘[have written] the book’). Once the latter meaning became the default reading, the construction spread to all classes of verb (Kurylowicz 1965: 58-9).

²⁴ ‘Reanalysis’ in the sense of Langacker (1977: 58): “change in the structure of an expression or class of expressions that does not involve any immediate or intrinsic modification of its surface manifestation”.

²⁵ Brief histories of the approach now known as grammaticalization theory can be found in Harris and Campbell 1995: 16ff, Lehmann 1982/1995: 1-8, Hopper and Traugott 1993: 18-30 and Heine et al 1991: 5-23.

²⁶ The inspiration for these nineteenth-century scholars apparently came from observations by Condillac (*Essai sur l’origine des connaissances humaines*, 1746) according to which “verbal inflections such as tense suffixes are historically derived from independent words” (cited in Heine et al (1991: 5)).

1991: 7). But there is little consensus among scholars on which of these phenomena belong to grammaticalization, nor on how many or which of them need to co-occur for a case to be considered to constitute a grammaticalization. The concept is therefore rather loose. Kurylowicz's definition now seems inadequate in that it begs questions concerning 'grammatical status', it focuses on the morphological level and it implies that grammaticalization concerns only lexical or word-level items. Until recently, grammaticalization research did focus almost exclusively on tracing the development of grammatical affixes out of lexical material: case affixes that have descended from nouns, aspect markers from verbs, and so on (e.g. Lehmann 1982/1995, Heine and Reh 1984, Bybee et al 1994)²⁷. But three important developments have occurred in the last decade which have shifted the emphasis and made grammaticalization more relevant to the study of discourse connectives. (i) Research on word order change and lexical change found significant similarities with the development of grammatical affixes, particularly in the semantic changes. There are now wider definitions of grammaticalization, as research has broadened from its focus on morphology to include these phenomena and to examine the early stages of grammaticalization. (ii) As a result of this more encompassing view of grammaticalization, the morpho-syntactic changes have come to be seen more in terms of reanalysis and decategorialization than of fixedness and morphological fusion. (iii) More emphasis is now placed on the semantic changes such as increase in abstraction and subjectivity. Despite its name, grammaticalization research is now as much concerned with semantic and pragmatic change. Alternative definitions of grammaticalization have been mooted recently. Four of these are outlined below.

²⁷ In addition to the lexeme>affix line of research, is a discourse>syntax approach (e.g. Givón 1979, Haiman 1985) which seeks to show that syntax evolves out of discourse practise. One view is that these are two competing approaches to grammaticalization (Heine et al 1991: 238-9). Another is that these are independent paths of change (e.g. Herring 1991: 255). For others, the two are related: Harris and Campbell note that both involve reanalysis, the essential feature of grammaticalization (1995: 92). Hopper goes furthest in viewing the two as aspects of a single phenomenon. He claims that the lexical/etymological approach and the discourse/textual approach complement one another: "the first explains *what* is grammaticalized and the second *how* this occurs" (1996: 232).

2.4.2 *Recent approaches to characterising grammaticalization*

Lehmann suggests a set of six parameters, which he also calls criteria, of grammaticalization (1982/1995: 123ff). He emphasizes the movement from autonomous lexical item to bound morpheme and the movement towards closed-class categories. These two parameters he calls ‘cohesion’: “as grammaticalization increases, the parameters of cohesion increase as well” (1982/1995:124).

Lehmann’s focus is on the emergence of grammatical affixes. Other linguists have suggested that this is only one, final stage of a much wider phenomenon.

Hopper (1998) suggests that the data typically addressed by grammaticalization studies (i.e. the change from lexical category to grammatical category) belong in fact to part of a wider process, for which the term ‘grammaticalization’ may not be appropriate. Often only a part of the process takes place, but the characteristics are the same regardless of whether grammatical affixes emerge.

Cristofaro also argues for extending the term to cover wider patterns of change: “the development of complementizers and subordinators in general out of full lexical items ... offers some evidence that grammaticalization must be considered as a broad parameter pertaining to the reanalysis not only of lexical as grammatical material, but also of discourse patterns as grammatical patterns and of discourse-level functions as sentence-level, semantic functions.” (1998:60).

Hopper, in particular, has drawn attention to how incipient grammaticalization is to be recognised. He proposes a set of ‘principles’ of grammaticalization as follows (1990, 1991).

‘Layering’ (or ‘accumulation’), whereby, within a broad functional domain, older layers co-exists with newer ones; ‘divergence’, whereby the original lexical form remains lexical, while the new form becomes grammatical; ‘specialization’, whereby the variety of choices narrows to smaller number of forms with more general grammatical meanings; ‘persistence’, whereby the lexical history of a grammaticalized element continues to impose constraints on its distribution; and ‘de-categorialization’ whereby an item changes from a primary to a secondary syntactic category. Like Lehmann’s, Hopper’s principles “characterize aspects of change in general, and are not distinctive for grammaticalization” (1991: 21). They are meant to identify areas of potential grammaticalization. Early

grammaticalization, according to Traugott, is a “complex set of correlated changes”. These are: structural decategorialization, entering a closed set (a functional category), and semantic shift from more referential to less referential via context-induced inferencing (Traugott in press). This suggests a view of grammaticalization as a label for a set of coinciding changes rather than for a unified process.

More recently, in an attempt to delimit the field of study, Tabor and Traugott (1998: 235) have proposed three ‘hallmarks’ of grammaticalization. These are: morphosyntactic change, pragmatic/semantic change and gradualness. The hallmarks cover a very wide range of changes (only coinages, borrowings and lexicalizations seem to be excluded), so that by these criteria grammaticalization becomes almost synonymous with (internal) language change. Indeed, it may be that most such change can be attributed to grammaticalization factors (Haspelmath 1998, Croft 2000: 156).

A fourth recent redefinition of grammaticalization emphasizes the roles of the immediate contexts in which grammaticalizations take place. That is, it is increasingly recognised that often it is the construction in which a lexical item is set which grammaticalizes, rather than the lexical item itself (Bybee in press; Traugott in press). This idea is already mooted in Lehmann (1991, 1992), who suggests that “phraseology is at work in grammaticalization” (1991: 503) and that grammaticalization seizes “the whole construction formed by the syntagmatic relations of the element in question” (1992: 406). A new periphrastic expression, for instance, does not immediately spread to all the contexts that it grammatically could, but tends to remain in particular collocations. Gronemeyer’s diachronic study of English *get* “provides strong evidence that grammaticalization affects constructions and not isolated items” (1999: 35). The current shift towards viewing constructions rather than lexical items as the units that grammaticalize owes much to recent work in construction grammar. Traugott (in press) likewise emphasizes the role of constructions and emphasizes that “the development of lexemes in context-specific constructions ... potentially expands the boundaries of what is often considered grammaticalization”.

In sum, there is no consensus on how many or which of the changes associated with

grammaticalization are necessary and sufficient conditions for the label grammaticalization to be used. Tabor and Traugott's (1998) 'hallmarks' are rather vaguely defined. Other recent, broadened definitions of grammaticalization also risk turning it into a synonym of internal language change in general. If there is a common denominator of the various approaches, it seems to be the co-occurrence, not necessarily simultaneously, of semantic change with morpho-syntactic reanalysis and gradual generalization, plus, finally, some phonological change.

2.4.3 Sentence-based vs discourse-based grammaticalization

These disparate views of the nature of grammaticalization have led to the suggestion that there are two paths of grammaticalization: the sentence-level, lexical development of grammatical affixes from lexical material, and the discourse-level, pragmatic development of grammatical patterns and clause linkage strategies. The mechanisms of change might be different, it is suggested: predominantly metaphorical extension in the former case, and conventionalization of implicature based on frequency in the latter (Traugott and König 1991: 190 and 213). But for others, the similarities are sufficient to unite them under a broader definition of grammaticalization, since both involve the expression of more abstract domains in terms of more concrete ones, and the coding of grammatical relations on the basis of context-dependent strategies (Cristofaro 1998:61). Brinton proposes evidence from the development of pragmatic markers in Old and Middle English in favour of this single wider view of the grammaticalization process (1996: 272).

The view taken here is that the two overlap. Recent studies within grammaticalization framework of changes in comment expressions (e.g. Matthiesen and Thompson 1988, Thompson and Mulac 1991, Traugott 1995b, 1998, in press, Brinton 1996) suggest that sentence-based and discourse-based grammaticalization are different, but share some important characteristics. The main differences are the absence from the latter of features associated with the final stages of grammaticalization, viz. phonological reduction, mutation into formatives such as clitics and affixes, reduction in scope and entering into a paradigm.

In the case of the discourse markers analysed in chapters 4-8, there is little phonological reduction or phonetic erosion (although there are prosodic modifications), there is no coalescence (although there is ‘binding within the phrase’ (Hansen 1998a: 225)), nor affixation nor loss of scope. The chief characteristics of the development are gradual semantic change (abstraction and subjectification), reanalysis (involving increase in scope), generalization and often the emergence of a polysemy structure. It is therefore arguably an instance of discourse-based grammaticalization, within the broad view of grammaticalization.

The main morpho-syntactic changes claimed as grammaticalization criteria are reanalysis, decategorialization (change to a more closed, more grammatical syntactic category), and coalescence (morphological bonding). Thus, the tendencies for two constituents to be reanalysed as one, for lexical expressions to be recategorized as functional, and for words to become clitics or bound morphemes characterize the morpho-syntactic aspect of grammaticalization.

Reanalysis, the major mechanism of morpho-syntactic change associated with grammaticalization²⁸, “changes the underlying structure of a syntactic pattern and ... does not involve any modification of its surface manifestation” (Harris and Campbell 1995: 50). In its narrow sense, reanalysis involves interpreting a string as projected from a new hierarchical structure, i.e. a re-bracketing. Reanalysis is said to lead to grammaticalization when it works in the direction of less grammatical towards more grammatical.

Following reanalysis, the new analysis tends to be generalized to more contexts, largely by analogy²⁹. While reanalysis is discrete, the generalization process is more gradual: “thus syntactic change has both discrete and gradual aspects” (Harris and Campbell 1995: 49).

Decategorialization involves an expression changing from a less grammatical category (such

²⁸ There is some disagreement over whether reanalysis is in fact necessary to grammaticalization: Heine et al (1991: 219), Fischer (1997: 176-7) and Haspelmath (1998) all suggest that grammaticalization can take place without reanalysis, though the argument hinges largely on what constitutes a category change and whether a category change is a reanalysis. On the other hand, many scholars define reanalysis more broadly to include semantic change too (e.g. Hopper and Traugott 1993: 32; Brinton 1996: 53-5; Forston in press).

²⁹ This phenomenon is variously termed ‘generalization’, ‘actualization’, ‘extension’ or ‘analogy’. To avoid confusion with semantic phenomena such as metaphorical extension, the term ‘generalization’ will be preferred.

as noun or verb) to a more grammatical category (such as pronoun, adposition, auxiliary or conjunction). It often involves loss of status as an independent word and coalescence into a grammatical affix.

Reanalysis, decategorialization and coalescence are all relevant to the development of DCs, as will be seen in the discussions of individual expressions in chapters 5-8. All these expressions involve reanalysis from VP-Adv to S-Adv. Reanalysis is often conceived of as discrete, especially by formal linguists. For example, Roberts' 'diachronic reanalysis' is a local change (usually triggered, according to Roberts, by a more global parametric change) whereby "a string is assigned structure S at period P and structure S¹ ≠ S at period P¹" (1993: 228). However, reanalysis is gradual in two senses. First, what Harris and Campbell call the 'basis of reanalysis' (1995: 72) is a syntactic pattern that has "the potential for multiple structural analyses" (1995: 72), and reanalysis involves multiple analysis of that pattern. In any one instance, in fact, there may be a dual analysis (Hankamer 1977), where both interpretations are made simultaneously.³⁰ Second, the new analysis starts in certain pragmatic contexts and then may extend gradually to others. It is, after all, occurrences that are interpreted, i.e. analysed, not abstract patterns. Indeed, "each instance redefines the system, however infinitesimally, maintaining its present state or shifting its probabilities in one direction or the other" (Halliday 1991: 34). In these senses, then, reanalysis can be described as gradual.³¹

Nevertheless, while semantic change is fine-grained and appears gradual, reanalysis appears sudden. It seems likely that meaning change can occur before, with and after reanalysis (in its narrow sense of re-bracketing). A reanalysis can only be seen to have taken place once the item starts to be used in contexts where only the new analysis makes sense. Heine (1999) calls the stage prior to this a

³⁰ Hankamer (1977) is an early proponent of multiple analyses. He suggests that "in the quite frequent cases where the analysis ... is underdetermined by the available data, the mind of the speaker does not, at least not always, choose one at random" (1977: 584). He claims further that "one of the most instrumental factors in syntactic change is the existence of multiple analyses in the grammars even of adult speakers" (1977: 601).

³¹ Haspelmath (1998) claims that reanalysis and grammaticization are disjunct phenomena, respectively abrupt and gradual. He argues against the 'gradual actualization' of a reanalysis idea (Harris and Campbell 1995: 176-82) on the grounds *inter alia* that (i) it is implausible that speakers have two analyses of a reanalysed-but-not-yet-actualized construction - such an ambiguity could not be sustained, and (ii) it is implausible that a reanalysis that is only partially actualized (one that is not overtly manifested) could spread. He does subscribe to the idea that an expression can gradually become a member of a different word class. However, Haspelmath does not seem to take account of the type/token distinction: reanalysis occurs in context, not in abstraction. There is no reason why long periods of ambiguity cannot be sustained. It is not a form that is reanalysed, but an occurrence of that form.

‘bridging context’ stage. A ‘bridging context’ is a context that gives rise to an inference that becomes foregrounded *in that context*. It therefore suggests a new meaning. From there it is a short step to a ‘switch context’, i.e. a new context in which only the new meaning makes sense. Once a switch context is introduced, the new meaning is autonomous (semanticized) and no longer needs the support of the context that originally gave rise to it. This phenomenon can be seen in Contrastive *after all* (chapter 5) and in Evaluative *at least* (chapter 7), where it is no longer necessary for any contrasting idea to be recoverable from the context.

During the ‘bridging context’ phase, slow meaning change can be taking place via the gradual reweighting of implicatures. However, it has just been seen that there is a sense in which reanalysis is gradual too: there may be long periods where a dual interpretation of a construction is probable. The data discussed in chapters 5-8 suggest that there can be prolonged periods of latent structural ambiguity. Only careful consideration of the contexts of use can shed light on the probable duration of this ambiguity, that is on the likely times at which the new interpretation emerged and at which it later eclipsed the old in the relevant contexts. The data also suggest that the boundary between (gradual) extension and (discrete) reanalysis depends largely on the theoretical stance adopted.

2.4.4 Semantic / pragmatic change

Lexical semantic change tends towards greater abstraction and greater subjectification. It can usefully be seen as potentially three-stage: (i) extension (by metonymy and/or metaphor) to new contexts, with the resulting increase in frequency possibly leading to loss of pragmatic force; (ii) semanticization of the new implicatures generated in these new contexts; (iii) development of a polysemy structure with possible eventual loss of the older meaning. This three-stage model is applicable to the development of the English modal adverbs discussed in chapters 4-8.

Subjectification can be modelled as a number of clines or dimensions along which change can take place and each of which involves a move towards greater subjectivity. Traugott calls them ‘correlated diachronic continua’ (1995a: 47) and suggests the dimensions in fig. 2.4.

Propositional function	→	Discourse function
Objective meaning	→	Subjective meaning
Non-epistemic modality	→	Epistemic modality
Non-syntactic subject	→	Syntactic subject
Syntactic subject	→	Speaking subject
Full, free form	→	Bonded form

Fig. 2.4. Subjectification in grammaticalization (Traugott 1995a)

These dimensions, according to Traugott, characterize semantic change in grammaticalization. The first three of these are relevant to the development of discourse connectives from P+N adverbials.

Abstraction is often associated with ‘bleaching’, or loss of semantic content. Bleaching is the type of semantic change usually associated with grammaticalization. With increased frequency of use, and increase in the contexts in which a lexical item occurs (‘generalization’), the ‘strength’ of the item is said to diminish. The bleached meaning is seen to be more abstract. Bleaching is said to occur in discourse marker development (Ariel 1994:3251). Other scholars are less committed to the centrality of bleaching to grammaticalization. Indeed, Traugott and König suggest that “bleaching and grammaticalization must be uncoupled if we are to understand the semantic-pragmatic processes of early stages of diachronic grammaticalization” (1991: 190).

The main mechanisms of semantic change in grammaticalization are said to be metaphorical extension (e.g. Claudi and Heine 1986, Willett 1988, Sweetser 1990) or conventionalization of implicature (after Geis and Zwicky 1971; e.g. König and Traugott 1988; Traugott 1999)³², also known as metonymic extension, or ‘conceptual metonymic change’ (Schwenter and Traugott 1995: 264). It will be argued that conventionalization of implicatures operates in the case of DC development. That is, co-occurrence of an expression with certain rhetorical contexts leads to the expression being interpreted as performing a particular discourse function and this function eventually becomes a part of its conventional meaning. As Levinson puts it, “certain linguistic expressions will tend to be associated with specific pragmatic inferences across a broad range of contexts, so that these associated inferences can be predicted in a systematic way, and play a systematic role in shaping patterns of

³² For discussion of these and other mechanisms of semantic change, see Bybee et al 1994, chapter 8.

lexicalisation and grammaticalisation” (1995: 93). These inferences he calls ‘utterance-type meanings’ and posits a cline from nonce-inference to conventional semantic meaning, with “a diachronic path from speaker-meanings to utterance-type meanings to sentence-meanings” (1995: 95). These utterance-type meanings, or default inferences, can be robust, and the diachronic process can be lengthy. It will be seen in chapter 6, for example, that the epistemic inferences associated with *in fact* from the turn of the C18th continue to be defeasible well into the C19th.

According to Bybee “it is possible that new occurrences arise on the basis of very local analogical processes” (1998: 269). The extension of the phrases *at least*, *in fact*, *after all* and *of course* to epistemic modal and discourse functional senses exemplifies this local, fine-grained nature of much semantic change. The view that DC development involves shifts from the content plane to the attitudinal and/or speech-act planes suggests metaphorical leaps across distinct conceptual domains. As noted in 1.3, however, the domains interact. Moreover, comment expressions cover a range of more or less subjective, more or less abstract ideas. The data on DC development do not support the idea of sudden metaphorical applications of content expressions to new, metatextual contexts, but rather point to generalized metatextual implicatures being firmly in place before such extensions take place.

Another phenomenon posited in grammaticalization studies is ‘pragmatic strengthening’ (König 1988, Traugott 1988, Traugott and König 1991, Brinton 1996), whereby an expression acquires new implicatures and seems to gain in salience. This is linked to a strand of research that has sought to show how pragmatic discorsal features become structural grammatical features: the ‘discourse > syntax’ model. Du Bois describes grammaticalization as “a prominent mechanism for imbedding pragmatic dimensions within linguistic structures” (1994: 3259). However, no clear criteria for establishing semantic content or its absence in discourse markers have been established, and the effect ‘pragmatic strengthening’ is intended to describe can be accounted for in terms of degree of subjectification, or of degree of referentiality or context-dependence.³³

³³ Some scholars distinguish ‘pragmaticalization’ from ‘grammaticalization’, the former being recruitment of lexical items to express speaker attitude (such as discourse markers and modal particles), the latter the derivation of grammatical forms (such as tense and mood markers); although sharing many features of change, the two modes are distinguished by the truth-conditional criterion (Aijmer 1997: 3, Amovick 1999: 117).

A correlate of the morpho-syntactic and semantic changes is often phonetic erosion and sometimes phonological reduction and/or coalescence. This may be the result of frequency, or of reanalysis of two constituents as one. The changes associated with the development of S-Advs and DCs from VP-Advs tend to be prosodic, with new stress or intonation patterns developing for an expression being used in a new context. In particular, expressions such as *at least* or *in fact* tend to develop their own tone group in some uses.

2.4.5 *The grammaticalization cycle*

Controversy surrounds the order of morpho-syntactic and semantic changes and the role of frequency in grammaticalization. For some scholars, morpho-syntactic change occurs first, triggering semantic change. Gronemeyer, for instance, in her study of the grammaticalization of English *get*, suggests that syntactic change precedes semantic change (1999: 37). Hock and Joseph (1996:194-6) argue that the developments that gave rise to modal *hopefully*, *presumably*, *actually* were syntactic, on the grounds that they were probably driven by analogy with adverbs such as *happily*. Others believe that semantic bleaching and syntactic reanalysis can co-occur: “‘bleaching’ is the result of reanalysis or, perhaps better said, it is the essence of the reanalysis itself” (Harris and Campbell 1995: 92). For Fischer, grammaticalization is usually seen as “a type of semantic change ... which is intertwined with or followed by grammatical restructuring” (1997: 173), but she argues, on the basis of the development of English *have to* constructions, that the semantic and syntactic changes in grammaticalization are quite independent of one another. But most grammaticalization scholars would claim that the evidence points to reanalysis following semantic change: “conceptual/semantic shift precedes morphosyntactic and phonological shift ... the result is asymmetry between meaning and form” (Heine et al. 1991:213, 260; see also Bybee et al. 1994). That is to say, gradual meaning shift runs on ahead of syntactic structure, which periodically undergoes sudden reanalyses to ‘catch up’. Data on the development of *instead of* suggest that the semantic change from the locative to the substitution meaning took place before the structural change which allowed the expression to take

clausal complements (Schwenter and Traugott 1995 and Traugott forthcoming). Traugott (1999: 178) shows that semantic change must have occurred in *in fact* before it was able to move to a new syntactic position.

Fig. 2.5 shows a possible cycle within the discourse-based grammaticalization purview.

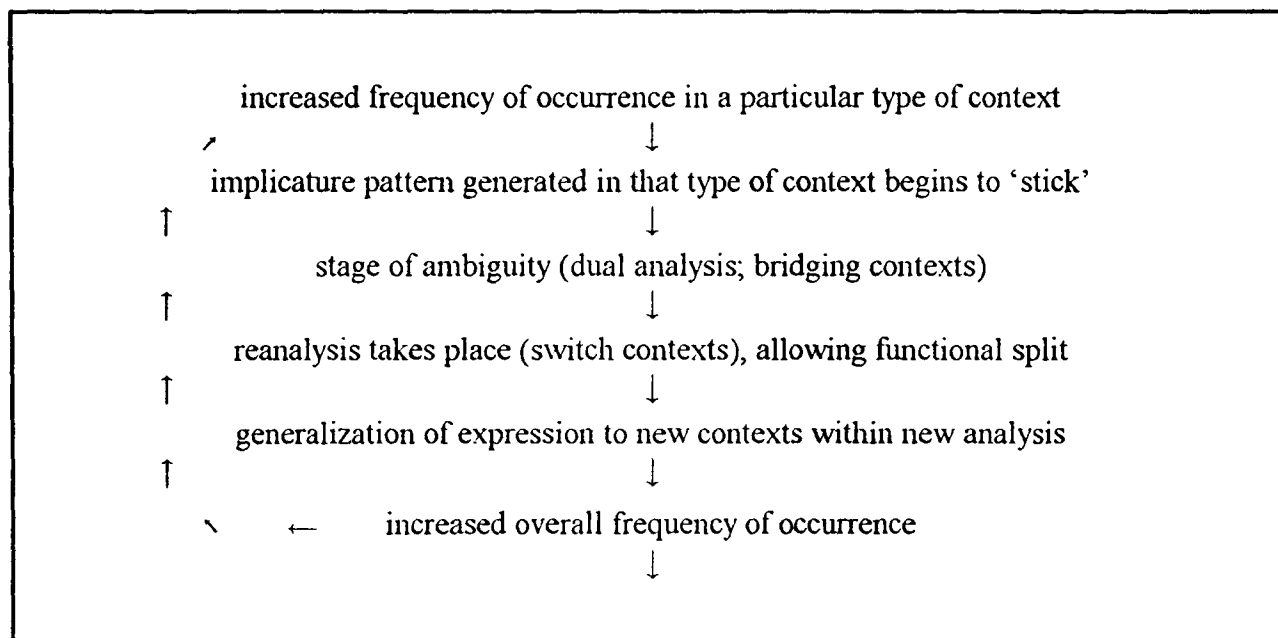


Fig. 2.5. A possible discourse grammaticalization cycle

There need be no starting point, nor need the cycle be completed. As Heine and Reh point out, “grammaticalization is an evolutionary continuum” (1984: 15). Sociolinguistic and variational catalysts may affect a construction after a long period of stability. For example, an expression or variant may acquire local social prestige via its association with particular contexts or speakers and then extend to new contexts and speakers. (Conversely, a variant may lose prestige and fall out of use.) This model does not imply a causal link between the morphosyntactic and semantic changes, but it does suggest unidirectionality. Change tends to start slowly, in a limited range of contexts, then to take off and spread rapidly until either it loses prestige or some level of saturation is reached (i.e. it has spread to one or more sets of relevant contexts), at which point it stabilizes. This is akin to the ‘S-curve’ by which sound changes in language are usually described (Aitchison in press).

An interesting view of the relation between semantic change and syntactic reanalysis is that the latter can often result from type coercion. Type coercion “captures the semantic relatedness between syntactically distinct expressions” (Pustejovsky 1995: 58). It is defined as “a semantic operation that converts an argument to the type which is expected by a function, where it would otherwise result in a

type error” (Pustejovsky 1995: 59, 111). That is, a lexical item or phrase is “coerced to a semantic interpretation by a governing item in the phrase *without change of its syntactic type*” (Pustejovsky 1995: 61, emphasis added)³⁴. Gronemeyer suggests that certain changes observable in grammaticalization can best be explained by type coercion. Her hypothesis is that “synchronic semantic processes of enriched composition can ... feed diachronic reanalysis” (1999: 16-17), the enriched composition being facilitated by type coercion. She argues that the mechanism behind the development of inchoative *get* as in (5) is type coercion: the adverbial and adjectival complements are initially coerced into goal readings. Later, *get* is reanalysed to take complements denoting change-of-state rather than change-of-location.

(5)	construction	<i>get</i> + PP/particle	>	<i>get</i> + Adverb	>	<i>get</i> + AP/adjectival PP
	sense	spatial		spatial		spatial or mental
	example	<i>get to, from</i>		<i>get back, ashore</i>		<i>get alone, dressed</i>

(adapted from Gronemeyer 1999: 28)

Gronemeyer claims that “the category of the complement changes, while the semantics remain constant, and then the syntax of the new construction generalizes allowing for a novel interpretation. Reanalysis occurs when type-coercion no longer takes place” (1999: 36). At first sight, this analysis appears merely to blur reanalysis: it is not immediately clear how to distinguish between a reanalysis and a type coercion³⁵. But it may be that a coercion model can usefully be applied to some developments affecting English modal adverbials. For instance, in (6), the VP complement can be seen

(6)	construction	<i>at least</i> + NP	>	<i>at least</i> + VP	>	<i>at least</i> + S
	sense	scalar		scalar/evaluative		evaluative
	examples	<i>at least nine</i>		<i>who would not at the least have chosen them a time (1563)</i>		<i>at least it's sunny today</i>

as coerced to a scalar entity by *at least*, and able to extend to the novel evaluative sense once the

³⁴ Pustejovsky’s aim is to provide a lexical semantics that, inter alia, accounts for words taking on novel meanings in novel contexts and that caters for polysemy instead of resorting to multiple meanings. Type coercion is one semantic operation that allows for a particular interpretation in a given context. An expression can “coerce an argument to the appropriate type only if that word or phrase has available to it an interpretation of the expected type” (1995: 59).

³⁵ In Pustejovsky’s model, constraints on type coercion are built into the lexical entry of the element that is coerced, which must have available to it “an *alias* of the appropriate type” (1995: 116). Pustejovsky’s example is *begin*, which requires an event complement. In ‘John began a book’, *book* can be coerced to an event reading by *begin* because its agentive (‘write’) and telic (‘read’) qualia roles make the event type available to it as an alternative type.

reanalysis to allow it to take a VP complement has taken place.

To return to the order of changes, Gronemeyer does not explain what might trigger type coercion, e.g. what might cause *get* suddenly to start coercing adverbial complements. It is hard to see how the appearance of a new syntactic category of complement for *get* might be motivated, unless semantically or pragmatically. It is more plausible, though purely speculative, that increased frequency of occurrence with some of its normal complements could strengthen certain implicatures and provoke a spread to semantically similar, syntactically distinct complements. In any case, a type coercion analysis of data such as (5) does not show conclusively that syntactic change comes first. In fact, type coercion can be regarded as another way of looking at meaning/form asymmetry.

As mentioned in chapter 1.4, the types of notion expressed lexically by discourse markers in English are often grammaticalized in other languages: there is a clear universal tendency towards grammaticalization of modality and other speaker comment. It will be argued that the development of some English modal adverbs can profitably be seen as grammaticalization.

2.4.6 Grammaticalization, discourse connectives and polysemy

It has long been observed that interpretation in context can give rise to semantic shifts. ‘Grammaticalization chains’ (Heine et al 1991: 220ff; Heine 1992) develop via such shifts. They create form-meaning asymmetries which result in polysemy. As the semantic changes and extension associated with grammaticalization take place, there emerge what have also been called ‘meaning chains’ (Taylor 1995: 108) or ‘sense-spectra’ (Cruse 1986: 71), wherein the various readings of a single form seem to fall along a continuum rather than be discrete senses. As an expression extends into new contexts, so its meaning becomes ‘contextually modulated’ (ch 1.4.6) and “the process of contextual modulation contains the seeds of polysemy” (Taylor 1995: 124). Eventually, the extensions of the form may crystallize into new polysemies that are non-arbitrary (they are transparently related) but nonetheless unpredictable (knowledge of one sense does not allow the speaker to predict the other(s)). This crystallization of new meanings is often referred to as ‘functional split’ (Heine and Reh

1984: 57-8). Finally, the senses may diverge into what Jackendoff terms ‘opaquely chained concepts’ (1996: 112): senses that are only seen to be related by knowledge of their historical development; i.e. the motivation for the multiple senses is no longer transparent. Once the relation is opaque to the speaker, of course, the senses become homonyms. The older meaning co-exists alongside the new, often for several centuries, and may continue to influence the selectional restrictions on the new use. From the point of view of mental representation in the speaker, then, it is suggested that a form can have a single sense, or a range of transparently related, contextually differentiated senses built around a single concept, or distinct unrelated senses.

This presupposes that, as discussed above, semantic change is gradual. Those who reject this notion, preferring to view semantic change as discrete reanalysis, also tend to reject the notion of polysemy in favour of homonymy: “a fundamental flaw of most categorizations of semantic change is that they rest upon the assumption that the old meaning *becomes* the new meaning, that there is some real connection between the two ... a connection between the new and old meanings is illusory” (Fortson in press).

However, meaning extension is not random: polysemy structures show regular interrelations of meanings within and across languages. To account for these, it is necessary to analyse both the synchronic and the diachronic situation. As Ullmann notes, “certain phenomena, such as for example conflicts between homonyms, can be handled only by a judicious combination of descriptive and historical viewpoints” (1964: 31). Tracing such patterns can be revealing of the influence of cognitive structure on language. For Hopper and Traugott, “from the perspective of grammaticalization it is methodologically essential to assume polysemy if there is a plausible semantic relationship ... because otherwise relationships between more and less grammaticalized variants of the same form cannot be established.” (1993: 71). It is the visibility in synchronic polysemy structures of the operations of language change that has led to the view of grammaticalization as ‘panchronic’ (Heine et al 1991: chapter 9, Nichols and Timberlake 1991: 130, see also Sweetser 1988: 9).

The pattern of development of modal S-Advts out of VP-Advts in English seems to follow the

model outlined above. In most cases, a polysemy structure has developed for the adverb, arguably via metonymic extension. In some cases, such as *surely* or *indeed*, the relation between fossilized VP-Adv usages and connective S-Adv usages may already have become opaque for most speakers. In other cases, such as *at least* or *in fact*, the relation is probably transparent (though the newer senses are not thereby predictable from the old).³⁶

What is layering from the diachronic perspective is polysemy from the synchronic perspective (fig. 2.6).

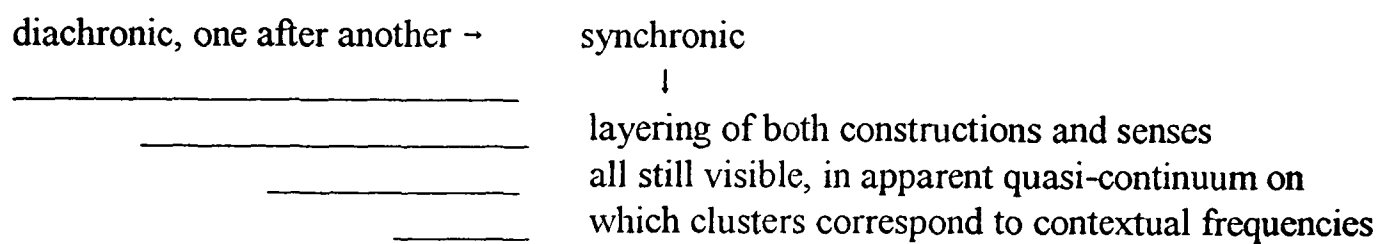


Fig. 2.6. Layering and polysemy

Not only are diachronic processes visible in synchronic language, but synchronic (sociolinguistic) variation is relevant too. Sociolects and dialects interact, as speakers tend to participate in more than one variety of language; innovations in one variety can spread to other varieties. While written language is undoubtedly more conservative than spoken, in highly literate and specialized communities there is much scope for innovation to arise in primarily written sublanguages and spread or not to other varieties.

2.5 Conclusion

What emerges from the above discussion is that there is no clear boundary between grammaticalization and internal language change in general. Indeed, the introduction to a collection of papers on the bounds of grammaticalization suggests that the only language changes that are not grammaticalization are some phonological changes, lexical replacements and noun-verb conversions (Ramat and Hopper 1998: 4). This state of affairs has engendered a more fundamental debate on the

³⁶ Polysemy is traditionally said to exist when a single form has two or more related meanings in the same syntactic category. For cases of 'polysemy' involving different categories, Lichtenberk (1991) introduced the term 'heterosemy'. However, some grammaticalization scholars extend the term 'polysemy' to include these cases (e.g. Traugott 1986). Moreover, if syntactic category boundaries are fuzzy (v. ch. 1), the criterion must be applied flexibly. Haspelmath (2000) suggests that fuzziness of word classes is particularly evident in expressions that have recently grammaticalized.

nature of ‘grammaticalization’ and on its place in language change studies. There is controversy over the status of the posited ‘unidirectionality’ of grammaticalizing change (e.g. Lass in press, Heine in press), over whether it constitutes an approach or a type of change (Newmeyer 1998, Traugott in press), and over whether, as a framework, it is explanatory or merely descriptive (Kroch 1989, Roberts 1993, Fischer 1997). Differing responses to these questions have produced two main characterizations of grammaticalization: one strong and one weak. The strong characterization holds that it constitutes a unified process having its own motivation, rather than a collection of changes that are independently motivated. It remains to be shown whether the three main aspects of grammaticalization always interlock in a unified process or whether their co-occurrence is contingent. It would need to be demonstrated that the set of observed changes was bound by a single motivation. The weak characterization sees grammaticalization as “an approach, a way of construing the data” (Traugott in press); or as a cover term for “a conjunction of familiar developments from different spheres of language, none of which require or entail any of the others” (Newmeyer 1998:295).

The term will be understood here in its weaker sense to refer to the tendency for semantic change, phonetic or prosodic change, and syntactic reanalysis to co-occur in the direction of extending a particular linguistic element or sequence to new (more subjective and/or more abstract) usages. None of the parameters of change is claimed to be unique to grammaticalization.

Adverbial expression of speaker commitment is often claimed to be a ‘late’ development (Hanson 1987, Swan 1988) or, rather, a development favoured by high levels of literacy. Ramat and Ricca claim, for instance, that in the Middle Ages, the re-emergence of a strong rhetorical and literate civilization in Europe meant that “Romance languages again had a need for discourse-oriented communicative means” and had therefore to “reinvent the whole strategy for adverbial formation which got lost in the crisis of the classical tradition” (1998: 244). These claims, however, may be influenced by a politically-motivated view of what ‘a language’ is (in terms of bundling registers and sociolects), and by a tendency to categorize expressions typical of spoken language ‘discourse markers’ differently from those typical of written language ‘sentence adverbials’ (v. ch 1.1). They do not take sufficient

account of the interrelationships between spoken and written language and among different registers. Brinton has shown for Old and Middle English that what looks like “a lack of literary sophistication and a ‘primitive’ paratactic style” (1996: 1) does contain plenty of discourse orientation in the form of “oral residue” (1996: 8). Stein (1985: 300) suggests that in fact expressions of attitude and commitment are relatively unstable and fast-changing. It is at least possible that rather than a ‘late’ development, modal adverbials are ‘early’ and the notions they express can in time be expected to grammaticalize into more stable formatives. This hypothesis is consistent with the fact that English both has an unusually heavy reliance on adverbials to express attitude and modality and is a relatively young language.

This chapter has presented a framework for analysing DCs in terms of their expression of the rhetorical relations that help to create discourse coherence. Second, it has claimed that such rhetorical relations can only be understood in the wider context of the information structuring of a discourse and ultimately of the overall purpose of the discourse. Finally, it hypothesized that both the polysemy and the multifunctionality of discourse marking expressions can usefully be viewed as resulting from a grammaticalization-like course of development.

CHAPTER 3

OUTLINE OF METHOD AND DATA

3.1 Introduction	99
3.2 Corpus linguistic method	99
3.2.1 The approach	99
3.2.2 Corpus linguistics and lexical semantics	101
3.2.3 Use of statistical method in corpus linguistics	104
3.2.4 Corpus design and sampling issues	105
3.2.5 Analysis of the data	105
3.3 The data used in this study	105
3.3.1 Selection of data	106
3.3.2 The synchronic data	108
3.3.3 The diachronic data: full-text sources	109
3.3.4 The diachronic data: non-full-text sources	110
3.4 Summary	111

3.1 Introduction

Corpus linguistics can be described as a new method for the exploration of traditional descriptive linguistic concerns. ‘Corpus’ is understood as ‘body of machine-readable text’. This chapter looks at the advantages and problems of using corpus data and at corpus design and sampling issues. The chapter also describes the research data, which is drawn from the British National Corpus and from electronic versions of some British newspapers.

The next section, 3.2, gives a brief description of corpus linguistics and its relevance to lexical semantics and lexicography. The data used in this study are described in 3.3.

3.2 Corpus linguistic method

3.2.1 The approach

Corpus linguistics has been described as "a new philosophical approach" to linguistics (Leech 1992: 106). Language variation and language change are areas that can benefit particularly from large computer-readable corpora and from software tools with which to analyse them. But corpus linguistics is also useful in the development of natural language processing (NLP) software.

As a data-driven, empirical approach, corpus linguistics resembles conversation analysis, first

in that it seeks to discover regularities, and second in that it makes fewer theoretical assumptions than other methodologies. In conversation analysis "search is made for recurring patterns across many records of naturally occurring conversations ... as many instances as possible of some particular phenomena are examined across texts" with the aim of discovering "the systematic properties of the sequential organization of talk" (Levinson 1983: 287). Conversation analysis focuses, of course, on spoken dialogue.

Both corpus linguistics and conversation analysis (CA) challenge conventional syntactic and semantic classifications that assume mutually exclusive functions. They tend to ignore language structure, at least initially, in favour of a naive approach to the data: "CA is reluctant to assign either meaning or function to a form apart from a thorough investigation of how that form is sequentially embedded in actual talk" (Schiffrin 1994: 279). Schiffrin (1994) suggests that there are advantages to analysing an expression in two different ways: first to examine all occurrences, without taking account of its conventional syntactic and semantic categories, and then with categorisation. This method enables the analyst to discover what, if any, are the underlying common functions across categories. Her example is '*there* + BE + ITEM': "That one construction can fit the resolution of simultaneous tasks also suggests that these tasks themselves might be related to one another" (Schiffrin 1994: 272).

Typical of the corpus linguistic approach is Johansson and Oksefjell's analysis of *get*, which takes a similar approach to Schiffrin's. Using corpus data together with established grammatical categorization schemes (such as lemmatisation and syntactic structures), they seek to build a unified account of *get*, and claim that "the main uses can be connected to one general prototype" (Johansson and Oksefjell 1996: 74). The corpus linguistic method is primarily inductive. However, although a few corpus linguists adopt a radically atheoretical stance, most adopt this dual approach and use corpora as test-beds for theories.

Above all, corpora provide data on frequency and distribution. As Croft points out, corpus data provide evidence for the conventionality of uses in terms of frequency of occurrence: "form-meaning pairings that have a significant token frequency are highly likely to be entrenched uses in

speakers' minds" (1998: 169). Conventionalized uses point to polysemous or homonymous senses. Corpus data can therefore reveal nascent polysemies by showing which uses are becoming entrenched. This is what Halliday calls "the transformation of instance into system" (1991: 34), a process which large corpora are beginning to allow us to observe.

Bybee (in press) distinguishes type frequency from token frequency. While token frequency is the frequency of occurrence of the target expression, type frequency refers to the range of contexts in which the expression can occur. Bybee's example is the increase in type frequency of *be going to*: first, the range of possible NP subjects has increased from animate, mobile ones to almost any NP subject; second, the range of following verbs has increased from dynamic ones to almost any type of verb. She suggests that increase in type frequency often leads to increase in token frequency, and that it is token frequency that ultimately triggers the changes associated with grammaticalization. A major aim of the analyses of the modal adverbs described in chapters 5-8 is to investigate type frequencies for those expressions, to test for a relation between quantitative shifts in types (context patterns) and semantic shifts.

What is of interest is relative frequency: "possibly the most serious limitation at present on a semantic interpretation of textual frequency data is the lack of normalized comparative statistics from a range of text types" (Stubbs & Gerbig 1993: 77). Data on a few modal adverbs in PDE (chapter 4) show how much variation there is among texts, and point to the dangers of extrapolating frequency data. Perhaps the main challenge of corpus linguistics is reliability. Nonetheless, corpora have already had a huge influence on descriptive linguistics.

3.2.2 *Corpus linguistics and lexical semantics*

The sense of a lexical item is traditionally defined by reference to its relations with other lexical items, its syntactic behaviour, and its relationship to conceptual models. Different approaches to lexical semantics have placed greater or lesser emphasis on each of these aspects of lexical meaning.

One approach focuses on lexical relations. Much lexical semantics has thought of lexical

meanings largely in terms of their paradigmatic and syntagmatic relationships with other lexical meanings; that is, to study the relations of synonymy, antonymy, hyponymy and so on, among lexemes. This is the ‘semantic field’ approach, exemplified in, for example, Cruse (1986), who suggests that “an extremely useful model of the meaning of a word, which can be extracted from the contextual relations, is one in which it is viewed as being made up, at least in part, of the meanings of other words” (1986:16). Particular attention is paid to verbs, nouns, and adjectives. This approach tends to take a fine-grained approach to discreteness of sense and produces what Pustejovsky (1995:29) calls a ‘sense enumeration lexicon’.

For example, Cruse (1986) delimits lexical units using paradigmatic and syntactic criteria such that wherever a meaning can participate in meaning-relations, and wherever grammatical difference correlates with meaning difference, there is a discrete sense (1986:ch.3). Moreover, he distinguishes very clearly between ‘lexical unit’, which he defines as “the union of a lexical form and a single sense” (1986:77) and a ‘lexeme’, which is a “family of lexical units” (1986:76). It is the lexical unit that is “the primary operational semantic unit”, while the lexeme is secondary (1986:80)¹. Cruse thus eschews the notion of polysemy on the grounds that it suggests a view of lexical meaning that privileges the lexeme over the lexical unit (1986:80). The lexical relations approach, then, tends to favour maximization of homonymy at the expense of polysemy. This not only leads to redundancy in the lexicon, but allows for distinctions of sense to be multiplied indefinitely (cf. Lyons 1977: 554-560). This proliferation of senses is in fact a danger of the corpus linguistic approach: the tendency to see a distinct sense in every context type an expression occurs in.

But it is rarely possible to draw rigid boundaries between senses of forms. First, as will be seen in the next chapter for *clearly* and *surely*, words gradually acquire new contexts and new senses.

¹ Lexical units are allocated to a lexeme (a) when, given sense ‘x’ of lexical unit X, it is possible to predict the existence of lexical unit Y with sense ‘y’ or (b) when the senses of two or more lexical units all belong to the same ‘sense-spectrum’ (meaning continuum) (1986:77-9). Cruse’s examples include *jacket*, which, by analogy with *skirt*, *coat*, etc., can be predicted to have a type reading if it has a unit reading, and *mouth*, which has a sense-spectrum running from the mouth of an animal to the mouth of a river. Thus, *jacket*-unit and *jacket*-type are two lexical units which can be allocated to a single lexeme on the basis of prediction of one sense from the other (the existence of the prediction rule is inferred from the recurrent semantic contrasts between other pairs of similar lexical units). And *mouth*-of-animal and *mouth*-of-river are distinct lexical units which can be allocated to a single lexeme on the basis that they both belong to the same sense-spectrum.

Second, the different senses of a word overlap and/or contain one another. This is particularly true of modal adverbials, as will be seen further in chapters 5-8. Third, there is no necessary correspondence between the sense of a word and a particular syntactic construction it enters into. It is “arbitrary to create separate word senses for a lexical item just because it can participate in distinct lexical realizations” (Pustejovsky 1995:50). Moreover, such an approach assumes a strict separation of diachronic and synchronic study that is artificial and unhelpful (ch. 2). For these reasons, what Pustejovsky calls ‘sense enumerative lexicons’ (1995:39ff) are not the best way to approach lexical semantic meaning.

A second approach emphasizes the way an item fits into syntactic patterns (e.g. Levin 1993), grouping lexical items according to which patterns and sub-patterns they can participate in, i.e. by their valency. Much work here has been done on verbs and their argument structures. The application of this approach to sentence adverbs was seen in chapter 1.

A third approach focuses more on the relationships between lexical forms and the conceptual models they reflect and groups lexical items by the models they share or interact with. “The semantics of natural language”, as Pustejovsky puts it, “should be the image of nonlinguistic conceptual organizing principles”. Word meanings, therefore, “should somehow reflect the deeper conceptual structures in the cognitive system” (1995:6). One such approach is that of frame semantics (Fillmore and Atkins 1992, 1994), which focuses not on syntactic distribution or lexical relations but on conceptual frames. Fillmore and Atkins (1994) provides an analysis of *risk* along these lines. The sense of *risk* is described as the relations it signals among the elements (such as Protagonist, Decision, Goal, etc.) of a schema, or conceptual frame. This approach allows both for the contexts of use to be incorporated into the sense of the expression and for a unified account of the various uses of the verb. For this type of approach, corpus data is invaluable.

We have seen that in the grammaticalization process the frequency and distribution of a linguistic expression help to forge its meaning. Using corpus data, “a discourse-oriented statistical analysis can suggest how to recognize possible ongoing grammaticalization processes” (Hopper &

Traugott, 1993: 201). This is what we shall try to do in the cases of *after all*, *in fact*, *at least* and *of course*. There is a sense too in which the frequency and distribution of an expression is a part of its synchronic meaning. Pustejovsky et al. (1993) report a number of distinct patterns of use emerging from corpus data on the verb *insist*, and suggest that "if such knowledge is in fact so systematic, then it must be at least partially represented in the lexical semantics of the verb" (1993: 350). Corpus data can provide the basis for describing those patterns and uncovering their co-occurrences with other aspects of discourse.

3.2.3 Use of statistical method in corpus linguistics

The statistical quality of the data can be estimated, by calculating the confidence interval, or error bars, for each derived result, assuming an infinite population:

$$\text{Confidence interval} = p \pm z \sqrt{p(1-p)/n}$$

where

- n = sample size
- p = sample proportion (no. of occurrences/sample size)
- z = a factor chosen according to the required confidence level. For a confidence level of 99%, z takes the value 2.58.

As an example, we can calculate the confidence interval of the finding that 75 out of 412 tokens of *after all* in the Periodicals data are temporal. If our data were known accurately to represent the target population, then 18.2% of tokens of *after all* would be temporal, since 75 is 18.2% of 412.

For

$$\begin{aligned} n &= 412 \\ p &= 75/412 = 0.182 \\ z &= 2.58 \text{ (for 99\% confidence)} \end{aligned}$$

the confidence interval

$$\begin{aligned} &= 0.182 \pm 2.58 \sqrt{0.182(1 - 0.182) / 412} \\ &= 0.182 \pm 0.049 \text{ or } (18.2 \pm 5)\% \end{aligned}$$

So we can be 99% sure that the proportion of temporal *after all* in the wider population (in this case, the language of British newspapers, magazines and journals) lies between 13.2% and 23.2%. The confidence interval scales primarily as the inverse square root of the sample size, and is

proportionately greater for low frequencies within the sample.

3.2.4 Corpus design and sampling issues

The data sets chosen for this study reflect the need to look at both diachronic change and synchronic variation to understand polysemy. PDE, like any so-called language, is a continuum of varieties, some more advanced, some more conservative. Hence attempts to develop ‘diversified’ corpora such as the BNC.

Conversation, itself heterogeneous, is usually taken by linguists to be the prime or most natural manifestation of language and therefore the principal object of study. It is also likely to be the most advanced and the most context-bound. Written-to-be-read texts on non-specialist topics and aimed at a widely dispersed audience are likely to be the most conservative and the most context-independent.

In practice, few corpora are truly register-diversified. In the first place, it is not clear what this means. Even supposing the issues of register definition and dialect variation were solved, should each identifiable register within a dialect be represented? Or should registers be represented to the extent that they contribute to the total output of the dialect? Or should the selection match the composition of typical language inputs? Secondly, spoken language is vastly more expensive to include in a corpus than written language, so that the latter predominates: it comprises most of very large corpora (90% of the BNC for instance). There are more written corpora available, and spoken corpora tend to be small. Corpus design is still severely curtailed by these practical constraints.

3.2.5 Analysis of the data

The software used was WordSmith Tools, written by Mike Scott at the University of Liverpool. Tokens of the target adverbs were extracted in 1,200-character spans, providing around 10-14 lines of co-text. The adverbs were then analysed, using parameters, depending on the adverb and textual source, such as date, source, sense, position (relative to host), host-category, verb, status

(factual or hypothetical), rhetorical relation, collocations, comments. A database of tokens was created (using Microsoft Access), each token being one record and each parameter value an attribute. This made it possible to search on any parameter as well as to run counts.

Where a token was unclear, as many in the Conversation data were, they are allowed for in the lexical density calculations but thereafter excluded.

3.3 The data used in this study

3.3.1 *Selection of data*

One of the aims of this study is to bring together synchronic and diachronic analysis. These are relevant to one another, for two main reasons. First, as discussed in ch. 2.4.2, diachronically emergent, ‘longitudinal’ polysemy structures are visible in synchronic ‘section’ (thus grammaticalization is often perceived as a ‘panchronic’ phenomenon). Second, written language is more conservative than spoken language. We can therefore expect to find similar differences between contemporary written and spoken language to those we find between language samples from different periods. The choice of data, in response to the above aim, includes (a) the three categories spontaneous spoken data, written-to-be-spoken data and written data for PDE and (b) historical data.

Another aim is to explore the use of modal adverbs of certainty in argumentation strategies. A second criterion of text choice is therefore to prefer texts likely to have significant evaluative and persuasive functions; hence the choice of political speeches and reviews. Literary language (fiction) is omitted as far as possible, due to its special characteristics.

The quantity of data needed depends on (a) the target population and (b) the linguistic features under investigation, i.e. how frequent they are in the target population.

On the whole, cross-register corpora are not as useful as sublanguage or mono-register corpora, because of the sampling issue: the wider the register range, the less representative the sample. Even what is nowadays considered a large sample (say, 100 million words) is too tiny to be representative of ‘English’. For sub-languages or particular registers, relatively reliable data can be

obtained from smaller corpora. However, replicability is still a problem because few studies have so far been replicated on a second or third corpus of the same target population.

Clearly, the greater the density of the target feature in the target population, the smaller the amount of data needed; the rarer the feature, the larger the amount of data needed. In other words, the amount of data needs to be calculated in terms of instances of whichever category provides the entry condition for the feature under study. In the case of discourse connectives, the relevant category is discourse segments; however, the feature is optional, so that in practice an amount of data is needed sufficient to yield enough occurrences. There is no method for establishing how many occurrences are 'enough'. Obviously "more than one instance of use of an item is needed as a basis for description" (Kennedy 1998: 67), and five is the minimum for which any statistical validity will be achieved (Manning & Schütze 1999: 161).

Overall, then, the more homogeneous (while preserving representativeness) the texts in the corpus are, the less data is needed; the more frequent the target features are in the target population, the less data needed. It is therefore not possible to propose that 20,000 word or 2,000 word samples are adequate without clearly specifying the population and the target features. On the assumption that the corpus is representative of the target population, and assuming an infinite population, 200 instances of a feature give a margin of error of around $\pm 5\%$. For example, if the frequency of occurrence of a feature is 1/10000 (one in ten thousand words), a 2-million word sample will provide results to within 5%. In practice things are not so simple. First, the nature of the target population is rarely so well described that it can be ascertained whether or not the data collected are representative. Second, scholars have to work with whatever corpora exist already or can easily be assembled, so the ideal data may not be available. Finally, research may be constrained by the software available. For example, the 67 linguistic features used in Biber (1988) were chosen as much for their ease of automated recognition as for their role in defining text type.

Most of the data used as input to this study is from corpora. But, particularly for the diachronic analysis, data has also been taken from the quotations in dictionaries and other non-full-text

sources. The statistics on relative frequencies and distributions are based on full text sources only.

However, many examples are taken from non-full-text sources.

3.3.2 *The synchronic data*

Three main corpora are used. They will be referred to as Conversation (Conv), Speeches (Sp) and Reviews (Rev). They are supplemented with data from a Periodicals corpus (Per).

The Conversation corpus consists of the 4.2 million words of demographically-sampled spoken dialogue from the British National Corpus (BNC). The Speeches corpus is 1.4 million words of political speeches given during the period 1996-1999. The majority are by ministers of the British Conservative and Labour governments; some are by Irish government ministers and some by other British politicians. There are 528 speeches in the database. The Reviews corpus contains 1018 book reviews totalling 1.1 million words. The reviews are of both fiction and non-fiction and they are drawn from the British and Irish press: newspapers and weekly journals.

The raw figures for the final data from these sources that was analysed, after discarding repeats, quotations and unclear instances, are given in table 3.1.

	Reviews	Speeches	Conversation
<i>after all</i>	91	63	70
<i>in fact</i>	223	121	333
<i>at least</i>	252	172	467
<i>of course</i>	254	680	502

Table 3.1. PDE data drawn from three corpora: numbers of tokens

Additional data is taken from the Periodicals section of the BNC, which totals 27.9 million words of text from periodicals (newspapers, magazines and journals).

There are important differences between the spoken dialogue data and the written-to-be-spoken and written data. In particular, there are several problems associated with using the demographically-sampled BNC data. First, the transcription quality is extremely variable, sometimes very poor, since it

was carried out by non-linguists and contains many errors. Where a recording has, by mistake, been transcribed twice there are wide divergences in the two transcriptions. The recordings are not available, so that there is no way to check the validity of the transcriptions. There is no prosodic mark-up, though long pauses, coughs and simultaneous speech are coded. Punctuation has been imposed by the transcribers. The transcriptions make it difficult to identify intonation units, and therefore discourse units. In the examples cited in chapters 4-8, the written-style punctuation has been removed, which makes the texts easier to read. Commas and full stops have been replaced with dots and capital letter retained only for proper names. The mark-up symbols used are listed on page viii.

The advantages of using the demographically-sampled BNC material are that the database is large, and the sampling is broad.

3.3.3 The diachronic data: full-text sources

The full-text sources of diachronic data are the Helsinki Corpus (Kytö 1993), the Lampeter Corpus, the Corpus of Early English Correspondence Sampler (Ceecs), and Historical Texts (Hist) — samples of written language (mostly discursive non-fiction) taken from the end of the fourteenth century to 1900 and averaging 900,000 words per century.

The Helsinki Corpus is a register-diversified corpus of texts from Old, Middle and Early Modern English, from the period 850-1710. The Middle English (608,570 words) and Early Modern English (551,000 words) sections have been used.

The Lampeter corpus consists of 120 texts totalling 1.1 million words of tracts and pamphlets published between 1640 and 1740. They are drawn from the fields of religion, politics, economics and trade, science, law and a few are miscellaneous. Only complete texts are included in this corpus.

The Early English Correspondence Sampler holds 450,000 words of letters written between 1418 and 1680, most of it in the latter half of this period. 194 letter writers are represented.

The Historical Texts have been culled from the Oxford Text Archive and from a variety of on-line text archives. They are predominantly non-fiction texts on economic, medical, philosophical,

political, scientific, social and theological topics. Some personal letters are also included. The list of texts is given in Appendix B.

The full-text sources have been used to calculate the frequencies of the adverbs in the study.

Table 3.2 gives a summary of the full-text sources.

	Historical texts	Helsinki ME	Helsinki EME	Lampeter	Ceecs	Total
1300-1399		192,470				192,470
1400-1449		158,310			38,323	196,633
1450-1499		118,100			55,110	173,210
1500-1549	113,526	119,390			36,147	269,063
1550-1599	660,165		148,090		103,762	912,017
1600-1649	193,705		123,950	129,431	136,132	583,218
1650-1699	510,459		159,410	618,081	80,594	1,386,544
1700-1749	226,686		11,630	445,941		684,257
1750-1799	579,199					579,199
1800-1849	341,114					341,114
1850-1899	860,960					860,960

Table 3.2. Sizes in words of full-text sources of historical data

3.3.4 The diachronic data: non-full-text sources

Historical examples have also been drawn from other corpora, citations databases and dictionaries. Examples are taken from the Corpus of Middle English Prose and Verse (CMEPV) and from the Michigan Early Modern English Materials (MEMEM). The latter consists of the citations collected for modal verbs and some other words for the *Early Modern English Dictionary*. These include many of the citation slips used by the OED editors. The database consists of around 50,000 records. The Newdigate corpus is the source of a few examples². The Newdigate Newletters is a collection of 3950 newsletters written between 1673/4 and 1715, most of them addressed to Sir

² The Newdigate corpus (1m words) was not used as a full-text source in order not to unbalance the data by introducing a very different text type that was not comparable with those from other periods.

Richard Newdigate, hence the name. The corpus contains 2100 of them, written between 1674 and 1692.

Examples are also taken from dictionary quotations and definitions. Modern dictionaries used are the Oxford English Dictionary, 2nd edition (OED) and the Middle English Dictionary (MED). Algeo notes that “the evidence of the OED has to be used cautiously ... The OED is biased in favour of literature and particularly of canonically enshrined authors” (1998: 63), but no glaring discrepancies were found between the distribution of the adverbs in the full-text sources and in the OED quotations. Historical dictionaries include Johnson’s dictionary (1755 and 1773 editions), and some of the dictionaries in the Early Modern English Dictionaries database (EMED).

Examples are followed by a code in brackets indicating the source. A list of these codes is in Appendix B, where further information about the data sources can be found.

3.4 Summary

Corpus linguistics allows language in use to be explored as it could not be before the advent of large computerised databases. Nowadays lexicography without machine-readable corpora is inconceivable. The size of the text samples described above in most cases provides tens or hundreds of tokens of the target adverbs. The margins of error are typically a few per cent. The figures therefore allow us to draw conclusions about their frequency and distribution that are statistically meaningful.

CHAPTER 4

MODAL ADVERBS AND POLYSEMY

4.1 Modal sentence adverbials of certainty	112
4.2 <i>Clearly</i>	115
4.2.1 Historical development	115
4.2.2 Contexts of use	117
4.3 <i>Surely</i>	118
4.3.1 Historical development	118
4.3.2 Contexts of use	120
4.4 Conclusion	121

4.1 Modal sentence adverbials of certainty

The formation of speaker-oriented sentence adverbs is fairly productive in English.. An obvious recent case is *hopefully*, along the lines of *thankfully*. The following example is along the lines of *surprisingly*.

- (1) Traditional Oxford handwritten notes now co-exist with answerphones and email, which *frustratingly* I do not yet have in Magdalen (1998, Don's Diary, THES)

As in many other languages, adverbs of certainty tend to derive from VP-adverbs and are often later recruited as discourse connectives. The development VP-adverb > sentence adverb > speaker-oriented sentence adverb is a regular one. Examples (2) illustrate this development for the English adverbs *clearly*, *surely*, *certainly*, *actually*, *at least*, *of course* and *indeed*.

- (2) a. .. vnto the tyme he knewe the mater *clerly* (1481)
They *clearly* have grave problems. (PDE)
- b. .. if he sitte nat *sueryly* .., the horse casteth him quickly (1531)
Surely you've heard of him? (PDE)
- c. .. if yow understand *certainely* the worth, the value, the tenure, the reparations and the price, I will tell you more (1598)
I don't know about warmer, but it's *certainly* going to be overcast. (PDE)
- d. Those who offend *actually*, are most grievously punished (1660, OED, Stanley)
Actually, of course, Britain is, in a sense, wholly an immigrant society (PDE)
- e. Lat se wher thou kans tellen aught in geeste, Or telle in prose somewhat, *at the leeste*, (c1395, Chaucer, Prologue to Sir Thopas)
Sir Alec is *at least* no fair weather convert (PDE)

- f. A pardon ... which either is graunted *of course*, or ye kyng of pytee and compassion geveth. (OED, 1548)
then we must face the problem and change the policy. *Of course* this is difficult (PDE)
- g. He doop synne bope *actuely in deede* and wilfully, or myntaly, in þe soule (c1425, MED)
They have *indeed* a great Army together (1674, London Gazette 937, 9-12 Nov)

Swan has documented the increase in the number of sentence adverbs (types, not tokens) in English from Old English to Present-Day English (1988: 514-5). She draws attention to the rise, during the Middle English period, of the truth intensifier class of adverbs, and the rise, since the start of the Modern English period, of the evaluative class (1991: 416-418). Very many of them develop out of VP-Advs. The process has similarities with grammaticalization (chapter 2.4) and with paths of unidirectional lexical-semantic internal change. It includes abstraction and subjectification of meaning, relaxation of co-occurrence restrictions, reanalysis and category change.

These changes often lead to an increase in frequency. Fig. 4.1 plots the frequencies of seven

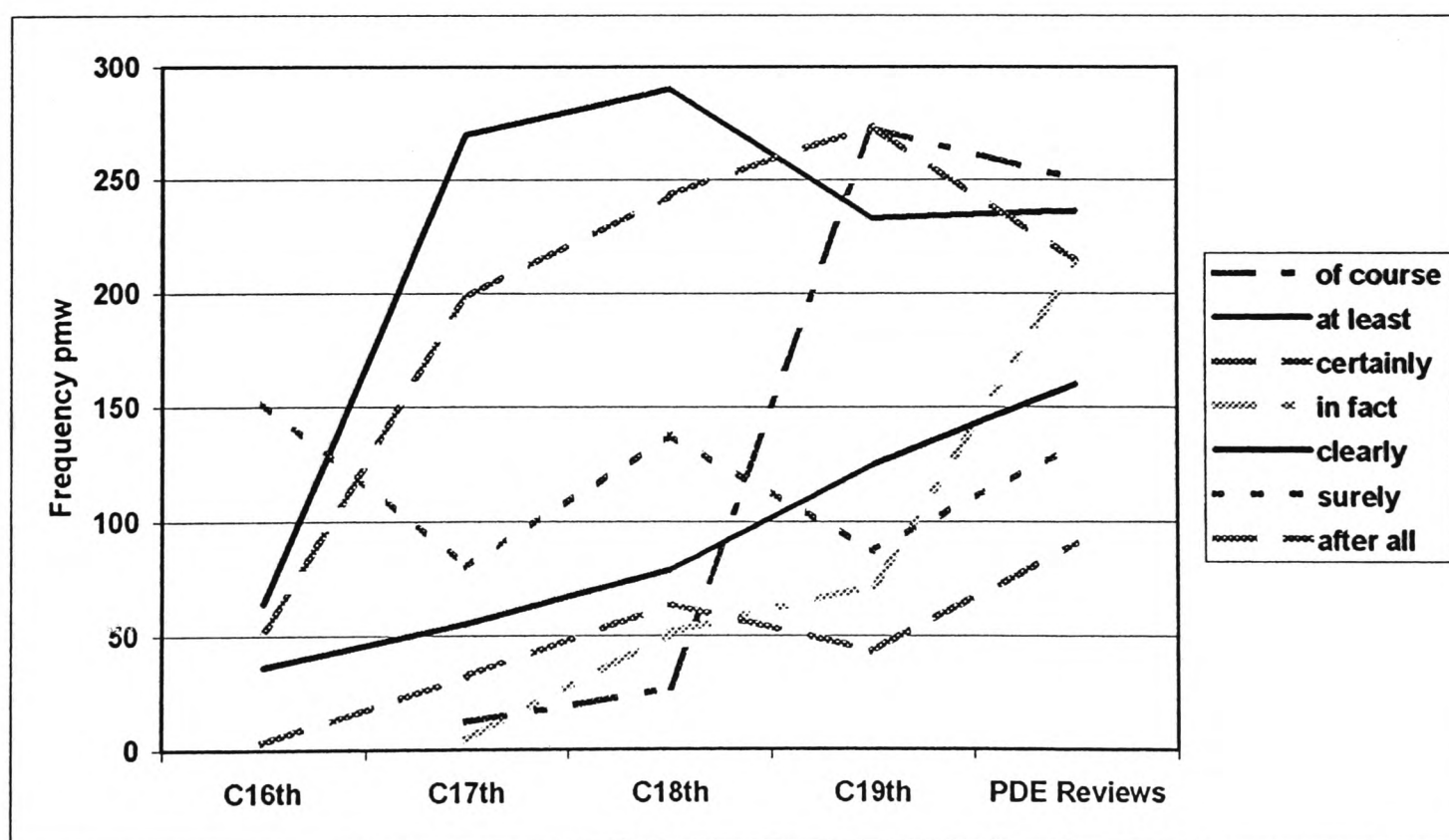


Fig. 4.1. Frequency trends of some modal adverbials

adverbs in predominantly discursive text types. It suggests that in those text types some have undergone sharp rises in frequency while others have risen slowly or remained steady. The life cycle of expressions varies greatly. For example, *of course* is a relatively recent adverb that has increased in frequency dramatically: its occurrence in PDE Conversation (based on the demographically-sampled section of the BNC) is 285 pmw. *Certainly* is a much older adverb which came into Middle English from French already functioning as a sentence adverb and has remained remarkably stable; its frequency in PDE Conversation is 102 pmw. The great differences in frequency between different text types in PDE (fig. 4.2) show that trends gleaned from historical data on particular text types cannot be extrapolated to any other types.

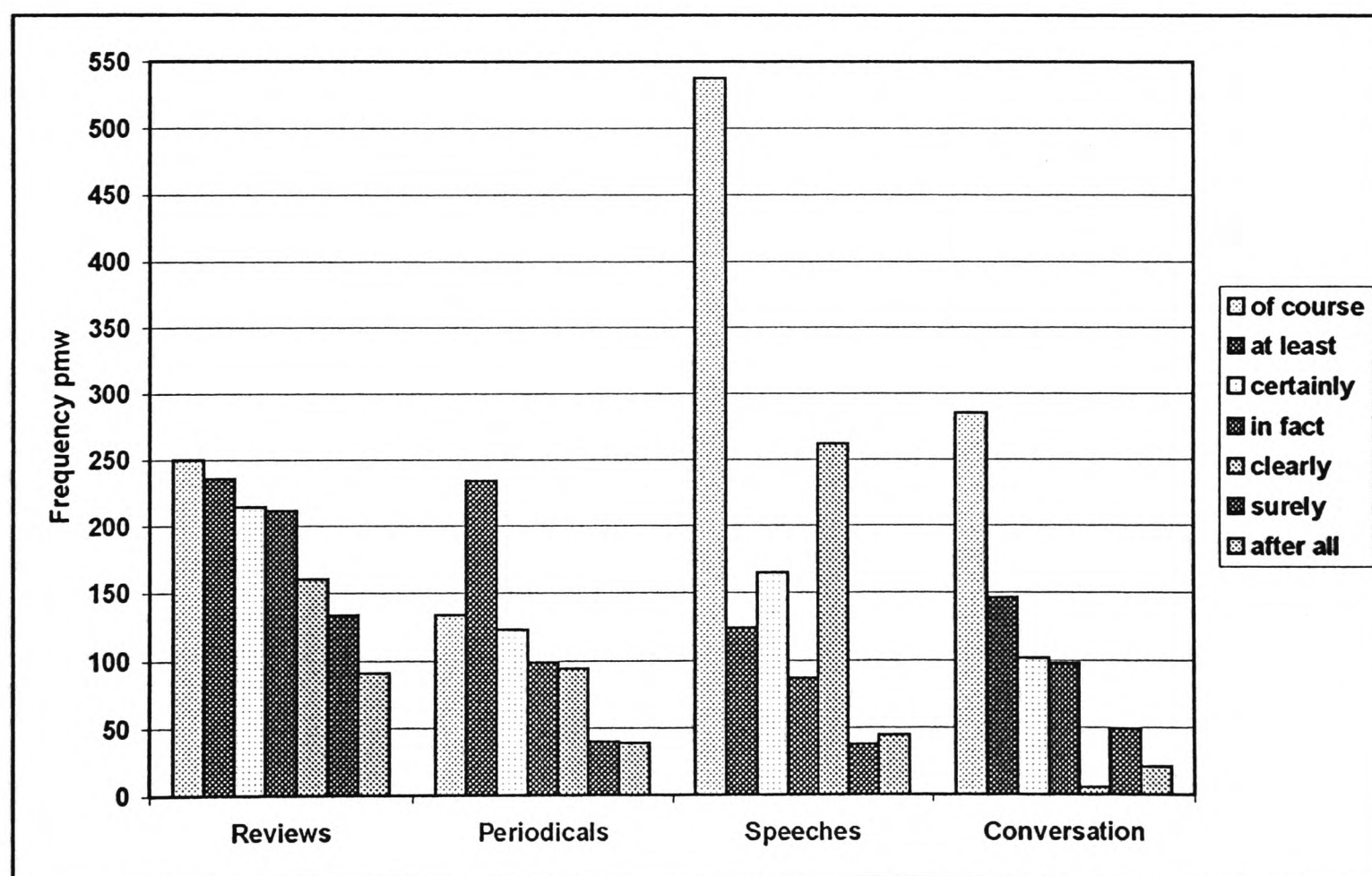


Fig. 4.2. Frequency of some modal adverbs in PDE

The semantic changes involved, it will be argued, result from usage patterns. Following usage-based accounts of language change (e.g. Croft 2000), the discourse contexts of target expressions will be

examined for clues to the bases of the meaning shifts, and for evidence of ‘utterance-type meanings’ (v. chapter 2.4.4 and Levinson 2000: 71).

This chapter looks at two *-ly* VP-adverbs, *clearly* and *surely*, which have become epistemic and then quasi-connective. It will not deal with the relationship between the *-ly* adverb and its corresponding adjective¹, but only with the adverbial PDE polysemies.

The data used (table 4.3) are taken from the full-text sources described in chapter 3 and from the OED and the MED.

Historical data						PDE data		
	C15th	C16th	C17th	C18th	C19th	Conversation	Speeches	Reviews
<i>clearly</i>	17	59	109	102	101	27	364	171
<i>surely</i>	22	180	118	175	82	207	54	144

Table 4.3 Data on *clearly* and *surely*

4.2. *Clearly*

4.2.1 Historical development

Clearly evidences the following line of development, whereby the expression gains in subjectivity:

VP-Adv ‘in a clear manner’ > S-Adv Emphasizer > ‘it is clear [to me] that’ > ‘I deduce that’

Example (3) shows the VP-Adv use.

- (3) ‘I wolde,’ quod I, ‘that thow madest me clerly to undirstonde what thou seist’ (1380)

¹ See Swan (1988: 530) and Nevalainen (1994, 1997) on adjective-adverb pairs in Middle and Early Modern English and Nuyts (1993a) on the PDE situation.

Later use includes examples of *clearly* that are epistemic; *clearly* acts as an emphaser². For example, in (4), which recurs at intervals throughout the text it comes from, the writer seems to treat the whole sequence as a unit modifying the following idea.

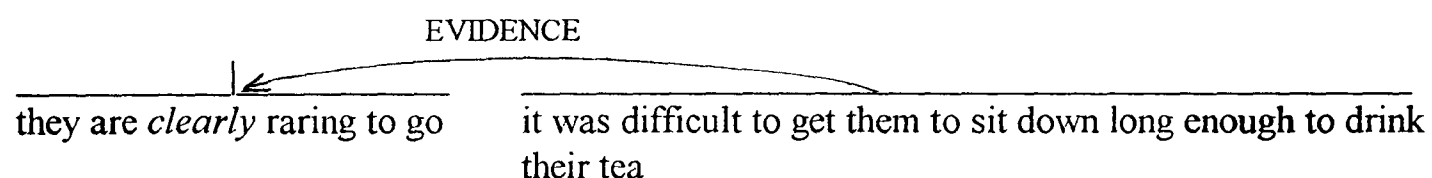
- (4) I am *clearly* of opinion.. (1797, Hist, Rumford)

At about the same time, we find examples of this epistemic sentence adverbial where it can be paraphrased as ‘it is clear that ..’ (5).

- (5) If the question is to be taken up, there is *clearly* no safe alternative but the settlement of it (1828, Hist, Peel)

In PDE *clearly* can be used as an evidential that implies ‘I deduce that’. (6) shows an example together with its RST schema..

- (6) They are *clearly* raring to go — it was difficult to get them to sit down long enough to drink their tea (PDE, Sp)



In the PDE data around a third of occurrences (27% - 36%) occur in an evidential context where one idea (the Satellite) serves as evidence for another (the Nucleus). The common pattern is

Clearly + Claim + Evidence
or Evidence + *Clearly* + Claim

Clearly is not connective where N-S order occurs in an EVIDENCE relation (chapter 2.3.3) as in (6), but in an S-N schema it begins to be: in (7a) the co-occurrence with *therefore* promotes an epistemic-causal interpretation and (7b) expresses the same relation without *therefore*.

- (7) a. Radiotherapy and radical prostatectomy are effective in treating locally confined prostate cancer. *Clearly*, therefore, some of the clinically detected locally confined tumours do not form an immediate threat to life. (PDE, Per)

² In the sense of Quirk et al., emphasizers are modal adverbs which “have a reinforcing effect on the truth value of the clause or part of the clause to which they apply” (1985: 583).

- b. Our chart more or less confirms that the more central bankers a country has, the slower is its growth rate. *Clearly* the key to economic success is to control not the supply of money, but the supply of central bankers. (PDE, Per)

4.2.2 Contexts of use

In Early Modern English, as a VP-Adv *clearly* can mean not only ‘in a clear manner’ or ‘lucidly’, but also ‘honestly’ (in expressions such as ‘deal *clearly* with someone’, ‘live *clearly*’), ‘cleanly’ (as in ‘cut/break *clearly*’ and the collocation ‘discharge *clearly*’ meaning ‘acquit’) and ‘completely’. The emphatic use exemplified in (4) may well have come from the ‘completely’ sense. In the Early Modern Period *clearly* collocates with the following verbs:

Perception, literal & figurative	see, perceive, know, understand, conceive, hear
Demonstration	show, demonstrate, manifest, appear, prove, make out
Expression	speak, declare, inform, specify, tell, answer, argue
Acquittal	discharge, pay, pardon
Accomplishment verbs	cleanse, etc.
Light	shine, etc.
Honesty	live, deal

However, the majority of verbs that *clearly* modifies, from the C15th to the present, are verbs of perception, demonstration, and expression (table 4.4).

	C15 th	C16 th	C17 th	C18 th	C19 th
Perception	46%	18%	22%	29%	37%
Demonstrat'n	5%	5%	20%	28%	18%
Expression	12%	11%	18%	12%	20%
Acquittal	12%	17%	5%		
Other	24%	46%	30%	31%	25%
Light	1%	2%			
Honesty		1%	5%		

Table 4.4. Verbs modified by *clearly* C15 - C19.

The class of ‘other’ verbs widens from actions performed *clearly* to any verbs, especially *be*. On one hand, *clearly* seems to specialize to perception, demonstration and expression contexts; on the other hand in its epistemic use it becomes suitable for modifying any verb. The inference from the former to the latter is that if something is seen clearly, it is clearly there; if someone speaks clearly, they have

clearly spoken. This inference can be seen in argumentation. For instance, Darwin (1859) repeatedly makes use of the phrase ‘we can *clearly* see/understand why..’ to introduce a deduction (8).

- (8) On my view of characters being of real importance for classification, only in so far as they reveal descent, we can *clearly* understand why analogical or adaptive character .. are almost valueless to the systematist. (1859, Hist, Darwin)

The implicature is that the conclusion is clear from the premise; that analogical/adaptive characters are clearly almost valueless.

Clearly, then, has come to signal ‘I deduce from this evidence’. It may in time become a fully-fledged connective on the attitudinal-epistemic plane (cf. epistemic *because*); in PDE it merely has connective implicatures.

4.3 *Surely*

4.3.1 *Historical development*

Surely follows a somewhat similar line of development, also gaining in subjectivity:

VP-Adv ‘in a sure manner’ > S-Adv Emphasizer > ‘[I think] it is sure that’ > ‘I assume/infer that’

Surely is widely used as a VP-Adv in Middle and Early Modern English in contexts of safety and security (9) (cf. PDE *slowly but surely*).

- (9) Item, that it be *surely* fencyd .. wyth a sure hedge, bothe hyghe and stronge (c1490)

An agent is implied who will build a proper, secure fence. Again, like *clearly*, there is an implicature that if something is done securely, it is definitely done. Moreover the VP-Adv > S-Adv shift is plausibly facilitated by the type of ambiguity that results from passives over whether the action or the resulting state is expressed. If a state, the ‘definitely’ reading is salient. The modifier can then be applied to any context where something may definitely be the case (10). In Johnson’s words, it is “used rather to intend and strengthen the meaning of the sentence, than with any distinct and explicable meaning” (1755, entry for *surely*).

- (10) this kynd of fornace is *surely* very commodious, bycause it niedeth litle fier (1559)

Where the event is hypothetical, *surely* is chosen precisely to counter the inherent uncertainty, to hedge against it (11).

- (11) ..far greater Numbers are Kill'd and Destroy'd by Physick, who would undisturb'd have *surely* recover'd without Any (1701)

Eventually a more subjective, epistemic sense arises, explicitly epistemic in (12).

- (12) they *must* have some meaning, *surely* they are not put there by Chance (1738)

Finally, *surely* occurs in overt deductions (13).

- (13) ..'cos if they're managers *surely* they get an allowance (PDE, Conv)

Here we understand that, far from an allowance being certain, the speaker is expressing doubt whether managers get an allowance. This shift in meaning from certainty to doubt has something in common with euphemism chains: an attempt by the speaker/writer to counter a default inference, with the inference proving stronger in the end.

From the C15th, *surely* occurs in initial position. In PDE Conversation it is in initial position in 62% of cases, which is curiously close to the figure of 65% cited by González-Álvarez for the Early Modern English period (1996: 230), while written PDE shows 70% in medial (pre-verbal) position. *Surely* is often treated as an interrogative marker, querying a stated or retrievable idea that runs counter to the speaker's belief. *Surely* seems to be moving to final position (29% in Conversation) in which case an answer is more likely to be forthcoming.

Where is it not interrogative, it acts as a hedge for evaluation (14).

- (14) a. it is *surely* unfair that .. (PDE, Sp)
b. M.S.'s witty book - *surely* one of the best ever on the subject - .. (PDE, Rev)

The addition of *surely* puts the claim beyond dispute, since paradoxically it reduces speaker commitment to its truth, by presenting it as the case only in the speaker's opinion, which only the speaker can know.

4.3.2 Contexts of use

As well as verbs of construction and physical activity, *surely* in Early Modern English collocates with accomplishment verbs. The contrast between the two adverb types VP-Adv and emphasizer S-Adv, the former involving an agent, the latter a patient, can be seen in (15).

- (15) a. .. the egge wyll kyll *surely* the worme in thee herball (1526)
b. he shall *surely* die that did it (1608)

Two typical discourse contexts of the emphasizer *surely* can be identified for the C16th to C17th. The first is occurrence with verbs of cognition such as *think*, *reckon*, *know*, *be convinced* etc. (16) and with expressions of value judgement (17). These account for between 11% and 15% of occurrences.

- (16) a. *surely* for mine owne part, i am perswaded .. (1599, MEMEM, Rainolds, Stage-Playes)
b. *Surely* I thinke myselfe mucche discredited, (1573, MEMEM, R. Lever, Witcraft)
- (17) a. I .. thought in my self, *surely* this is a very vayngloriows asse (1582, Helsinki, Madox, Diary)
b. *Surely* this ys a noble fellow (c1580, Ceecs)

The second is where an inference is made, *surely* being an evidential marker. In 8% - 10% of its occurrences in the data, *surely* signals an inference and the motivation for it is given, expressed by patterns such as 'if p, *surely* q', '*surely* q, for/because p', '*surely* q, otherwise p', etc. (18).

- (18) a. ande *suerty* they wolde nott have electidde hym yff fere off there persons hadde not dryven them thereunto (c1519)
b. For *surely* if his Highnes might inwardlie see my true minde such as God knoweth it is, it wolde (I trust) sone aswage his high displeasure. (1534, Helsinki, More, Letters)
c. and deliuered him to the King, who .. sayde, *surely* the Scottes mighte haue sente him to me, for I can speake French (1580)
d. *Surely* those brethren .. spake as they believed, .. Otherwise it had been needless for St Paul to admonish them .. (1590)

Surely then begins to occur without appeal to particular evidence (19).

- (19) a. *surely* his expences *cannott* be lytle (1586)
b. But *surely* they who make the exception *must* be Men either of weak understandings, or of very bad consciences (1689)
c. But *surely it might be expected* ... that he would have told us .. (1738)

All these are related in that all emphasize an opinion, attitude or belief of the speaker, increasingly subjective, till in PDE the sense is more conjecture than certainty. PDE *surely* continues to collocate with modal verbs in unrealized situations.

4.4. Conclusion

The *-ly* adverbs *clearly* and *surely* both undergo subjectification of meaning, category change and generalization to new contexts. They both ‘weaken’, from certainty to supposition. This weakening of the ‘probability value’ of modal adverbs is apparently common to very many European languages (Ramat and Ricca 1998: 236-237). Earlier, emphatic uses that express real certainty are still current. The collocational patterns with particular verbs or verb types are robust and durable. The wider discourse context can also be influential, especially typical patterns of argumentation and information structure. Overall, the data suggest very gradual change.

The following chapters discuss four non *-ly* modal adverbs. It will be seen that the overlap among ‘senses’ is considerable, or rather that there is a continuum of context types in which they occur, with frequency clusters of these facilitating the generation of ‘utterance-type meanings’.

CHAPTER 5

AFTER ALL

Then seek this path that I to thee presage,
Which *after all* to heaven shall thee send.
Spenser, *The Faerie Queen*, 1590

5.1 Introduction	122
5.2 The history of <i>after all</i>	124
5.2.1 The data	124
5.2.2 Rhetorical patterns of <i>after all</i>	124
5.2.3 Summary of historical development	128
5.3 <i>After all</i> in PDE	129
5.3.1 Distribution	129
5.3.2 Counter-expectation <i>after all</i>	130
5.3.3 Justificative <i>after all</i>	131
5.3 Summary	134

5.1 Introduction

Chapters 5-8 examine, via closer analysis of four expressions (*in fact*, *after all*, *at least* and *of course*), the hypothesis that modal adverbials tend to be recruited to discourse connective functions via a grammaticalization-like process which is rhetorically motivated.

Despite their diverse origins, the four expressions can all be said to be adverbs of certainty in at least some of their uses. All four expressions derive from PPs and all can still be used as PPs, if only in frozen or semi-frozen expressions or idioms (e.g. *as a matter of course*, *in fact vs in fiction*, *after all that*, *at (the) least vs at (the) most*) and with reduced relative frequency. All have become S-Adv and have taken on discourse connective functions. All occur in at least one of their senses in clause-initial position and all can be used parenthetically. All have increased in frequency. In other respects, however, they are different.

First, they show differences in their historical development. For example, *at least* was taken into Middle English from Old English and there are examples of clause-modifying *at least* dating back to the C14th, but the connective use may still be transparently related to the quantitative use for many speakers. *After all* and *in fact*, on the other hand, became sentence adverbials relatively recently (late C17th - early C18th), yet in each case the S-Adv has already developed a pronounced split and distinct

connective senses. Sentence adverbial *of course* dates from the C17th, but has so far developed less obviously connective senses than the other three adverbs.

Second, they are at different positions on a polysemy continuum (see chapter 1.4.6). While *in fact* and *after all* each have two uncontroversially distinct senses, the various senses of *at least* and *of course*, insofar as they can be distinguished¹, are closer together and more transparently related: the split is perhaps less advanced. Fig 5.1. illustrates the view that *in fact* and *after all* are polysemous to similar degrees, while the various uses of *at least* and *of course* are less clearly distinguishable and it is more appropriate to speak of contextual modulation.

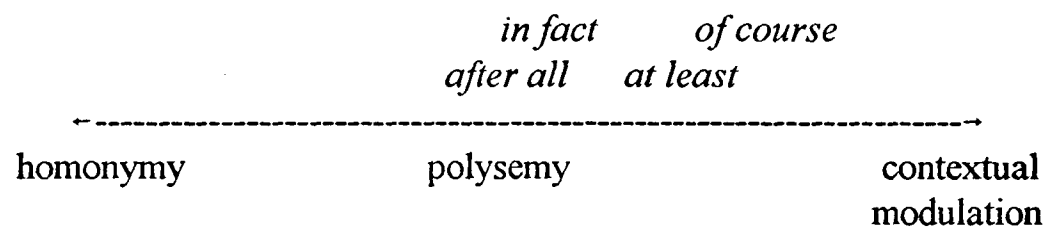


Fig. 5.1. Four expressions on a scale of polysemy

The four adverbs therefore have followed similar trajectories but to differing degrees and at different rates. By looking at both diachronic change and synchronic variation, it is possible to examine the nature of the functional splits in the expressions, and characterise the new polysemies.

This chapter deals with a fairly recently developed sentence adverb which has split and has two distinct uses in PDE. Three dictionaries were consulted for lexicographical definitions of *after all*: a traditionally-compiled synchronic dictionary (*The Collins Dictionary of the English Language*, first published 1979, 4th edition 1998), a traditionally-compiled historical dictionary (the *Shorter Oxford English Dictionary* (1983) and a corpus-based dictionary (the *Collins Cobuild English Language Dictionary*, 1987). The *Shorter Oxford English Dictionary* includes the citation *The Roman occupation was, after all, very superficial* to illustrate the sense of temporal *after* ‘following in time’, etc., but does not have an entry for an adverbial *after all*. The other dictionaries identify two distinct senses for *after all* as follows:

Collins Dictionary
after

¹ There is prosodic evidence for identifying more than one sense in all four forms.

11. **after all**. a. when everything is considered: *after all, why worry about it?* b. in spite of expectations, efforts, etc.: *he won the race after all!*

Cobuild Dictionary

all

8. **After all** is used 8.1 when you are stating a reason or opinion that relates to the previous statement.

... 8.2 when you are saying that something is or might be the case, in spite of things that have happened or that have been said or done.

The first we shall term ‘Justificative’, the second ‘Counter-expectation’. The origins of the Justificative and Counter-expectation senses are discussed in the next section. The following section looks at their frequency and distribution in PDE.

5.2 The history of *after all*

5.2.1 The data

348 occurrences of *after all*, listed in table 5.1, were examined. *After all* in both the historical and the PDE data is the least frequent of the four adverbials studied. The numbers are small, though on the scant evidence of these texts, the frequency has not changed very much over the last three centuries.

Period	No. of occurrences in full-text sources				No. in non-full-text sources					
	Hist	CeeCs	HelEME	Lamp	Total	OED2	Nge	Memem	Other	Total
1600-1649	1		1		2	11			1	12
1650-1699	16	1	3	22	42	25	38	4	1	68
1700-1749	23			32	55	11				11
1750-1799	27				27	11				11
1800-1849	14				14	38				38
1850-1899	40				40	20			8	28

Table 5.1. Breakdown of historical data on *after all*

5.2.2 Rhetorical patterns of *after all*

After all comes from the temporal *after* plus a NP. It is a somewhat unusual discourse marker in that it comes from uses of prepositional *after* occurring predominantly clause-initially and only later moves to final position. Unlike the other adverbials in this study, it does not seem to operate as a VP-Adv before becoming sentential.

From the early seventeenth century there is a pattern of ‘*After all* N + contrasting idea’ and this pattern continues into PDE. Common collocations are expressions of endeavour such as *pains*,

labour, travail (1).

- (1) a. Yet *after all heuynesse, penaunce, and dysconfyture*, She reioysed in soule. (1513, OED)
 b. Sit downe, And rest you, *after all this passed thrall* ['distress'] (1609, OED)

After all + N (prepositional *after* with a NP) is temporal. Over the C17th and C18th, two changes seem to occur apparently simultaneously: one is that *after all* acquires contrastive implicatures as a result of its collocation with contrastive contexts; the other is that *after all* as a semi-lexicalized expression generalizes to mean 'after some unspecified events', then 'in the end', and finally to introduce a conclusion or generalization.

In the following examples *after all* is still temporal and is part of a PP. The second event in the sequence is inconsistent with the first; that is, what would normally be expected, such as reward for endeavour, as in (1), does not happen (2).

- (2) a. tis now said that *after all Opinions to ye Contrary*, Genll Montecuculy will be againe sent to Comand ye army on ye Rhyne (1673, Newdigate)
 b. But, *after all his pains*, as much of his Ball lay bare as before. (1690, Hist, Locke)
 c. .. if *after all the Labour and Pains* of Acquiring that Knowledge .. their bread shall be taken off their Trenchers .. by every Illiterate Pretender (1712, Lampeter, Science)
 d. It must be a shock to find that, *after all* your effort and in spite of your qualifications, Price Waterhouse does not have any work for you. (PDE, Per)

After all in (2) is part of an adverbial phrase which provides the temporal circumstance in which to interpret the idea in the following clause. The adverbial can be considered a framework-setting Satellite discourse segment which precedes a Nucleus segment containing the unexpected second event. But in addition to its temporal sense, *after all* in these contexts already carries contrastive (concessive) implicatures and can be rendered as 'despite'.

After all as self-contained adverbial phrase is found in the C16th and C17th with a purely temporal sense: *all* refers to the immediately preceding events that have been described (3):

- (3) a. .. and doctur Whyt bysshope of Lynkolne dyd pryche at the sam masse; and *after all* they whent to his plasse to dener (1559, Helsinki, Machyn, Diary)
 b. there he made me right welcome, both with varietie of fare, and *after all*, hee commanded three of his men to direct mee to see his most admirable Colemines (1630, Helsinki, Taylor)

From about the mid C17th, this *after all* without an additional N occurs in contrastive

contexts where the temporal succession idea is less important and the earlier events are presupposed rather than mentioned. The temporal sense is thereby weakened (4).

- (4)
- a. .. my Lord avoided speaking with him, and made him and many others stay expecting him, while I walked up and down [with my Lord] above an hour ... And *yet, after all*, there has been so little ground for his jealousy of me, that I am sometimes afraid that he do this only in policy to bring me to his side by scaring me (c1666, Hist, Pepys, Diary)
 - b. I formerly told you it was said in Holland that 9 or 10 dutch men of War .. were gone out on some private designe as was beleived out of Europe, *but after all* it appears in reality these ships are gone to joyne Leiutenat Admirall de Ruyter in the mediterranean (1676, Newdigate)
 - c. I have revolved this Sentence in my Mind till I have quite tired myself, *but cannot, after all*, find any Meaning in it. (1739, Hist, Review of Hume)
 - d. His expressions there would not do for one who had rather more acuteness, penetration & taste, than love, which was your case. And *yet, after all*, I am surprised that the change in your feelings should be so great. (1814, Hist, Austen, Letters)
 - e. a letter leaves you more contented; and *yet, after all*, I am very glad to get notes (1857, Hist, Gaskell)

In (4) *after all* can be interpreted either as a PP or as an adverb depending on how far it may have lexicalized. The contrastive contexts look like a natural extension from the typical contexts of PP use exemplified in (2). The contrastive or antithetical relationship between the two ideas is marked by a conjunction such as *but* or *yet* preceding the *after all*. This co-occurrence with *but* would strengthen the contrastive implicature in *after all*, so that it could eventually indicate counter-expectation when there is no *but* present and when there is no explicitly mentioned contrasting event (it never becomes a full contrastive connective). *After all* becomes a signal of counter-expectation, meaning ‘despite what might have been thought’ (5).

- (5)
- a. .. they thought themselves past danger & were *after all* Cast away [at sea] (1688, Newdigate)
 - b. A man who has studied logic all his life may be, *after all*, only a petulant wrangler. (1774, OED)
 - c. ..when I first read that, I drew a deep breath of relief, and said to myself: ‘*After all*, there is a stretch of humanity beyond Franklin’s victorious good sense!’ (1882, Arnold, Culture and Anarchy)

Contrastive conjunctions continue to collocate with *after all* in PDE. This redundancy in the marking of rhetorical relations and evaluative comments is widespread and is arguably one cause of the ongoing reweighting of implicatures within modal expressions.

When *all* replaces NPs describing specific events, *after all* comes to mean simply ‘in the end’.

It occurs with both overtly contrastive contexts, as above, and others that are less obviously contrastive. The development of conclusive *after all* overlaps with that of the counter-expectation marker. The following examples (6), in which both the counter-expectation and the ‘in the end’ ideas fit, span two centuries. The *after all* host idea is typically negative, or smaller than or less than expected (*merely, but* in (6c,d)); that is, it suggests polarity or scalarity.

- (6)
- a. ...ils s’en revenoyent, & se mutinoyent. ... that made them *after all* into mutinie (1611, EMED, Cotgrave)
 - b. We do not now again talk so much as formerly of Don Juan’s journey into Italy, which, *after all*, it’s believed his Highness is not thoroughly disposed to (1675, The London Gazette, Nov 4-8)
 - c. The question is, perhaps, *after all*, a question *merely of words* (1798, Hist, Malthus)
 - d. Some incarnate Mephistopheles, to whom this great terrestrial and celestial Round, *after all*, were *but* some huge foolish Whirligig (1831, OED)

The contrastive and conclusive components of *after all*, one or the other more salient depending on context, continue to co-exist into the C19th. It is the conclusive use that develops into Justificative *after all*. ‘In the end’ naturally occurs at the end of an argument, especially to introduce a conclusion or a generalization (7).

- (7)
- a. ... it may be of dangerous consequence to them either to oppose or neglect us: Whereas by joining cordially in this matter, they may unite us inseparably to themselves for ever, enrich their own Nation, secure and advance the Protestant interest, keep the Ballance of Europe in their hands, and prevent the returns of its danger, their own expence of Blood and Treasure to save its being threatned with Slavery any more, either by the House of Bourbon or Austria. Therefore we cannot believe *after all*, but our wise and politick Neighbours will at last see it their Interest to protect and encourage us in this matter .. (1699, Lampeter, Politics)
 - b. they, perhaps, loved, where we only pity; and were stern and inexorable, where we are not merciful, only irresolute. *After all*, the merit of a man is determined by his candour and generosity to his associates, .. not by moderation alone. (1767, Hist, Ferguson)
 - c. The preference of brag over speculation [card games] does not greatly surprise me, ... but it mortifies me deeply, because speculation was under my patronage; and, *after all*, what is there so delightful in a pair royal of braggers? (c1813, Hist, Austen)
 - d. Reid in particular is ever appealing to men’s actions and language, as proof that there must be certain principles, beliefs, and affections in the mind. Still this evidence ever carries us back to consciousness, as *after all* both the primary witness and the final judge of appeal (1875, Hist, McCosh)

Here *after all* no longer refers to ‘after’ in narrative time, but rather to ‘after’ in the line of argument, and it can be paraphrased by ‘in sum’ or ‘after considering everything’². It seems to still behave as an independent segment: it occurs parenthetically and tends to focus its host idea. By the early to mid C19th, *after all* is only weakly contrastive in contexts where it introduces a conclusion or a generalization or a final judgment about a topic. But it still carries the implicature that its host is true despite everything, and it therefore conveys epistemic certainty. In examples (7) we can see incipient Justificative *after all*: in (7a) the conclusion follows from the arguments (*therefore*); but in (7b,c,d) it is possible to interpret the ‘conclusion’ as a truism used to justify the preceding idea. That is, an idea that has been shown elsewhere to be certain is invoked to support an argument. *After all* introduces a backgrounded idea: the addressee is reminded of the idea rather than presented with it.

Ariel refers to (Justificative) *after all* as a ‘sentential accessibility marker’ (1998: 243); that is, it marks its host as ‘given’. Appeal to a shared, non-disputed idea as a premise to support a new idea is a typical argumentation strategy (cf. *of course*, chapter 8). Hence the Justificative use.

Traugott (1997) claims that the Justificative use dates back to the early C18th; her example is

- (8) You need not be much concerned at it; for *after all*, this way of explaining things, as you called it, could never have satisfied any reasonable man (1713, Hist, Berkeley)

According to Traugott, “here *after all* justifies *you need not*, and follows *for*, which signals locutionary ‘because’” (1997: 6). There are a few instances of *for after all* in Berkeley’s writings (1710, 1713), and one *as after all* in Hume (1748). It seems likely, however, that at this stage the causal notion is carried solely by the *for* or *as*. Even if *after all* is indeed Justificative in these examples, they seem to be isolated cases and not to represent a generalized implicature. What is salient is the ‘in the end’ reading.

5.2.3 Summary of historical development

Early adverbial uses of *after all* (end of the seventeenth century) can be characterized as ‘in

² Johnson’s (1755 and 1773) definition is ‘when all has been taken into the view; when there remains nothing more to be added; at last; in fine; in conclusion’.

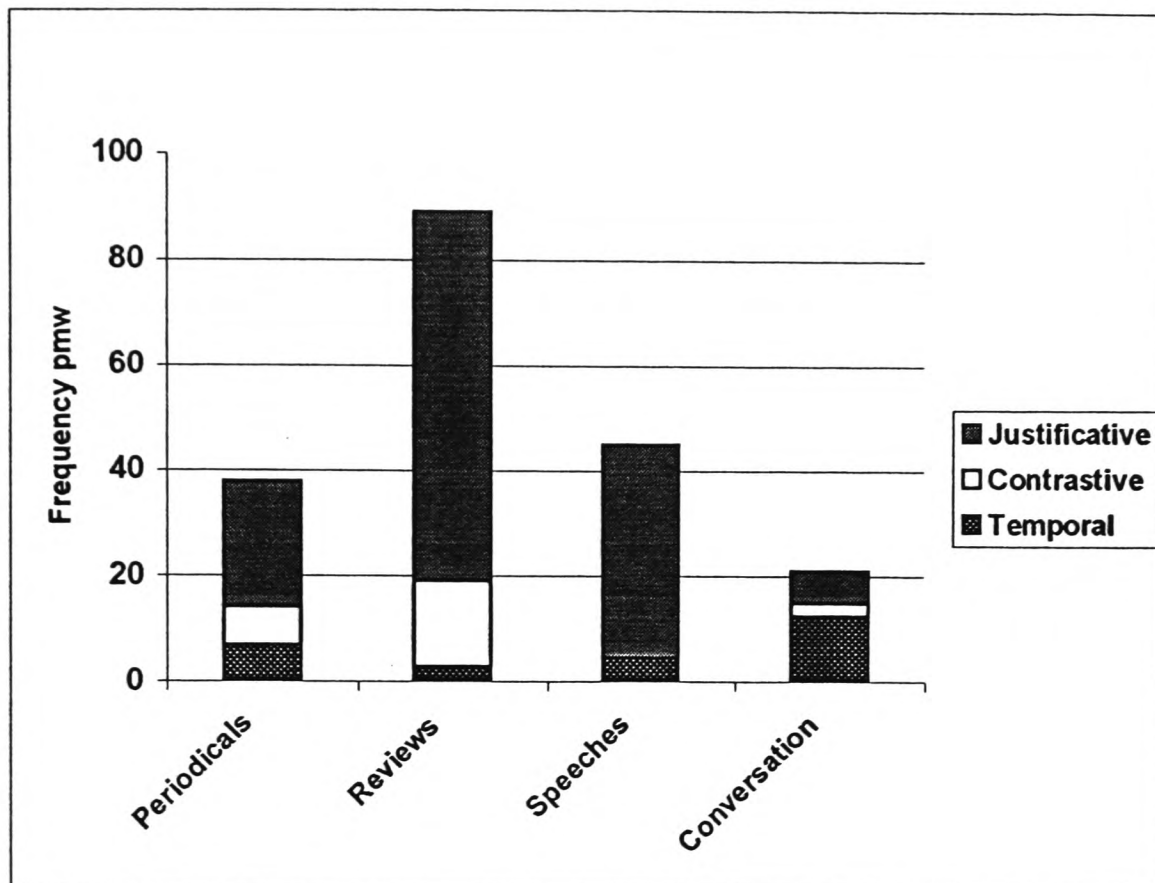


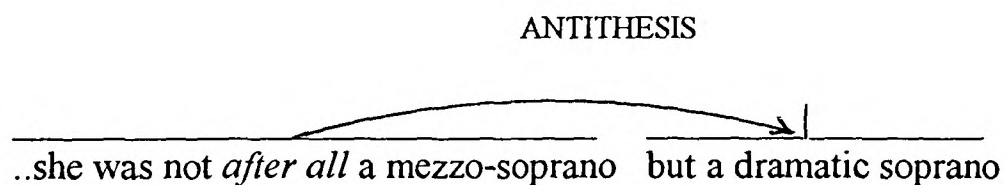
Fig. 5.4. Distribution of *after all* across three genres

In formal registers involving argumentation, therefore, S-Adv *after all* is primarily used as a Justificative.

5.3.2 Counter-expectation *after all*

Occasionally, the contrasting idea in Counter-expectation *after all* is mentioned (10), but more often it is presupposed. Contexts are still scalar or polar, and the contrast relation ANTITHESIS.

- (10) It was a shock .. to discover that she was not *after all* a mezzo-soprano but a dramatic soprano (Rev)



As mentioned above, *after all* starts out in initial position and later moves down the clause (cf. *though*). In PDE Counter-expectation *after all* can no longer occur in initial position (cf 5d above).

The changing information structure of its contexts provides a clue to this development. As part of a PP, it occurs in the Satellite of temporal CIRCUMSTANCE relations where the typical order is S-N, since in the line of narrative the Satellite is an anaphoric link. This referring back function continues when *after all* loses its reference to a specific event: *after all* itself becomes the whole Satellite. When it loses its reference function altogether and simply marks counter-expectation and ‘in the end’, there is no longer a clear S-N relation between two ideas on the content plane, but rather a single discourse unit with a comment attached. *After all* now operates at least partially on the attitudinal plane as a speaker comment. In initial position, *after all* without *but* is seldom strongly contrastive at any period, but once in final position it marks clear counter-expectation. It is not clear whether the change of position or the strengthening of contrastiveness came first. In final position it is not parenthetical in PDE: in contrast with final Justificative *after all*, it does not have its own intonation group. It is thus no longer a (satellite) discourse segment itself, but seems to be the focus of the clause. With its move to the end of the clause it becomes integrated into the clause and gains in information status. This seems to happen at the same time as it moves on to the attitudinal plane.

5.3.3 Justificative *after all*

Justificative *after all* follows the pattern Claim + *after all* + Justification (table 5.2).

Claim (nucleus)	<i>After all</i>	Justification (satellite)
a. Firms will often see merger as an ‘easy way out’;	<i>after all</i> ,	nobody in business prefers to face competitive pressure (Sp, Bridgeman)
b. I often hear ... the fear that politicians might still secretly hanker after a little inflation.	<i>After all</i> ,	a little inflation leads to higher interest rates for savers ... (Sp, Clarke)

Table 5.2. *After all* in the JUSTIFY relation

The second idea is not presented as the reason for which the first event is carried out (though it might be the reason), nor is it presented as circumstantial evidence for the conclusion in the claim. Rather, the justification consists of a generalization or truism which makes the claim seem reasonable.

However, the distinction between Causative, Evidential and Justificative relations is not always clear-cut.

Justificative *after all* frequently also enters into a wider rhetorical pattern of contrast. Sixteen per cent of Justificative *after all* in periodicals are followed by a segment introduced by an expression of contrast: usually *but*, also *however*, *yet*, *although* and other contrastive expressions. There is a pattern of argument and counter-argument:

[Argument + [*after all* + Justification]] + Counter-argument

This is illustrated in table 5.3. In each case the dispreferred argument expresses an expectation that is not to be fulfilled. *After all* introduces the justification for that expectation which is at the same time a concession (fig. 5.5).

Argument	<i>After all</i> justification	Counter-argument
a. <i>We're sure</i> Moby's behaviour is simply a phase.	<i>After all</i> , he's only been experiencing the outside world since his vaccination course was completed a few weeks ago.	<i>But</i> , with him already so large and strong, we're concerned ... (6,A17)
b. The layman's reaction is <i>probably</i> 'and a good thing, too'.	Who, <i>after all</i> , is in favour of conflict for its own sake?	<i>But</i> the dilemma, at least for the television journalist, is not quite as simple as that. (29,A3A)
c. Panama's President Francisco Rodriguez <i>should have been</i> there, too	(it is a Group of Eight meeting, <i>after all</i>),	<i>but</i> his country is in the doghouse ... (33,A3U)
d. It would <i>probably</i> unleash another burst of investment as firms took advantage of opportunities which had hitherto been regarded as too risky.	<i>After all</i> , businesses can insure against the exchange rate risk on particular deals ... by accepting less sterling today in exchange for the known future quantity of francs.	<i>But</i> they cannot hedge investments in the same way .. (61,A7T)
e. It <i>may seem</i> misguided and unnecessarily pedantic to take issue with the White Paper's notion of punishment.	Punishment is just a word <i>after all</i> : nothing more than convenient political rhetoric ..	<i>However</i> , this fails to acknowledge a fundamental axiom of social work practice; namely that language not only reflects but also shapes social reality. (138,ALP)
f. He has, <i>you might suppose</i> , some right on his side.	<i>After all</i> , why should he consent to be Helena's fairy-tale prize for curing the King, if he cannot love her?	<i>But</i> your ability to sympathize with him is powerfully lessened by the fact that this hero is also an immature prig and lying snob.. (39,A4S)

Table. 5.3. *After all* in counter-argumentation strategy.

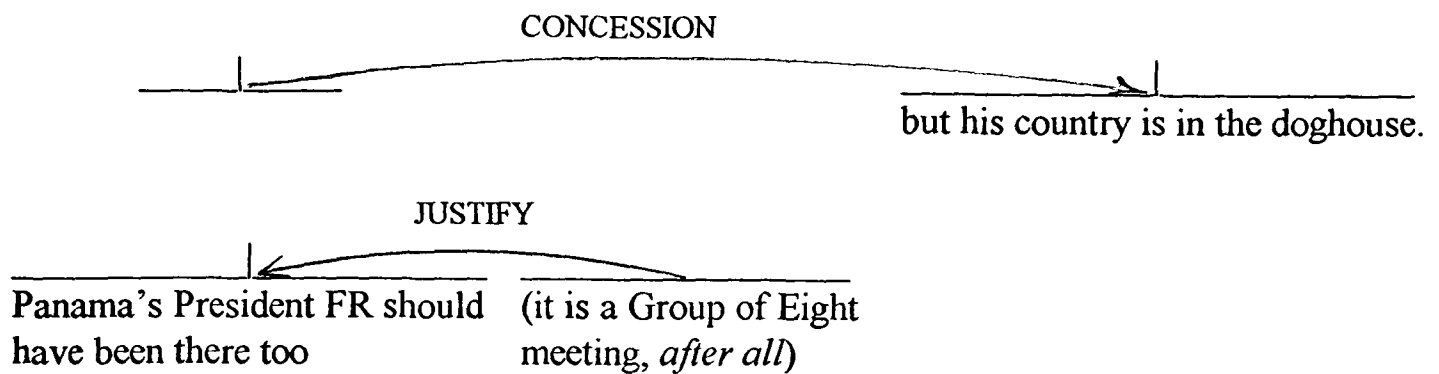


Fig. 5.5. RST schema for example c in table 5.3

After all, in this pattern, is associated with a weak, dispreferred assertion. *Of course* occurs in a similar, typical argumentation pattern (chapter 8).

Medial parenthetical position seems to be favoured for Justificative *after all* (table 5.4).

	Initial	Medial	Final	Total
Reviews	33	34	3	70
Speeches	21	34	1	56
Conversation	14	0	6	20

Table 5.4. Positions of Justificative *after all* in three genres

Traugott doubts “that there is a real distinction to be based on whether a DM [discourse connective] can or cannot initiate a turn, despite efforts to find such a distinction” (1997: 7). The evidence of *after all*, however, suggests there is a distinction between a marker that allows a presupposition to be retrieved and one which does not. Blakemore (1996: 337) claims that (Justificative) *after all* can be used in non-linguistic contexts. Her example is

- (11) [the speaker takes an extremely large slice of cake]
After all, it is my birthday.

(Blakemore 1996: 338)

But the point is that Justificative *after all* cannot begin a discourse. An immediately previous idea is required for *after all* to refer back to. The fact that in particular circumstances that idea can be non-verbally expressed does not alter this requirement. By contrast, a clause with Counter-expectation *after all* can begin a discourse because it contains a presupposition: the notion that the antithesis was

expected (12).

- (12) David Gower will tour India *after all* — as part of Sky Sports' commentary team
(Per. Opening sentence of an article in the *Today* newspaper, 1992)

5.3 Summary

After all has undergone abstraction, subjectification and layering, that is, the types of semantic change associated with grammaticalization. It has been shown to occur in very similar discourse contexts over several centuries as illustrated in (2), (6), etc. PDE usage of *after all* as part of an adverbial PP has changed little since the C17th. In around half the cases of temporal *after all* in the PDE data there is a contrast between the earlier and the later event, as in (2) above. *After all* therefore exemplifies layering of senses over time (fig. 5.3, see also chapter 2.4). Older senses do not change into newer ones; rather, there is a gradual reweighting of implicatures in certain contexts that eventually results in the synchronic polysemy we see in PDE. Although there has been a relatively clear split in *after all* to result in the current polysemous situation, the relatedness of the three uses is not opaque. It is easy to see that example (13a) (a paraphrase of (2d)) can become (13b) and imply (13c)

- (13) a. *After all your hard work*, Price Waterhouse has no work for you.
b. Price Waterhouse has no work for you *after all*.
c. Price Waterhouse should have work for you; *after all* you worked hard.

After all exemplifies subjectification in semantic change: from marking a temporal relation (content plane), it comes to express a speaker attitude, both as a counter-expectation marker and as an indicator of the Justification rhetorical relation.

The historical data suggest that contrastive and epistemic implicatures crystallized over a long period to produce the two lexicalized polysemies observed in PDE, and that this was promoted by regular occurrence with the same rhetorical relations and thus with the same information structure. It was seen in ch. 2.3 that foregrounding is the default while backgrounding needs marking. *After all*, as this chapter has shown, acts as a typical backgrounder. The information status of the Justificative *after all* host is reduced, for *after all* presents it as given or accepted, not as new. In the case of Counter-

expectation, what is new is usually the polarity, not the idea, which again is presented as accessible in the context. This backgrounding function is already present in the PP temporal adverbial, which presents an event (e.g. ‘they took pains’) as a given (e.g. ‘all their pains’). Haspelmath notes that “functionalists have tended to focus more on less tangible notions such as iconicity and metaphor, leaving economy accounts insufficiently explicated” (Haspelmath in press). Although Justificative *after all* meaning ‘after all the arguments’ may appear retrospectively like a metaphorical use of an originally temporal *after all*, there is no evidence of a metaphorical leap in the history of the expression. Nor is there evidence for the kind of ‘ad hoc innovation’ that according to Koch induces metonymic polysemy (1999: 140). The contexts in which *after all* develops are not quirky or innovative ones, but regular, established ones where the expression has its prototypical sense. The adverbialization of *after all* is better seen as an evolved shorthand, a form of economy (chapter 1).

The next chapter looks at an adverb that has also diverged into two relatively clear polysemies in PDE, but which has developed different information structuring functions too.

CHAPTER 6

IN FACT

To know the nature and powers of the human voice, is *in fact*
to know the matter or common subject of language.
(Harris, *Hermes*, 1751)

6.1 Introduction	136
6.2 The history of <i>in fact</i>	137
6.2.1 The data	137
6.2.2 Development of epistemic/evidential <i>in fact</i>	138
6.2.3 Contrastive contexts of <i>in fact</i>	140
6.2.4 Elaborative contexts of <i>in fact</i>	143
6.2.5 Summary	146
6.3 <i>In fact</i> in PDE	147
6.3.1 Types of <i>in fact</i>	147
6.3.2 Emphasizer <i>in fact</i>	148
6.3.3 Antithetical <i>in fact</i>	149
6.3.4 Elaborative <i>in fact</i>	150
6.3.5 Distribution of <i>in fact</i>	152
6.4 Summary	153

6.1 Introduction

PDE *in fact* receives very different treatment by three sample dictionaries. The *Collins Dictionary* gives a single sense as follows:

fact *n.*

...

6. as a matter of fact, in fact, in point of fact. in reality or actuality.

The corpus-based *Collins Cobuild Dictionary* devotes a long space to *in fact*, in recognition of its frequency, and describes three uses:

fact

...

4 The word **fact** is also used in the following expressions. **4.1 In fact, in point of fact, or in actual fact** means in reality or in truth; often used as emphasis, or when you are asking whether something really is true or really has happened. ... **4.2** You say **in fact, in actual fact, as a matter of fact, or in point of fact** to emphasize or to introduce a contradiction or an opinion which is different from something that has just been said. ... **4.3** You say **in fact, in actual fact, or in point of fact** to introduce or to indicate more detailed information which is related to what you have just been saying.

The *Shorter Oxford* does not define *in fact*. However, *OED2* has the following entry:

fact, *n.*

...

6.b. *in fact*: in reality (cf. sense 1 [a thing done or performed] and *indeed*). Now often used

parenthetically in an epexegetical statement, or when a more comprehensive assertion is substituted for that which has just been made.

It will be seen that it is the epexegetical (introducing further elucidation) use of *in fact* that is most frequent in the PDE data, especially in Conversation.

Previous lexical semantic work on *in fact* has distinguished three different uses in PDE corresponding approximately to the *Cobuild Dictionary's* senses (Aijmer 1996: 223, Lewis 1999, Schwenter and Traugott 2000, Traugott 1998, 1999, in press). In this chapter, however, a division into four senses will be defended: a PP comparable with *in feeling, in fiction, in theory, etc.*; an emphasizer; a contrastive (antithetical) connective; an elaborative connective. The next section examines the history of these types of *in fact*, and the following section considers their situation in PDE.

6.2 The History of *in fact*

6.2.1 The data

Table 6.1 shows the occurrences of *in fact* drawn from full-text and non-full-text sources.

Period	No. of occurrences in full-text sources			No. in non-full-text sources OED2
	Hist	HelEME Lamp	Total	
1550-1599				2
1600-1649				7
1650-1699	1	1	4	6
1700-1749	2		20	9
1750-1799	40		40	14
1800-1849	42		42	69
1850-1899	120		120	97

Table 6.1. Breakdown of historical data on *in fact*

The problems with a lot of the data being from dictionary citations are those typical of using any non-full-text sources, especially lack of adequate context, and therefore potential over-representation of *in fact* in sub-sentential contexts. In addition, OED2 seems at first sight to have netted a lot of *in fact's* in definitional contexts. In practice, however, the distribution was found not to differ much from that of the full-text sources (see note 2 below).

6.2.2 Development of epistemic/evidential *in fact*

Traugott (1999) and Schwenter and Traugott (2000) posit three senses for *in fact*: (i) VP-Adv ‘in practice’ (*in fact*₁), (ii) adversative adverb combining the semantics of *certainly* and *however* (*in fact*₂) and (iii) additive rhetorical discourse connective similar to *what’s more* (*in fact*₃) (2000: 11-12). The discourse connective *in fact*₃ is said to belong in a set with *in other words*, and *I mean* and it is claimed that *in fact* exemplifies a regular pathway of change VAdv > IPAdv > DM (Traugott 1999: 177. Traugott presents DM as a syntactic category.)

Traugott (1999) and Schwenter and Traugott (2000) do not quantify historical use of *in fact*; they provide some examples. Their claim is that connective *in fact* developed out of adversative IP-Adv *in fact*: “what, exactly, is being contrasted in the discourse? At first, propositions, ... But in contexts where the contrast is at a greater distance in the discourse flow, or requires greater work to recover, the implicature that q [the *in fact* host] is a better verbal expression than p [the previous idea] will be salient” (2000: 18-19). Referring to additive connective *in fact*₃, they claim that “its origins can only be understood using the epistemic adversative adverb *in fact*₂ as its point of departure” (2000: 21).

The historical data examined here suggest that, by contrast, *in fact*₃ developed from non-adversative uses of *in fact*; that it developed from the collocation of *in fact* with elaborative rhetorical relations. Our data also suggest that *in fact* did not have, until very recently, an intrinsic contrastive sense, despite its regular occurrence in contrastive contexts.

However, this account of the historical development of *in fact* still provides support for the claim that it undergoes progressive subjectification (Traugott 1999: 189-190, Schwenter and Traugott 2000: 7, 10) and for claims regarding a unidirectional diachronic path from the content domain through the epistemic domain to the speech-act domain (Sweetser 1990).

Adverbial *in fact* dates from around the end of the C16th. Early examples of *in fact* are VP-adv in final position. *Fact* in this period usually meant ‘deed’, ‘act’, ‘doing’; facts were ‘committed’.¹

¹ This sense is retained in PDE in the fossilized *before/after the fact*.

The manner adverbial *in fact* thus means ‘in actions’ or ‘in practice’, i.e. physically and concretely, as opposed to in words or in appearance or in any other abstract mode. It often occurs with another manner adverbial, which is either consonant in meaning (1), or, more usually, contrasts with *in fact* (2).

- (1) The yeare must be alledged *in fact*, | for it may be mislaid by the plaintiffe, | and therefore the defendants .. must alledge it *precisely* (1596, OED)
- (2)
 - a. If any whosoeuer will needes be offering abuse *in fact*, | or snip-snapping *in termes*. (1593, OED)
 - b. And there is a Release *in fact*, | and a release *in lawe*. (1607, OED)
 - c. leave it to a man himself to see if he can conceive, | even *in thought*, | what he holds to be true *in fact* (1713, Hist, Berkeley)

VP-adverbial *in fact* is in use continuously till the present day, usually, but not always, in final position (3).

- (3)
 - a. Her dress is extremely modest *in appearance*, | and yet very coquettish *in fact* (1792, OED, Wollstonecraft)
 - b. The English Government ... gave the custody and profit of its own money .. to a single bank, | and therefore | *practically* and *in fact* it is identified with the Bank of this hour. (1875, Hist, Bagehot)
 - c. Its opening is the key which must unlock the door to rapid decisions on ending 45 years of East-West confrontation, *formally* as well as *in fact*. (PDE, Per)

As the noun *fact* becomes more abstract in meaning, embracing what is ‘real’ rather than only what is an ‘act’ or ‘deed’, so *in fact* too becomes more abstract and modifies states of affairs as well as actions. It is advantageous to be believed, and explicitly attributing an idea to real-world occurrence increases its face value. Thus, *in fact* becomes a type of evidential, emphasizing the validity of the idea it attaches to. If it happened in real life, it must be true: the sequence *true in fact* becomes a collocation. Examples (4) show that by the turn of the C18th, *in fact* is moving away from the literal meaning as its epistemic implicatures strengthen.

- (4)
 - a. But it is evident *in fact* and experience that | there is no such universal Judge (1671, Helsinki)
 - b. If Bullion be exported, | and that hinder not the exportation of our English Manufactures, | as *in fact* doth appear: (1681, Lampeter)
 - c. If you know *in fact* that you have any number of Men of this Stamp, | you must have felt how much our whole Constitution suffers from such infected Members (1702, Lampeter)

None of the examples in (4) involves an action: the contexts are stative. In (4a) *in fact* strengthens the

evidential *it is evident*; in (4b) it occurs with *appear*, whereas in its manner use it is often contrasted with *in appearance*; in (4c) it modifies *know* without implying other modes of knowledge (cf PDE *know for a fact*). In all the (4) examples, *in fact* can be seen as an emphazier. Its use is motivated by the need to counter addressee expectation, or to cancel a previous presupposition. To a large extent, therefore, it is already a truth hedge, an epistemic marker. In the earlier examples (3), by contrast, *in fact* is used to specify the domain in which the idea holds, i.e. as a manner adverbial. However, the content reading is still available for examples (4) and continues to be so for many occurrences well into the C19th. The shift from content adverbial to counter-expectation / counter-presupposition adverbial is gradual (see ch. 2.4.4).

The data suggest that this reweighting of implicatures encouraged *in fact* to move from final position (5) to medial position (6).

- (5) a. neither ought we to believe that | what is *pretended* to be gotten by the East-India Trade | was so in Fact (1697, Lampeter)
 b. If this were true in fact, | I don't see any tolerable colour for such a conclusion (1707, OED)
- (6) a. Reason tells us, that | in Nature it should be so; | and Experience tells us, that | *in Fact it is so* (1721, Lampeter)
 b. the Enemies of the Practice [Inoculation] have not produc'd the Names of above 3 Persons that have died; | allowing their Deaths chargeable on this Practice, | which I believe is not in Fact true. (1722, Lampeter)

6.2.3 Contrastive contexts of *in fact*

The contrastive and elaborative senses seem to have developed not in sequence, but both from this emergent epistemic *in fact* meaning 'in reality'. A split begins to occur between the emphatic, counter-expectation use (7) and the contrastive, reality vs appearance, use (8).

- (7) We would .. intreat them to ask their own Consciences, .. | Whether our Principles do not tend to keep out Popery, | and Theirs to bring it in: | Whether *in Fact* many more, than is usual in the same Space of Time, have not been perverted to it, since the late Propagation of Their Doctrines (1721, Lampeter)
- (8) Were usefulness the principal foundation of morals, | justice ought to have a much higher degree of praise and approbation attending it, than generosity, magnanimity, heroism, or any of the exalted virtues. ... | But the contrary holds good *in fact*. (1753, Hist, Commentary on Hume)

From the mid C18th, *in fact* is commonly used in contrastive contexts. But there is no evidence that *in fact* carries in itself any contrastive meaning. In the data up to the end of the C19th, *in fact* does not occur in contrastive contexts without a marker of contrast such as *but*, *yet*, *though*, etc. or a clear lexical contrast. We can therefore conclude that *in fact* is an emergent emphatic/epistemic expression that co-occurs with contrastive contexts and probably carried defeasible adversative implicatures in those contexts. The most common contrastive relation in which it occurs is ANTITHESIS. The contrast is often between appearance or belief on one hand and reality on the other. The sense of *in fact* as ‘in the real world’ is still strong. Nonetheless, the seeds of PDE ANTITHESIS *in fact* can be seen in this period.

Examples (9) - (11) illustrate the three main patterns in which contrastive *in fact* is found in the C18th and C19th.

One is *concessive (conditional) + in fact + preferred idea* (9) (fig. 6.1).

- (9) a. Whatever we may be told of Cases in modern Languages, | there are *in fact* no such things (1751, OED)
 b. the lotto of Genoa, which, | though decorated with a smooth and splendid name, | is *in fact* no more than a Pharaoh table (1787, OED)
 c. Every banker knows that | if he has to prove that he is worthy of credit, | however good may be his arguments, | *in fact* his credit is gone. (1875, Hist, Bagehot)

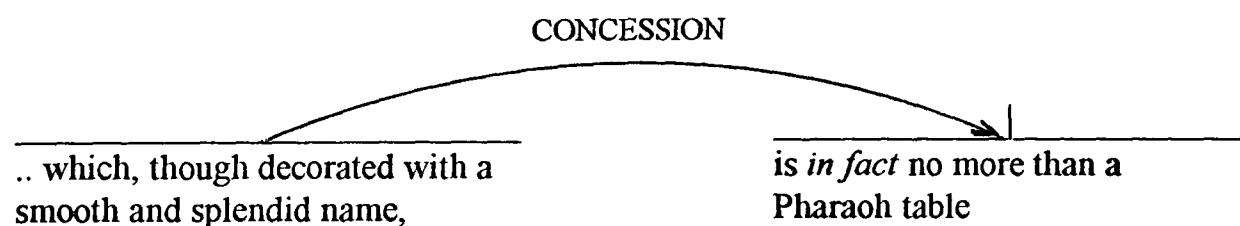


Fig 6.1. RST schema for (9 b)

A second pattern is *what seems to be x is in fact y* (10) (fig 6.2).

- (10) a. What is called a quintan, | is *in fact* a tertian (1747, OED)
 b. what resembled hares | were *in fact* hill-kids (1816, OED)
 c. His Muse is, *in fact*, a giddy wanton flirt (1818, OED)

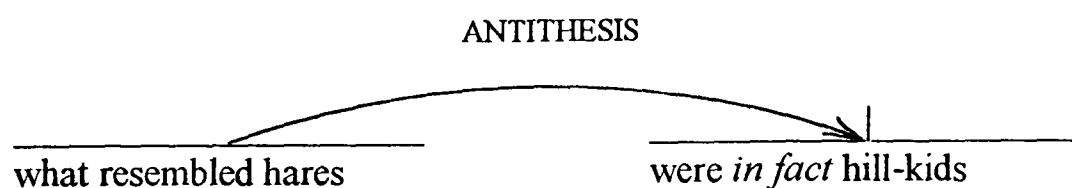


Fig. 6.2. RST schema for (10b)

The third is *dispreferred idea* + *but / when, etc.* + *in fact* + *preferred idea* (11) (fig. 6.3).

- (11) a. I made the most of the story | because it came in to advantage, | but *in fact* he only asked me whether I were to be at Sidney Gardens in the evening or not (c1813, Hist, Austen)
- b. We are apt to look upon cottages as incumbrances and clogs to our property, | when, *in fact*, those who occupy them are the very nerves and sinews of agriculture (1776, OED)
- c. [it] has sometimes been noticed ... that | we know nothing about the origin or history of any of our domestic breeds. | But, *in fact*, a breed, like a dialect of a language, can hardly be said to have had a definite origin. (1859, Hist, Darwin)

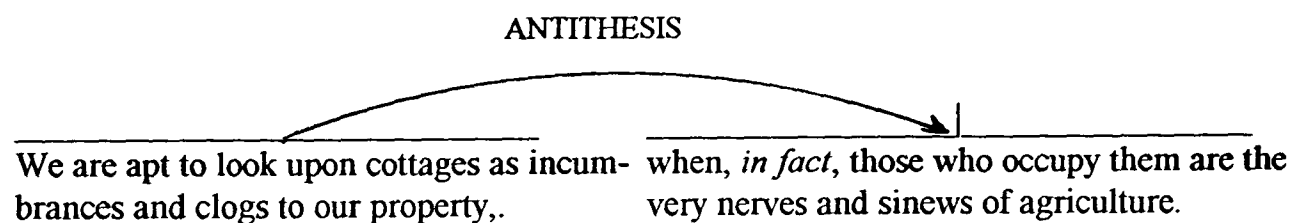


Fig. 6.3. RST schema of (11b)

In particular, as seen in (9-11), the co-occurrence of *in fact* with CONCESSION and ANTITHESIS relations in argumentation means that its host, being the preferred idea, is almost always Nuclear. Moreover, *in fact* is occurring in initial position and becoming parenthetical. It is thus poised to become the autonomous discourse segment with its own tone group that is PDE ANTITHESIS *in fact*, that is, where the *in fact* host is understood to replace a false idea.

However, in this period the *in fact* is not always attached to the Nuclear idea (12). This again suggests that *in fact* is not in itself contrastive.

- (12) It is not ... the paper that is, *in fact*, the substitute for money | but something still more exile; | the promise ... stamped upon it (1797, OED)

Table 6.2 shows that, in the data between the mid C18th and the end of the C19th², around a fifth of occurrences of *in fact* are in contrastive contexts, while around half of *in fact* tokens occur in elaborative contexts.

² The OED data and the full-text source data for the C19th compare as follows:

	VP-Adv	S-Adv (Epistemic)	Emphatic	Elaborative	Contrastive contexts
OED	8%	6%	65%		21%
Full-text	3%	15%	70%		12%

Period	Type	VP-Adv	S-Adv (Epistemic)			n
			Emphatic only	+ Elaborative context	+ Contrastive context	
1550-1599		2				2
1600-1649		7				7
1650-1699		4	1		1	6
1700-1749		20 (65%)	8 (26%)	2 (6%)	1 (3%)	31
1750-1799		5 (9%)	15 (28%)	22 (41%)	12 (22%)	54
1800-1849		4 (4%)	11 (10%)	75 (67%)	21 (19%)	111
1850-1899		15 (7%)	24 (11%)	146 (67%)	32 (15%)	181

Table 6.2. Types and contexts of *in fact* in the historical data, C16th - C19th

In sum, throughout the period mid C18th to end C19th, *in fact* seems not to signal contrast, but to have context-dependent contrastive implicatures.

6.2.4 Elaborative contexts of *in fact*

What is striking in table 6.2 is the preponderance of elaborative contexts: around two-thirds of occurrences of *in fact* in the C19th data. The seeds of Elaborative *in fact*³ can be found in the late C17th.

- (13) That the Turkey Merchants do Ship out much Cloth, I deny not; but as true it is, that they have Shipt out more Yearly since the great encrease of the East-India Trade, and since themselves have made this Complaint, than they did in former Years. | So that *in Fact* it doth not follow that the encrease of the East-India Trade ... doth hinder or diminish the Exportation of Cloth to Turkey, but rather the contrary. (1681, Lampeter)

In fact frequently appears in naturally-occurring definitions and specifications. A regular pattern is *NP + being + in fact + definition of NP* (14).

- (14) a. Gas, therefore, ... becomes a generic term, expressing the fullest degree of saturation in any body with caloric; | *being in fact*, a term expressive of a mode of existence (1790, OED)
 b. instead of composing the work in moveable type, it was set up in moveable copper matrices; | each matrix *being in fact* a piece of copper of the same size as the type, and having the impression of the letter sunk into its surface (1832, Hist, Babbage)
 c. The feature ... of placing the preposition after the noun | - making it, *in fact*, a 'post-position' - thus: He is come the village from. (1863, OED)

³ What we label Elaborative *in fact* signals relations that fall outside the classical RST definition of ELABORATION, but are not adequately described by the RESTATEMENT relation either (see Appendix A). The term Elaborative has been chosen following Halliday's slightly broader definition, in which elaboration involves restating, clarifying, refining, or adding a descriptive attribute or comment (1994: 225).

This 'definition' use of *in fact* is part of a wider use which is to introduce an Expansion of a current topic. In the C18th to C19th we commonly find *in fact* in contexts of ELABORATION, SUMMARY or INTERPRETATION. These are all Expansion relations (see ch. 2.3) and in practice they partially overlap, but they are useful categories with which to describe textual organization. An ELABORATION provides attributes of something just mentioned, often in the form of a further specification of it (15) (fig 6.4). These examples are a short step from the type in (14).

- (15) a. The beast [a rhinoceros] .. kept on an even and steady course, | which, *in fact*, was a kind of pacing (1785, OED)
 b. An Indian is always a 'heap' hungry or thirsty - loves a 'heap' - is a 'heap' brave | - *in fact*, 'heap' is tantamount to very much (1848, OED)
 c. To a University man, a Grind did not possess any reading signification, but a riding one. | *In fact*, it was a steeple-chase, slightly varying in its details according to the college that patronised the pastime. (1857, OED)

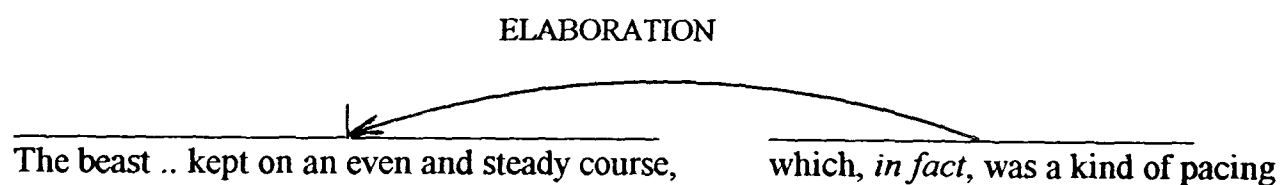


Fig. 6.4. RST schema of (15a)

A summing up of a topic consists of a shorter and often stronger expression of it (16) (fig. 6.5)

- (16) a. his business is to wash dishes, carry a lantern, | and, *in fact*, wait upon the Kitmutghaur (c1803, OED)
 b. The word theory has been perverted to denote an operation .. which .. consists in supposing and setting down matters supposed as matters observed. | Theory *in fact* has been confounded with Hypothesis. (1829, OED)

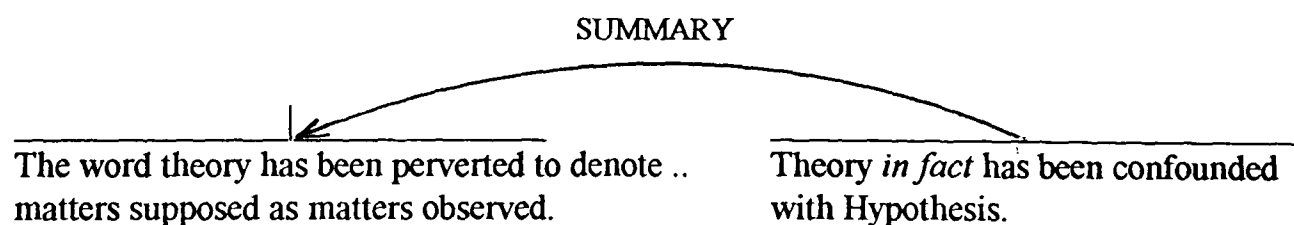


Fig. 6.5. RST schema of (16b)

Finally, *in fact* co-occurs with interpretations or conclusions drawn from the previous idea(s) in the discourse (17) (fig. 6.6).

- (17) a. .. so that, *in fact*, until reviews are established in which booksellers have no interest, they can never be safely trusted (1832, Hist, Babbage)
 b. In very many cases ... either the anthers burst before the stigma is ready for fertilisation, | or the stigma is ready before the pollen of that flower is ready, | so that these plants have *in fact* separated sexes .. (1859, Hist, Darwin)

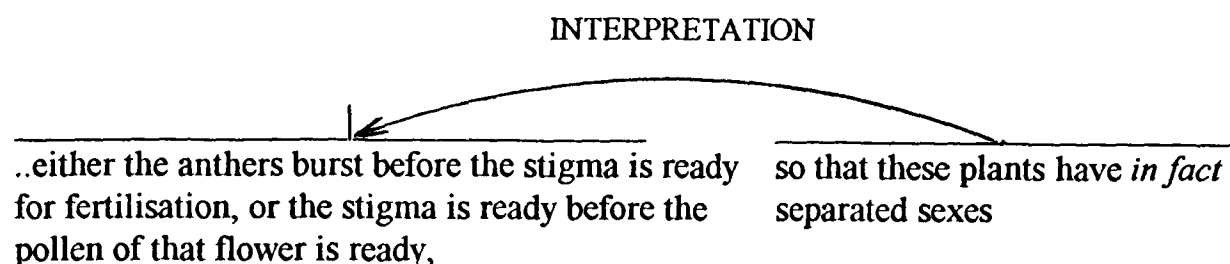


Fig. 6.6. RST schema of (17b)

In each case the *in fact* host consists of further material on the current topic, and this material is consonant with what has previously been said on that topic. In argumentation terms, the *in fact* host furthers the current line of argument. This much can be gleaned both from the sense of the contexts, and from observing that *in fact* very often follows *and*, the conjunction *for*, or *so that*. In none of these examples can *in fact* be said to be contrastive. Moreover, the *in fact* host is Satellite in the relations exemplified in (15-17) above, while in contrastive contexts it is Nuclear.

From the mid C19th, however, there is an increasing tendency for the *in fact* host not just to elaborate on the current topic, but to express a stronger version of the previous idea, although it remains a rhetorical Satellite. This is noticeable in Carlyle's pamphlets (1850), which are informal in style, and in later C19th texts (18).

- (18) a. At half-past Seven this Theatre was crowded in every part, by upwards of four hundred Students, of the most respectable description; | *in fact* we never before witnessed so genteel a Surgical class (1823, OED)
- b. the dissatisfaction with it is great, universal, and continually increasing in intensity, | - *in fact*, mounting, we might say, to the pitch of settled despair. (1850, Hist, Carlyle)
- c. As few persons now read, | or *in fact* ever did read, | through his weighty volumes, .. (1875, Hist, McCosh)

This type of elaborative *in fact* modifies an idea that is higher on a scale of strength than the previous idea. It is no longer a RESTATEMENT of the previous idea, but a reformulation of it as a stronger idea. It is the opposite of Retreat *at least*: compare *or in fact* with *or at least*, which modifies an idea that is lower on a scale of strength than the previous idea (chapter 7.2.3). In these occurrences, *in fact*, like *at least*, demotes its host to Satellite status, creating an asymmetry between two ideas. It is at this point that *in fact* can be seen as truly scalar, its primary function to signal the relative rhetorical strength of two ideas. From here develops the Additive speech-act level marker of PDE, with a meaning closer to 'moreover' than to 'indeed'. The weaker idea is by default entailed by the stronger one (Levinson

1983: 138). In (18b) above, for example, *mounting to the pitch of settled despair* entails *increasing in intensity*.

6.2.5 Summary

The development of *in fact* can be summarized as follows.

Stage I	<i>In fact</i> = in deed/act/action (cf. <i>commit a fact</i>). It often occurs with either a consonant adverbial such as <i>practically</i> or with a dissonant adverbial such as <i>in appearance, in law</i> ⁴ .
Stage II	<i>In fact</i> becomes an emphasizer, increasing epistemic commitment: if an idea is embodied in a visible or physical action, it must be 'really' true. It functions as a hedge against anticipated disbelief. A split starts to occur between the emphasizer used in dissonant contexts, i.e. in a contrast between a real idea and an unreal idea or some other mode of truth, and the emphasizer used in consonant contexts, where the current topic is elaborated upon.
Stage IIIa	Dissonant uses lead to contrastive use. Almost all contrastive contexts in the data contain a contrastive word (<i>but, though, etc.</i>) or a contrastive lexical item; this suggests that <i>in fact</i> does not indicate the contrast but merely collocates with it.
Stage IVa	<i>In fact</i> itself acquires contrastive meaning. The <i>in fact</i> host counters a false idea contained in a previous segment.
Stage IIIb	Consonant uses lead to elaborative use. To define or specify something further is often to explain what it is like 'in fact', i.e. in action or in physical form. Definitions, descriptions, specifications, summaries, restatements are all elaborations of the current topic.
Stage IVb	Elaborative <i>in fact</i> is extended to contexts where 'in reality' as in 'in the real world' no longer makes sense, and <i>in fact</i> can modify any idea however unreal. Its meaning is now closer to elaborative <i>indeed</i> , or, as an Additive speech-act marker, to <i>moreover</i> .

Fig 6.7 outlines the functional split of *in fact* into elaborative and contrastive markers.

CONTENT PLANE	ATTITUDINAL PLANE (EPISTEMIC)	(RHETORICAL)	SP ACT PLANE
		reality vs expectation →	ANTITHESIS
'in practice' →	'in reality' →	emphasizer <	
		'in its very nature' →	ELABORATION INTERPRETATION RESTATEMENT
			ADDITIVE

Fig. 6.7. Functional split of *in fact*

The data suggest that the VP-Adv *in fact* occurred more often with another contrasting

⁴ *In fact* vs *in law* no doubt stems from Lat. *de facto* vs *de jure*.

adverbial than with another consonant adverbial. However, there is no evidence that the elaborative sense of the S-Adv developed out of the Contrastive sense of the S-Adv. Elaborative, non-contrastive contexts are more frequent for the early S-Adv and continue to be more frequent in PDE. Moreover, the two senses can be seen to diverge from early on in their different information structures: *in fact* in elaborative contexts modifies a Satellite host, whereas *in fact* in contrastive contexts modifies a Nuclear host.

The development is gradual. Fig. 6.8 assumes that *in fact* becomes unambiguously contrastive when it starts to occur in clear antithetical relations without a contrastive marker. But this is a simplification, for there is no clear-cut dividing line to be drawn between implicated contrast and coded contrast in this case.



Fig 6.8. Development of *in fact*

Both the elaborative and contrastive senses of *in fact*, then, seem to have developed from the emphatic modal *in fact* meaning ‘in reality’, which in turn emerged from the VP-Adv. That is, *in fact* acquired an epistemic modal sense, after undergoing an abstraction from ‘in action’ to ‘in reality’. Then two different epistemic uses develop: one, via counter-expectation, becomes Antithetical *in fact*; the other, via a truth hedge, comes to introduce further detail about the current topic or a stronger claim.

6.3 *In fact* in PDE

6.3.1 *Types of in fact*

681 occurrences of *in fact* were analysed: 227 from the Reviews corpus, 121 from the

Speeches corpus and 333 from the Conversation corpus.⁵ The data show the semantic spectrum of PDE *in fact* to be more complex than the three-sense division outlined in section 1 above. There is not such a clear-cut distinction between contrastive and elaborative senses as at first sight appears. PDE *in fact* reflects a history of gradual layering. As mentioned in 6.1, four main clusters of uses can be identified as follows:

(i) PP used as manner adverbial:

- (19) you can say it entirely symbolically and spiritually and and things don't have to have happened *in fact* (Conv)

(ii) an emphasizer:

- (20) it is *in fact* North Radstock isn't it? (***) the actual town is North Radstock (Conv)

(iii) a contrastive:

- (21) The river just to the east of Tarsus is marked as the Goksu River; *in fact* it is the Seyhan River. (Rev)

(iv) an elaborative:

- (22) We are seeing above trend growth all with a strong emphasis on our export performance; - *in fact* we are breaking records with our exports (Sp, Aitkin)

This PDE polysemy results from the layering seen in section 6.2. The following sections examine emphatic, contrastive and elaborative *in fact*.

6.3.2 Emphasizer *in fact*

In fact is an epistemic modal adverbial of certainty; it reinforces commitment to the truth of its host. In some measure this function is common to all occurrences of S-Adv *in fact*. It is comparable to *actually*, with which it is often interchangeable. In the following examples (23), the principal function of *in fact* is to emphasize the veracity of its host idea.

- (23) a. So from London to King's Lynn is .. they say .. a hundred miles and I'll take their word for it ... London to Ipswich *in fact* is seventy miles (Conv)

⁵ All 428 occurrences in the demographically-sampled section of the BNC were examined. Removal of 16 repeats and 79 unclear instances left 333 occurrences for analysis.

- b. she's been living in a culture for four years .. where [cough] that's the way they .. they sell and buy [.....] but because she's back here .. she still sometimes does it .. and *in fact* they do do it for her (Conv)
- c. ... the group were free to consult with whatever experts they chose and ... did *in fact* contact every E.N.T. consultant in Ireland (Sp, Smith)
- d. Whether this is *in fact* the case is open to debate (Rev)

Emphatic *in fact* is a truth hedge that attaches to an idea that might run counter to addressee expectations. In examples (23b,c) above, it works with emphatic *do* to emphasize the truth of the idea.

6.3.3 Antithetical *in fact*

In initial position, antithetical *in fact* forms its own tone group: it typically bears a stress on *fact*, a higher pitch and a rising-falling intonation. Elaborative *in fact* can have this prosody too, but more often is a part of the host tone group and has a level intonation. Antithetical *in fact* does not occur in final position. It occurs with the ANTITHESIS relation, in the pattern:

False idea + *In fact* + Correct idea

ANTITHESIS is a subtype of contrast. In the ANTITHESIS relation with *in fact*, a 'correct' claim is contrasted with a claim marked by the speaker/writer as hypothetical (table 6.3, fig. 6.9).

False idea	<i>In fact</i>	Correct idea
a. The mystery virus, <i>it was thought</i> , would spread into the healthy tree.	<i>In fact</i> ,	the opposite was the case (Per)
b. <i>You're tempted to conclude</i> he's a little under-rehearsed.	<i>In fact</i> ,	it's part of a shrewdly calculated subtext. (Per)
c. .. the age profile of our troops, frequently <i>seen as</i> a problem,	was <i>in fact</i>	a benefit in SFOR (Sp)
d. <i>There were those who thought</i> we had lost our appetite for it ..	<i>In fact</i> ,	we simply wanted sufficient time to devise a truly effective regime (Sp)
e. <i>He tells us that</i> comets would "burn up" in the thin atmosphere of Mars;	<i>in fact</i> ,	they would strike its surface with explosive violence (Rev)
f. <i>I thought</i> I'd used eleven but	[<i>in fact</i>]	I've only [] used ten (Conv)

Table 6.3. *In fact* in ANTITHESIS relation

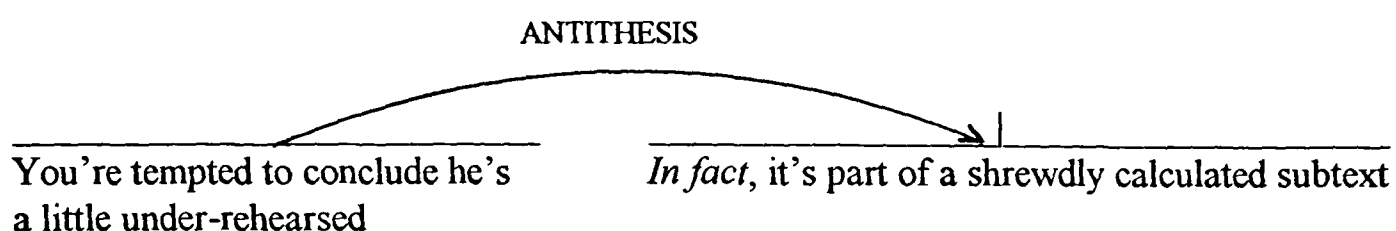


Fig. 6.9. RST schema for example b in Table 6.3.

The examples in table 6.3 are similar to those in (11) above, except that most of them do not have a contrastive conjunction such as *but*. The motivation for distinguishing this *in fact* from the Emphatic use is that it now clearly functions as a contrastive connective like *but*: its removal creates incoherence, since the two ideas are incompatible, and it is replaceable by *but*.

Schwenter & Traugott claim that *in fact*₂ (which they term an adversative sentential adverb and which combines Emphatic and Antithetical) ranks its host higher on a scale of speaker belief than a preceding proposition (2000: 13) but is not connective. Their definition of connective seems to correspond to what RST refers to as presentational relations only. Under our broader definition, Antithetical *in fact* is connective, for it expresses a rhetorical relation between two ideas. There must be a previous idea for it to refer back to.

6.3.4 Elaborative *in fact*

Example (24) can be interpreted in two ways:

- (24) we sell ... shall we say a million bottles of aspirin a year .. it is *in fact* considerably more than that (Conv)

The phrase *considerably more than that* may be interpreted contrastively, with stress on *fact*, to contradict *a million*, or it may be interpreted epexegetically as a Satellite providing further information about the quantity of aspirin. These two interpretations are incompatible. In this instance, the wider context provides the clue that *in fact* is meant contrastively: Boots sells a million bottles a year of its aspirins through its own stores, but many more than that when sales through other outlets are taken into consideration. This occasional ambiguity between Antithetical and Elaborative *in fact* shows that they are distinct.

The pattern of *in fact* in the ELABORATION relation is:

Idea + *In fact* + Elaboration of idea

Table 6.4 gives examples of this pattern from Reviews, Speeches and Conversation.

The elaboration sometimes involves giving more detail about a topic (table 6.4c). More often, the *in fact* host consists of a stronger claim. In around one fifth of cases of Elaborative *in fact*, the pattern consists of a negative claim followed by *in fact* introducing a positive stronger claim, as in the last two examples of table 6.4. The *in fact* host acts as a Satellite in the relation, but the order is always N-S, so that the Satellite acquires end-focus. Elaborative *in fact* is also connective; it must have a previous idea to refer back to.

Claim	<i>In fact</i>	Stronger claim
a. that Michael Watson give him some stick ...	<i>in fact</i>	if he hadn't have knocked him out and put him in a coma in last round ... that Michael Watson would have won it (Conv)
b. that's enough now David ...	<i>in fact</i>	it's too much of it (Conv)
c. .. was wearing a green (***) top yeah ..	<i>in fact</i>	it was quite luminous (Conv)
d. This tradition goes back longer than many of the Christian Democrat parties themselves -	<i>in fact</i>	it begins with Bismarck in the 19th century (Sp)
e. The JREI ... has proved to be an outstanding success ...	<i>In fact,</i>	the JREI has been so successful that I was pleased to have announced earlier this year that it is to be an annual event (Sp)
f. There is no evidence that Rumbold was particularly hostile to Jews by the standards of his day and class.	<i>In fact,</i>	he was later to be appalled by Hitler's violence towards the Jews (Rev)
g. A: have you got any stamps? B: no I don't think I have	<i>in fact</i>	I know I haven't (Conv)

Table 6.4. *In fact* in ELABORATION relation

In the Conversation data, *in fact* sometimes introduces a new point that is not directly related to the previous point (25).

- (25) a. Oh I'll do it myself ... hundred and fifty ohm ... think I've got one of them .. yeah
 ...*in fact* this particular chassis I've never had er ... never done any work on
 (Conv)
- b. A: has John got a new car do you think?
 B: oh well I saw that on Friday ... *in fact* yesterday we said I wonder if he's
 got a new car or .. whether he's got visitors (Conv)
- c. but he were shook up when he come in ... *in fact* he didn't know whether it were a
 child .. he says you don't know when it's dark (Conv)

Here, rather than an addition to the content of the previous idea, it introduces an additional speech act.

This often shifts to a new sub-topic within the current topic, but no particular relationship between the two ideas is signalled, nor is *in fact* emphatic here. It acts as type of presentative, along the lines of 'another thing..': it is always in initial position. It seems therefore to be becoming an additive marker introducing a new speech act. In this function it resembles certain discourse-management uses of *of course* (chapter 8.4). But whereas *of course* backgrounds its host, reminding the addressee of something rather than informing them of it, *in fact* implicates that the information in its host is new.

6.3.5 Distribution of *in fact*

The distribution across the three corpora is shown in table 6.5 and fig. 6.10.

Text types Frequency pmw	Reviews	Speeches	Conversation
Literal	4	1	1
Emphatic	28	3	20
Contrastive	87	33	7
Elaborative	89	49	72

Table 6.5. Distribution of *in fact* across three text types

In each corpus, *in fact* occurs with the ELABORATION relation more frequently than it occurs in an ANTITHESIS.

Curiously, the density of *in fact* is twice as high in the Reviews corpus as in the Speeches or Conversation corpora, although the distribution is similar. The figure for the Periodicals section of the BNC is 99 pmw, similar to Conversation and Speeches, so that the high frequency in Reviews is not

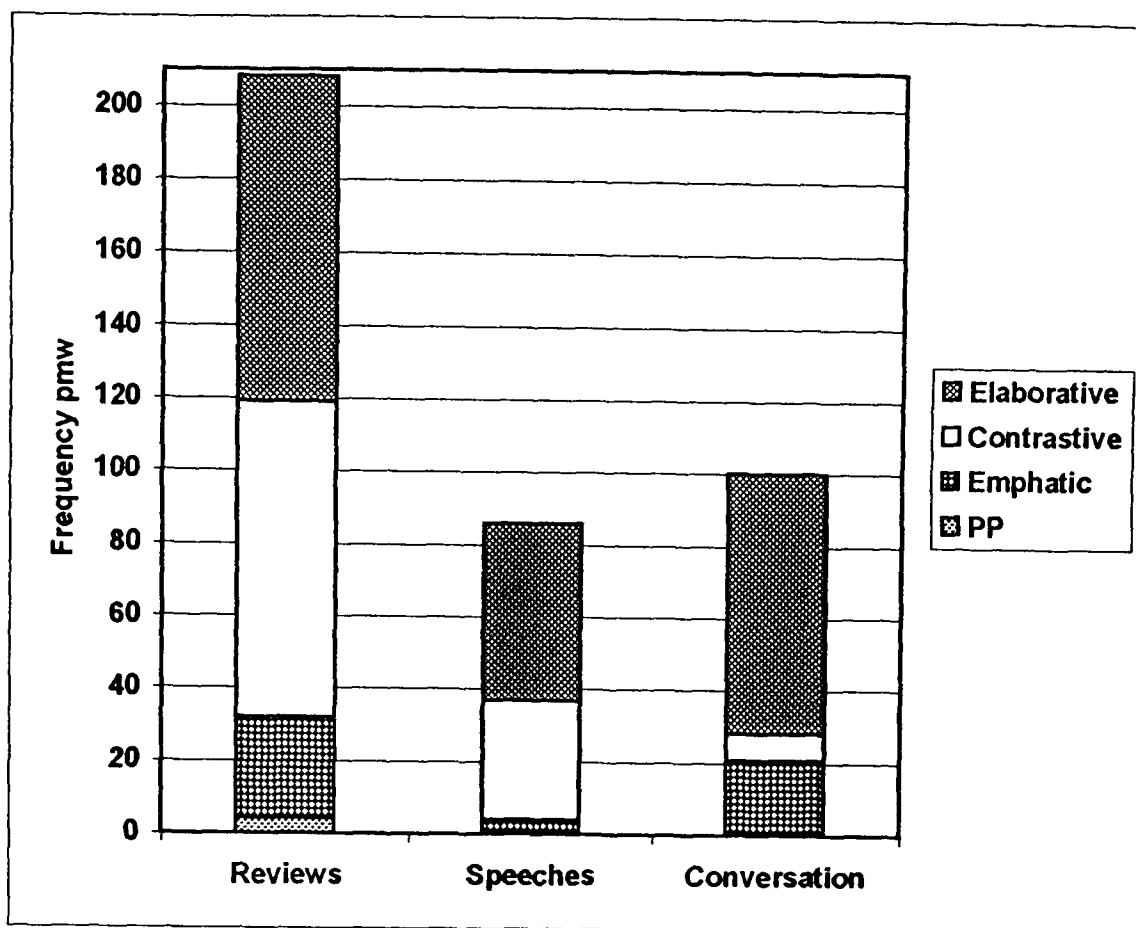


Fig. 6.10. Distribution of *in fact* across three text types

characteristic of written language in general, but perhaps of literary criticism or academic argument.

6.4 Summary

Like *after all*, *in fact* generalizes from actions and events to states. It undergoes a similar subjectification, indicating speaker commitment, then speaker judgement of the relationships between ideas. Finally, it seems to be acquiring a more abstract discourse organizational function in PDE conversation. The successive uses that have branched off move gradually further away from reference to the content of the host discourse segment, and towards more metatextual comment.

Like *after all*, *in fact* develops two connective uses that seem almost opposites: one dissonant, in which an idea that was expected or thought or claimed turns out not to have been the case at all; and one consonant, in which an idea is justified (*after all*) or further developed (*in fact*). In each case, the two uses develop from an epistemic marker via a functional split.

The next two chapters look at adverbs that have polysemies with a stronger underlying shared sense, and where the divergence has been more one of contexts.

CHAPTER 7

AT LEAST

Youre Ennemys .. schall be fayne to desire your peas and
frendeschip, *or atte lest* to bryngge hider their Merchandises
(1439, Rolls of Parliament¹)

7.1 Types of <i>at least</i>	154
7.1.1 Current definitions	154
7.1.2 Scalar <i>at least</i>	157
7.1.3 Retreat <i>at least</i>	158
7.1.4 Evaluative <i>at least</i>	161
7.2 The history of <i>at least</i>	163
7.2.1 The data	163
7.2.2 VP-adv > S-Adv	164
7.2.3 The emergence of Epistemic-Retreat <i>at least</i> : the <i>or at least</i> pattern	166
7.2.4 Towards Evaluative <i>at least</i> : the <i>if (not) p, (then) at least q</i> pattern	170
7.2.5 Summary of historical development	173
7.3 <i>At least</i> in PDE	174
7.3.1 Distribution of <i>at least</i>	174
7.3.2. Rhetorical contexts of <i>at least</i>	176
7.3.3 Layering	180
7.4 Conclusion	183

7.1 Types of *at least*.

This chapter examines the expression *at least*. It begins by distinguishing three uses of *at least*, and then looks at where these have come from historically (7.2) and at patterns of use in PDE. (7.3). It argues that *at least* provides evidence that polysemies develop gradually, via repetition in certain contexts which encourages the crystallization of implicatures salient in those contexts.

7.1.1 Current definitions

Contemporary dictionaries agree that *at least* is lexicalised, but they disagree over how many senses it expresses. In the *Collins Dictionary*, two senses are given, one emphasizing the contrastive aspect of the expression and the other emphasizing its scalar nature:

¹ Source: Middle English Dictionary, from *Rotuli Parliamentorum; ut et Petitiones et Placita in Parlamento*, 1783.

least

...

2. at least. a. if nothing else: *you should at least try*. b. at the least. **3. at the least.** Also: **at least.** at the minimum: *at the least you should earn a hundred pounds*.

The entry of the *Shorter Oxford*, under the headword *least*, glosses both the VP-Adv and the S-Adv as ‘at any rate’:

At l., at the l.: [1] qualifying an expression of amount or number: = ‘(so much or many) at any rate, if not more’; [2] hence, characterizing a statement as certainly valid, even if a wider one be not allowable; = ‘at any rate’, ‘at all events’.

The corpus-based *Cobuild* dictionary places *at least* at the top of the entry for *least* (reflecting its frequency) and states that you use *at least*:

- 1.1 to say that the number or amount mentioned is a minimum and that the actual number or amount is greater.
- 1.2 to indicate that something is the minimum which can be done, even if nothing else is done.
- 1.3 to indicate an advantage that exists in spite of the disadvantage or bad situation that was previously mentioned.
- 2. when you want to modify or correct something that you have just said.

These dictionary senses can be summarised as follows:

- ‘if nothing else’ (Collins 2a, Cobuild 1.2)
- ‘at the minimum’ (Collins 3, Cobuild 1.1, SOED [1])
- to indicate an advantage (Cobuild 1.3)
- ‘at any rate, at all events’/ to correct a previous assertion (SOED [2], Cobuild 2)

Despite much overlap, then, the dictionaries group usages of *at least* into senses in different ways.

We turn now from the lexicographic viewpoint to a linguist’s account of *at least*. Kay’s (1992) analysis of *at least* distinguishes three *at least* constructions, which he terms Scalar, Evaluative and Rhetorical Retreat. Kay’s examples are:

- (1) a. Mary received calls from *at least* three soldiers. (Scalar)
 - b. *At least*, this one’s cooked. (Evaluative)
 - c. I see her every day, *at least* when I’m in town. (Rhetorical Retreat)
- (Kay 1992:311)

Kay argues that the relatedness of the three usages amounts to non-productive, ad hoc polysemy that is not a part of the grammar (1992: 326-330). That is, no usage is predictable from the other usages;

each of the three constructions must therefore be learnt separately². However, Kay does not discuss how the usages relate to one another, or how they relate to contexts. His analysis describes the form *at least* as representing three constructions, but it does not explain how these constructions arose, or how they function within higher level constructions. Moreover, Kay does not explore the frequency of the three usages or their prosody.

Our PDE corpus data bear out Kay's division of the expression into three usages. It will be convenient to use Kay's labels, while bearing in mind that, as Kay also points out, all three are scalar.

Scalar *at least* (2) modifies a quantitative or scalar expression, such as a number, a frequency, or an intensity.

- (2) She must have been there thirty years *at least* (Conv) (Scalar)

The second pattern (3) presents an alternative to a constituent.

- (3) there is probably no substitute for a presence or *at least* an agent on the ground in the region (Sp) (Retreat)

This sense is labelled Rhetorical Retreat by Kay (1992), but to avoid confusion with the discourse-coherence meaning of 'rhetorical' and with the term as used in Rhetorical Structure Theory, this pattern will here be termed Retreat.

Finally, where *at least* modifies a host idea evaluated as beneficial by the speaker, it is Evaluative (4).

- (4) in 1904 Sicilians welcome him back to Palermo. ...| But *at least* he is not reinstated in office. (Rev) (Evaluative)

Table 7.1 shows the way these schemas are realized in the examples above.

² Kay's analysis is designed to illustrate Construction Grammar (Fillmore et al. 1988, Kay 1997), which framework holds that a grammar consists of a repertory of 'constructions', that is "conventional associations of lexical, syntactic, semantic and pragmatic information" (Kay 1992:310). Lexical items are also constructions — ones "with minimal constituent structures" (Kay 1997:125).

Type	Implicit scale	Contrasting higher point	Contrasting lower point	<i>At least</i> point
Scalar	quantity			thirty years <i>at least</i> (Conv)
Retreat	strength, domain or epistemic commitment	a presence		or <i>at least</i> an agent (Sp)
Evaluative	value		in 1904 Sicilians welcome him back to Palermo But <i>at least</i> he is not re-instated in office. (Rev)

Table 7.1. Schemas of the Scalar, Retreat and Evaluative constructions of *at least*.

The next three sub-sections describe the syntactic, semantic and rhetorical characteristics of Scalar, Retreat and Evaluative *at least*.

7.1.2 Scalar *at least*

Scalar *at least* is basic in that, as noted in section 5.1, every use involves scalarity. The examples in (5) show that in PDE Scalar *at least* occurs with different categories of host: NP, AP, AdvP, VP, PP.

- (5)
- a. *at least* one of the spouses will have to have parental responsibility (Sp) (NP)
 - b. ... despite those sides of Durrell's character which are unappealing *at least* ... (Rev) (AP)
 - c. ... in whom the ferocity of the dispossessed survives *at least* as viciously as in any layabout of the pubs. (Rev) (AdvP)
 - d. ... every eligible poor country should *at least* have embarked on the process of securing a sustainable exit ... (Sp) (VP)
 - e. We promised to keep the basic state pension | and uprate it *at least* by prices. (Sp) (PP)

The syntactic frame of *at least* is straightforward. The host must be a maximal projection, as Kay points out (1992:314), and *at least* can occur either before or (in around 5% of cases) after its host. Other types of *at least* are rare in final position. Prosodically, Scalar *at least* generally forms a single intonational phrase with its host³.

³ Kay claims (1992:313) that when it occurs to the right of its focus it takes parenthetical intonation. But this seems not to be the case. For example, in 'I hope that some of you *at least* have had time to look at the consultation paper' (Sp), it can form a single intonational phrase with 'some of you'. Prosodic markings of *at least* constructions in the London-Lund corpus support the claim that final-position Scalar *at least* does not always take parenthetical intonation. See also Allerton and Cruttenden (1974).

Scalar *at least* can modify any constituent whose meaning can be construed as scalar. It indicates that the host represents a minimum (on whichever scale is appropriate to the meaning of the host) and implies ‘possibly more’. In other words, it is used to cancel the default implicature that when a speaker says ‘q’ she means ‘not more than q’.

In terms of rhetorical structure, the host of Scalar *at least* is not an independent discourse unit. So although *at least* focuses its host, the resultant constituent can assume any rhetorical role within the discourse unit of which it is a part. For example, in (5c,d,e) above, the host of *at least* is part of rhematic material, whereas in examples (5a,b) it is not.

7.1.3 Retreat at least

Retreat *at least* highlights the contrast between a point on an implicit scale, marked by *at least*, and a higher point on that scale. In fact, three configurations of Retreat *at least* are distinguishable. They will be called Alternative, Restrictive and Epistemic-Retreat and are exemplified in (6).

- (6)
- a. if you make an appointment you should jolly well go and see it .. or *at least* pick the phone up and say erm you know .. (Conv) (Alternative)
 - b. ... these [operas] ... are largely unrevivable, | *at least* in the present climate of musical taste. (Rev) (Restrictive)
 - c. Raymond Seitz was the most popular and successful US ambassador to the Court of St James’s since David Bruce more than 30 years ago. | Or *at least* that represented the view on this side of the Atlantic. (Rev) (Epistemic-Retreat)

All three involve a retreat from a stronger assertion to a weaker one; what differs is the relation between the two assertions, described below. Each configuration also has a typical syntactic realization.

In the first configuration, *at least* modifies an element that is put forward as an alternative to another element in the discourse. Syntactically, the Alternative Retreat consists of two maximal projections of the same type, one of which is host to *at least*. Alternative *at least* and its host almost always form a single intonational phrase.

Examples are:

- (7) a. But the overall perspective and objectives of each of the major political parties were set, | if not in stone, | ***at least in conventions which operated like steel contours*** (Sp, Field)
- b. A **-these meant to be archived are they?**
B -well or ***at least sorted*** (Conv)
- c. Some of the designers of “Europe” would positively prefer us not to know, | *or at least would not care what we thought about it if we did.* (Rev)

Semantically, Alternative *at least* operates on a scale of strength or intensity. Two distinct ideas are mentioned, one of which, the *at least* point, is less strong than the other. There is an element of concession in the construction with *if not*: the stronger claim may not be the case, but the weaker claim certainly holds. *Or*, when it is present, is usually inclusive: both ideas could hold. For example, (7b) allows that the papers might need to be both sorted and archived. The pattern is therefore:

{(stronger idea) + (*at least* weaker idea)}.

While the semantics always involves stronger and weaker ideas, different rhetorical structures correspond to the different constructions. The *or at least* host (7b,c) is a rhetorical Satellite, while in the *if not p, at least q* pattern (7a), *q* is a rhetorical Nucleus (the Nucleus in each case is bolded). The typical rhetorical structure patterns are:

<i>or at least:</i>	N, or <i>at least</i> S
<i>if not p, at least q:</i>	if not S, <i>at least</i> N / <i>at least</i> N, if not S

The host of Alternative *at least* is thus an independent discourse unit.

The second ‘retreat’ pattern has *at least* modifying an element which expresses a restriction on the applicability or the scope of an idea. Examples are:

- (8) a. In a short story Jorge Luis Borges resolved the mystery | — ***at least imaginatively*** — | by suggesting that, throughout his life, Shakespeare concealed a terrible secret. (Rev)
- b. we have not, | ***in Britain at least***, | valued it at its true worth in recent times. (Sp)

Restrictive *at least* and its host tend to form a single intonational phrase. Syntactically, the host of Restrictive *at least* is typically an adjunct or a parenthetical or a fragment, usually an adverb or adverbial phrase. The adjunct occasionally follows its clausal host, but it also occurs initially and very often medially.

Restrictive *at least* implies a scale of applicability: an idea is expressed and is then modulated

by the *at least* host, which defines the degree or the domain(s) of applicability of that idea, cancelling the unlimited, default reading. The semantic pattern can be broadly characterized as:

{(assertion *p*) + (*at least* for some D [assertion *p*])}

where D is the relevant domain (degree, time, manner, place, participants, etc.). Like Alternative *at least*, a stronger idea is replaced by a weaker one; but whereas Alternative *at least* deals with two distinct ideas, here the constraint is on the same idea. For example, (9a) implicates ‘only in America’ (by Horn’s Q-Principle ‘say as much as you can’). *At least* in (9b) cancels that implicature.

- (9) a. Betty Friedan discovered early on that in America, | when a man’s wife dies, | unless he remarries | he is more likely to follow suit ...
 b. Betty Friedan discovered early on that, | *at least* in America, | when a man’s wife dies, | unless he remarries | he is more likely to follow suit ... (Rev)

The host of Restrictive *at least* is an independent discourse unit. Indeed, *at least* functions to create that additional discourse unit, as can be seen in the following contrast with Scalar *at least*:

- (10) Foresight aims to help all those involved to reach agreement on *at least* some important areas (Sp) (Scalar)
 Foresight aims to help all those involved to reach agreement, | *at least* on some important areas. (Restrictive)

Rhetorically, Restrictive *at least* introduces a Satellite, while the unrestricted idea has Nucleus status .

The typical rhetorical structure is ‘N, *at least* S’ or ‘*at least* S, N’. As in the case of Alternative *at least*, many Restrictive patterns involve parentheticals, which typically encode rhetorical Satellites.

In its third configuration, Retreat *at least* introduces evidence for the previous idea. This amounts to a comment on the epistemic status of that idea. Some examples are:

- (11) a. ... you do feel the flush of another dimension, | *At least*, | you might with the help of a few cans of Boddies. (Rev)
 b. Boyd ... appears not to know of my own friendship with Tate; | *at least* | he fails to mention one of the great vignettes of the New York art scene. (Rev)
 c. A -it’s .. it’s not in this week then .. that one?
 B -no it doesn’t seem to be. | *at least* | I, I didn’t see it. (Conv)

As Epistemic-Retreat *at least* is connective, linking one clausal constituent to the previous one, it occurs most commonly in initial position. It is the only type of *at least* to usually form a separate intonational phrase and to constitute an independent discourse segment.

It introduces either a modality or a warrant for the previous assertion. Broadly, the frame is:

{(assertion *p*) + (*at least* reduced speaker commitment to assertion *p*)}

Epistemic-Retreat *at least* is similar to Alternative and Restrictive *at least* in that a strong statement is restated as a weaker one. In Epistemic-Retreat *at least* it is speaker commitment that is weaker. The *at least* unit retracts the default full commitment of the first discourse unit and replaces it with a weakened commitment.

Kay comments that his example of what is here called Epistemic-Retreat “seems to have a metalinguistic flavour” (1992:319), while examples of (Restrictive) *at least* “add a qualification to the content of the proposition already asserted and do not address the issue of the speaker’s warrant for its assertion”. The difference is one of discourse plane: Epistemic-Retreat *at least* applies at the attitudinal-epistemic level, while Restrictive and Alternative *at least* apply at the content level. Omission of Epistemic-Retreat *at least* can result in a contradiction or a non sequitur, whereas it is usually possible to omit Restrictive or Alternative *at least* and be left with a well-formed sentence and a coherent discourse.

Rhetorically, the *at least* host is a Satellite to the previous idea. The Epistemic-Retreat *at least* cannot be interpreted without knowledge of the previous assertion. It is therefore a discourse connective: it encodes the coherence relation between the two discourse segments: N. *At least*, S.

7.1.4 Evaluative *at least*

Evaluative *at least* indicates the speaker’s favourable evaluation of some idea. Examples are:

- (12) a. It is a less common experience for me to speak after the second course. | ... *At least* speeches do not start at the beginning of a dinner, | as is I believe the custom in parts of Scandinavia. (Sp)
 b. -*at least* I haven’t got school tomorrow .. | I’ll have a lie in tomorrow (Conv)
 c. ..he promptly settled upon persuading a group of women to castrate Rasputin. | *At least* it cannot be said that Russians lack resource or imagination (Rev)

Syntactically, *at least* typically modifies a clause or fragment and it occurs in the usual S-Adv slots (see ch. 1.4). Its syntactic behaviour differs from that of Epistemic-Retreat *at least* in that it is not parenthetical but is integrated into the clause. Evaluative *at least* is also prosodically integrated with its host.

Evaluative *at least* implies a scale of benefit, on which the content of the host contrasts favourably with some lower point. The less favourable state of affairs need not be expressed. This means that Evaluative *at least* can initiate a discourse without causing incoherence. For example (13) suggests there was a risk the speaker might not have got her card back.

- (13) ... it came out erm .. incorrect pin number .. please remove your card .. I thought (***)
at least I'm getting it back you know so I whipped it out .. I said to mum ..
 I'm gonna have to pay for this with a cheque .. I said ... (Conv)

It does not signal a rhetorical relation. Unlike the Epistemic-Retreat pattern, therefore, Evaluative *at least* is not connective. When the contrasting, negatively valued idea is explicit, the *at least* host is rhetorically nuclear and the adjacent contrasting idea is its satellite: [S]. *At least* N.

The occurrence of these distinct patterns of use shows that *at least* is polysemous (v. chapters 1.4 and 2.4). Its core sense is scalar. It will be argued below that the polysemies have arisen from conversational implicatures generated in particular contexts.

Table 7.2 summarizes the characteristics of the different *at least* patterns.

Type		Discourse plane	Sense	Typical rhetorical structure
Scalar		content	Scale: quantitative 'This much and maybe more' (cancels implicature 'this much and no more')	Host is not independent discourse segment Host may be Nuclear or not
Retreat	-Alternative	content	Scale: strength of concept/degree of intensity Presents a weaker alternative to the previous assertion	Host can be discourse segment or <i>at least</i> : host is Satellite if not p, <i>at least</i> q: host is Nuclear
	-Restrictive	content	Scale: domain of applicability (time, place, manner etc) Introduces a restriction on the applicability of the proposition just expressed	Host is usually discourse segment Host is Satellite
	-Epistemic-Retreat	attitudinal-epistemic	Scale: epistemic commitment Reduces commitment to an idea; states evidence for it	<i>At least</i> is discourse segment (S) Host is discourse segment (N) Host is S to previous segment N <i>At least</i> is discourse connective
Evaluative		attitudinal-evaluative	Scale: benefit Signals a positive evaluation of the proposition expressed compared with the previous idea or with a presupposed idea	Host is discourse segment Host is N when contrasting idea is expressed.

Table 7.2. Summary of PDE *at least* patterns

7.2 The history of *at least*

7.2.1. The data

The historical data used are described by table 7.3.

Period	No. of occurrences in full-text sources						No. in non-full-text sources				
	Hist	HelME	HelEME	Lamp	Ng	Ceecs	Total	Dicys	Memem	Other	Total
1300-1399	28	3					31	1			1
1400-1449	5	3					8	2		1	3
1450-1499	14	2					18		2		2
1500-1549	7		7				18	6	27	1	34
1550-1599	23		21				57	3	53	1	57
1600-1649	14		21	21			83	2	30		32
1650-1699	99		26	153	68		355	7	32	2	41
1700-1749	78			99			177				
1750-1799	186						186				
1800-1849	104						104				
1850-1899	131						131				

Table 7.3. Historical data on *at least*

Fig. 7.1 shows the evolution of the frequency of *at least* by type from the C15th to the C19th.

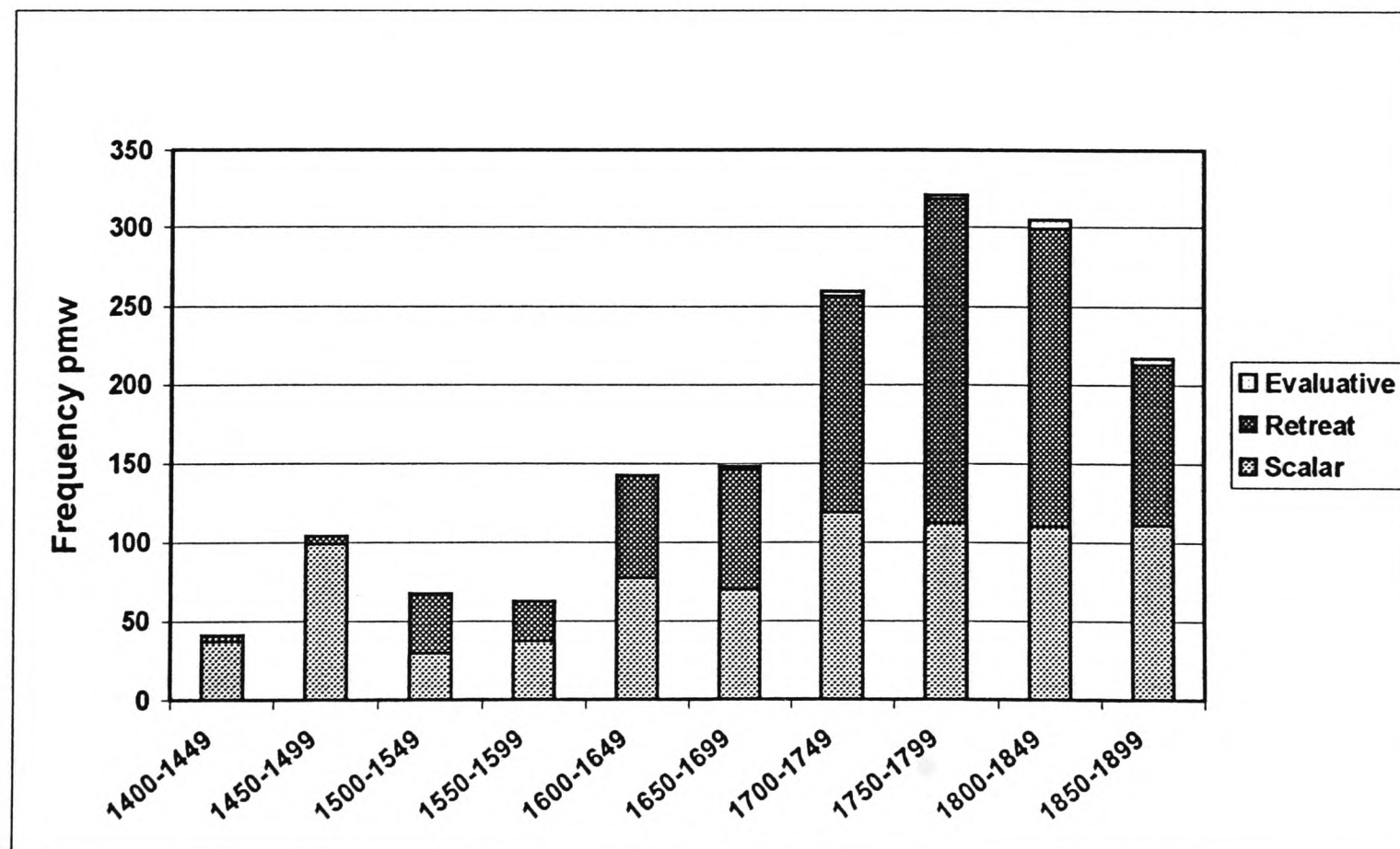


Fig. 7.1. Frequencies of *at least*, C15th-C19th

The increase in frequency, peaking in the late 18th/early 19th centuries, is very largely due to Retreat *at least*. The frequency of Scalar *at least* has remained fairly stable. Evaluative *at least* has very low frequency. This no doubt reflects the written language bias of the corpora: in PDE Evaluative *at least* is still low frequency in written language, but high in spoken (see section 7.3).

Scalar *at least* dates back to Old English, while the Epistemic-Retreat and Evaluative senses are more recent (fig. 7.2).

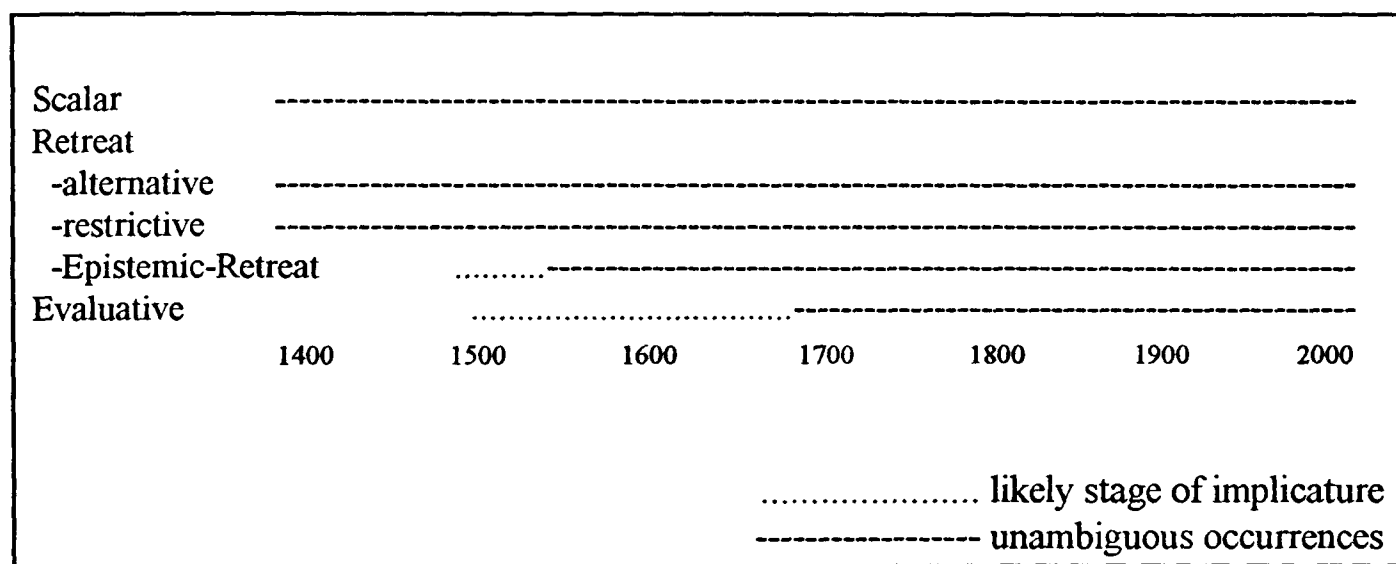


Fig. 7.2. Emergence of types of *at least*

7.2.2 VP-Adv > SAdv

All uses of *at least* are scalar in some way and involve a contrast between two incompatible ideas. *At least* has stayed remarkably stable: the thirteenth-century use illustrated in (14) is still current.⁴

- (14) a. eihte reisuns *et te leste* (a1225, OED, MED)
 b. schriue hire euche wike eanes *ed te leaste* ['each week once'] (c1230, MED)

Although *at least* has now largely pushed out *at the least*, the latter is still found in PDE.⁵ Equivalents of present-day English (PDE) *at least* have been, and in some dialects still are, *at the least*, *at the least wise*, *leastwise*, *at the least mark*, *at the least way*, *by the least (way)*, *leastways*, and variants of these. Those with *wise* and *way* seem to occur more often with non-numeric hosts. However, both are

⁴ In Middle English and Early Modern English, *atte leste*, etc. is sometimes conflated with *atte laste*; Thomas (1587) uses *at the least* and *at the least way* to translate Lat. *tandem*, 'when all is done'.

⁵ *Least* and *lest* are of the same origin but *lest* will not be discussed.

found with numbers. In Chaucer, for instance, the use or omission of *way* is clearly governed by rhyming requirements. *At the least* is often a filler brought in to provide rhyme for expressions such as *biheeste, feeste, geeste*.

- (15) And but I have hir mercy and hir grace,
That I may seen hire atte leeste weye,
I nam but deed; ther nis namoore to seye. (c.1380s, Chaucer, Knight's Tale)

The earliest uses of *at the least* with a numerical host in the corpus are in final position as in (16):

- (16) a. An hundred *at þe leste* (c1300 MED)
b. trie *at the leste* (c1300, Helsinki)

But from the turn of the C15th, *at the least* appears both in final position separated from the number itself (17) and in initial position (18):

- (17) a. To make odour parfyte, .. foure þinges nedep *atte leste* (a1398, MED)
b. the seid elleccion stond by xv of hem *atte leste* (1444, MED)
- (18) a. þis wounde schal be holden open *at þe leeste* fourty daies (a1400, MED)
b. *at þe leeste* twies in þe day (c1475 MED)

At around the same time, *at the least* occurs in numerical contexts, but raised to precede a PP (19a,b) or to the medial VP-Adv position (19c,d) or to initial position (19e).

- (19) a. shall *at þe lest* before þe ende of euery terme (1425, Hist)
b. and loke þat þou holde þe wounde open, *at þe leeste weye*, vnto þe fourti daies ende (c1475 MED)
c. he had *at the leest* in his company .v. thousand feighters (1500, CMEPV)
d. vnles he would *at the least* geue them .iii. li., | they woulde smite of his arme from the body (1567, Hist)
e. *at the lest weye* lete him purveye x li., for owyth be my reknyng at Myhelmesse last passed, be-syde youre faderes dette, xvij li. xiiij s. viij d. (1475, Hist)

This leftwards movement is common to other English adverbs (cf. Nevalainen 1991). In the same period, Scalar *at least* occurs as a VP-Adv and manner S-Adv (20).

- (20) The flode was so felle, with fallyng of Rayn, Hit was like, *by the lest*, as oure lorde wold With water haue wastid all the world efte (c1400 OED)

In the above examples, *at least* implies 'definitely', by contrast with a presupposed higher quantity that is only possible or desirable.

Few unambiguous occurrences of Restrictive *at least* appear in the corpora before the end of

the C16th, but there is evidence that it arose much earlier (21).

- (21) a. Hote and moyste membres beep saide, nameliche, þe blood (*at þe leste* materialy), þe spirit, and þe flesche (1425 MED)
 b. Men muste haue haunt and vse of þo fleischly deedis *at þe leest* in matrimonye for þe peplis multipli yng (1443 MED)

The entry in the IVth edition of Johnson's dictionary claims that *at least* 'implies doubt', citing an example from Milton of 'if *at least*'. 'Implies doubt' is a sense quite contrary to that of 'certainly' or 'at any rate' seen above⁶. There are indeed many examples of *at least* in conditional contexts, but rather than indicate doubt, the *at least* seems to imply that the condition clause is an elaboration of the previous idea; i.e. it demotes the condition clause to the status of modifier. These occurrences can therefore be seen as Restrictive (22). It is unlikely that *at least* could occur if the *if*-clause were in initial position (cf. chapter 2.3.3).

- (22) a. It is therefore our parts, if *at least* we be desirous of the health and salvation of our owne selues, most earnestly to pray vnto our heauenly Father .. (1563, Hist)
 b. Truly I could wish, if *at least* I might be so bold to wish, in a thing beyond the reach of my capacity ... (1595, Hist)

From Middle English onwards, two regular, contrastive *at least* patterns are noticeable. The first is *[either] p or [else] at least q* (section 7.2.3). The second is *if/since [not] p, [then] at least q* (section 7.2.4).

7.2.3 The emergence of Epistemic-Retreat at least: the *or at least* pattern

In PDE, *or* and *at least* collocate so readily that they can be said to constitute an independent construction. The collocation has a solid history: its strength as reflected in these corpora is shown in table 7.4. The co-occurrence with *or* dates back at least to the 14th century (23).

- (23) a. til þat foure daies ben passid, | *or ellis þre at þe leeste wey* (1392, MED)
 b. This litel cercle shal be perced ful of smale holes .. | in 360 holes yif it be possible | *or in .180. or in 90 atte leste* (c1392, Helsinki)

⁶ To claim that the *at least* implies doubt may be to confuse the sense of *at least* with the frame in which it occurs (cf. Sweetser 1990: 14).

Period	<i>or at least</i> as % of all <i>at least</i>	density of <i>at least</i> (pmw)
1400-1499	16% (5/31)	73
1500-1599	34% (56/166)	65
1600-1699	26% (135/511)	145
1700-1799	18% (65/363)	290
1800-1900	14% (33/235)	261

Table 7.4. Proportion of *or at least* in total *at least* in historical data.

Occurrence of *at least* with *or* and in initial position coalesces into the *or at least* pattern. The examples in (24) express alternatives: the conjuncts linked by *or at least* are of the same constituency and are various types of complement or clauses. (24) exemplifies the pattern of repetition (e.g. *synnes ... synnes*, etc) that often occurs in Alternatives.

- (24) a. .. that he confesse nakedly and openly al his synnes ..| *or at leste* the synnes that he remembrey may .. (c1450 MED)
 b. .. to drinke with colde water, *or at the leest* colde water yf it be somer, and in wynter warme (1526, MEMEM)
 c. .. putte the chylde in to the water, | *or at the leest* caste water on the chylde (1528, Hist)
 d. .. than the verbe ... muste be negatyue | *or at the leste* the verbes muste be of contrarye sence (1530, EMED, Palsgrave)

The presence of *at least* in the expression of alternatives adds scalarity. The conjuncts are asymmetrical: *p or at least q* is not the same as *q or at least p*, because the *at least* host is necessarily lower on the scale than the other conjunct (cf. *or even*).⁷ The relation between the *at least* host and the other conjunct is best seen as ELABORATION, because the host provides more detail or expresses a refinement of the previous idea. The *at least* host comes to act as a modifier of the other conjunct, to tone down or hedge it. The pattern involves two ideas, the second of which is usually more likely than the first. It therefore highlights the implicature of certainty: '*p*, well, surely *q*'.

The *at least* host therefore starts to become a Satellite. The use in this construction of *either*, which implies equal alternatives, dwindles, as the asymmetry between the two conjuncts increases.

⁷ Sweetser (1990: 93-100) describes uses of *or* across the content, epistemic and speech-act domains and shows that on all three levels there are both symmetric and asymmetric uses.

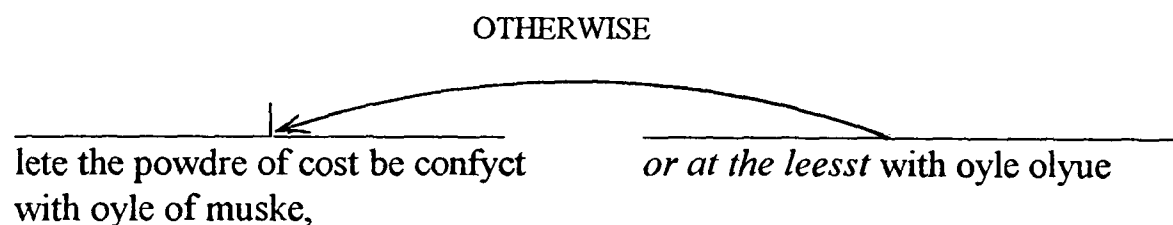
From the 16th to the 18th centuries, the construction with *either* — ‘*either p or at least q*’ as in (25) — accounts for between 10% and 15% of the occurrences of *or at least*.

- (25) a. that they may thereby *either* bee brought to goodnesse, | *or at the least* that GOD and the common wealth may be lesse hurt and offended (1547, Hist)
 b. *eyther* they be ryche, | *or at the leaste wyse* be verye well pleased wyth that little that they haue (1557, Hist)

But occurrences with *either* are extremely rare in PDE. The *or at least* pattern now expresses less an alternative than a correction or hedge of the previous idea.

All the corpus occurrences from the period C15th to early C16th are in irrealis⁸ contexts so that either idea is a possible outcome: *or at least* signals a real-world alternative, usually an OTHERWISE relation. In every case, the idea in the host of *at least* is not asserted, as evidenced by a modal or an imperative etc. (26).

- (26) lete the powdre of cost be confyct with oyle of muske, | *or at leest* with oyle olyue (1526, MEMEM)



But from the early C16th, *or at least* starts to appear in realis contexts, and can be seen as an unambiguous epistemic marker. In example (27), the host clause *we haue no worde to countreuayle it* is almost certainly asserted.

- (27) .. this worde si... semeth to signyfye nothyng with vs | *or at the leste* we haue no worde to countreuayle it [‘we have no word of equivalent value to it’] (1530, EMED, Palsgrave)

It can be interpreted as EVIDENCE for the previous segment, rather than an ELABORATION of it.

Whereas in the previous examples, *p or at least q* implicates *if q not p* (i.e. an exclusive *or* interpretation), no such implicature holds for (27). That is, in (26) *with oyle of muske, or at leest with oyle olyue* implicates ‘if *with oyle of muske*, then not *with oyle olyue*’, whereas in (27) the

⁸ The term ‘irrealis’ here refers to an unrealized situation, i.e. to an event or situation that the speaker / writer perceives as unreal, as not having taken place. The term ‘irrealis’ will not be used here as a grammatical term to refer to a verbal form (cf Palmer 1986: 26).

interpretation is not ‘if we have no word for it, then it doesn’t mean nothing’, but rather ‘if my evidence is only that we have no word for it, then I cannot claim that it means nothing to us’. This is not a real-world, content relation, but an epistemic one. (27) can be seen as a bridging context between Alternative and Epistemic-Retreat *at least*.

There seems, then, to be a split in the *or at least* construction: its use to indicate a real alternative continues, but a new, epistemic use branches off. Alternative *at least* indicates a lower limit. It can often be glossed as ‘certainly’, because it indicates that although the first idea may not be realized, the *at least* host idea is the minimum that can be expected, thus very likely to be realized. From there, the extension of use to the attitudinal-epistemic plane is a small step. The types of relation expressed by *at least* on this plane are broadly evidential (28) (fig. 7.3).⁹

- (28) a. my Lady Castlemaine ... supposed to have miscarried, | but for certain is as great in favour as heretofore, | *at least* Mrs Sarah at my Lord’s, who hears all from their own family, do say so. (1666, Hist)
- b. ... Enforced him to take the Chaire to which he was very unwilling *or at least* Immediatly soe appeared (1675, Newdigate)

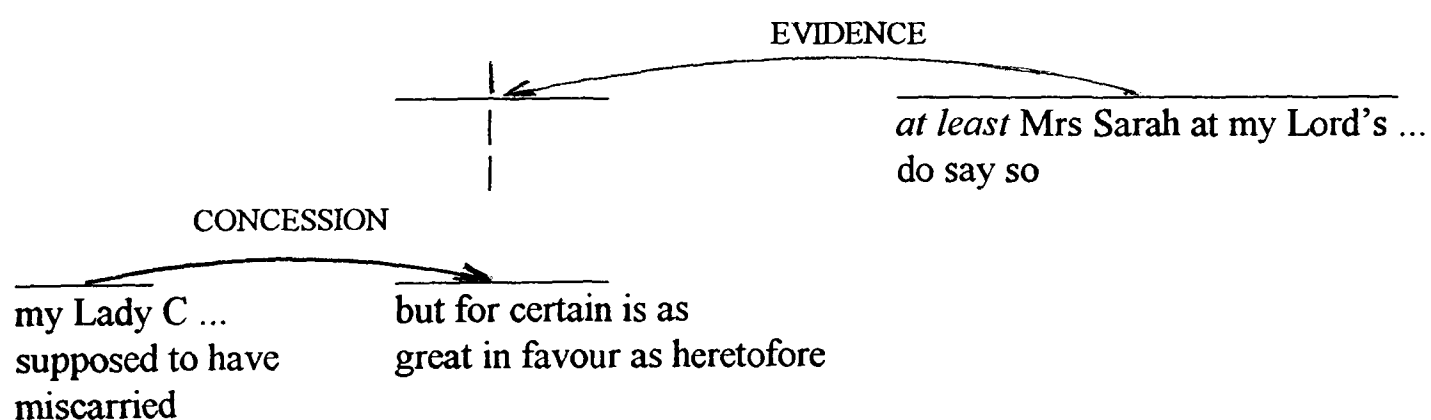


Fig. 7.3. RST schema for example (28 a)

The implicature in (28b) is not ‘if he appeared unwilling, he was not unwilling’, but ‘if he was *apparently* unwilling, he was not *definitely* unwilling’. (28) may be regarded as switch contexts, as they are clearly epistemic.

The same epistemic retreat sense results from Restrictive *at least* where the domain is a subjective perception (29).

⁹ It may be argued that EVIDENCE is inadequate to describe examples (28); that they display differences that call for a finer-grained analysis of their rhetorical relations (v. ch. 2.3.5) and of their discourse plane (v. ch. 1.3). The purpose of this analysis, however, is to distinguish ‘autonomous’ uses of *at least*, i.e. stable patterns that can generate unambiguous instances.

- (29) a. lett him commend you for his successor to the M^r and fellowes, | which | being an ordinary curtisie in all Colleges, | will add something to their unkindnes which shall withstand it, | *at least* in your Lord's opinion (1625, Ceece)
- b. the Bath-water will have relation to, | or comprehend in its largest extent, | eight distinct Substances, | *at least* according to my Observations, | which will be the business, God willing, of the first Book of my Treatise. (1676, Lampeter Science)

Epistemic-Retreat *at least* is therefore an Alternative or Restrictive on the attitudinal-epistemic plane; it retains the same 'minimum' scalar sense, but in the domain of speaker commitment. From the turn of the C18th, *at least* occurs to indicate a lowered commitment to the previous idea by claiming that there is merely some evidence for it. The sense of [*or*] *at least*, then, shifts from an expression of a real alternative to a Quality hedge (30).

- (30) a. William the Conqueror proceeded much further; ... directing all Pleadings to be in that Language, and endeavouring to make it universal in the Kingdom. This, *at least*, is the Opinion generally received. (1710, Hist)
- b. Lady Ladd has been very handsome, but is now, I think, quite ugly — *at least* she has a sort of face I like not (1780, Hist)
- c. We dine to-day at Goodnestone, to meet my aunt Fielding from Margate, and a Mr Clayton, her professed admirer; *at least* so I imagine. (1795, Hist)

In (30), *at least* is fully discourse connective.

7.2.4 Towards Evaluative *at least*: the *if (not) p, (then) at least q* pattern

From around the middle of the 16th century, *at least* often occurs in the context of a strong explicit contrast. The most obvious exponent of this contrast is the pattern *if (not) p, (then) at least q*, where *p* and *q* are incompatible ideas. It accounts for around 5-9% of occurrences of *at least* (table 7.5). This construction can be seen as the typical exponent of a use of *at least* that emphasizes the contrast between the *at least* point and the lower bound.

Period	<i>if (not) p, at least q</i> pattern as % of all <i>at least</i> occurrences
1500-1599	9% (15/166)
1600-1699	5% (24/511)
1700-1799	8% (29/363)
1800-1899	6% (14/235)

Table 7.5. Proportion of *if (not) p, (then) at least q* patterns in total *at least* in the corpora.

The pattern is equivalent to the early *or at least* pattern in that *p or q* can often be paraphrased as *if not p, q* or *q, if not p*. Indeed, the two work together in example (31).

- (31) Pe rote of reedis .. is seid to drawen out boonys or yren or oþir sicke wiþouten ony greuauce, | *or, at þe leeste*, | *if* it may *not* fully be drawen oute, | it remollip & largip þe place | þat þei moun be drawen out wiþouten hurtyng. (c1475, MED)

[‘...is said to draw out bones or iron or other such without any pain, or at least if it cannot be completely drawn out, it softens and enlarges the place ..]

Some C15th and C16th examples of this pattern are:

- (32) a. *Though* thou be now a prisoner, *Keþ atte leste* thyn herte to me (a1425, MED)
 b. *yf* thou wille hym *no* good | *yet at the leste* wille hum no harme ner encombraunce (c1475, Hist)
 c. *Thoughe* he be *nat* your frende | *yet* sythe he cometh to you | *at the leest* you ought to welcome him (c1530, Hist)
 d. *if* quint essence can *not* be made, *at the leest* the second or the third, or as many as may be (1559, MEMEM)
 e. *seing* thou hast vsed so small care to succour my lyfe, *at the leest* discharge the office of a frende in reuengyng of my death (1569, Hist)

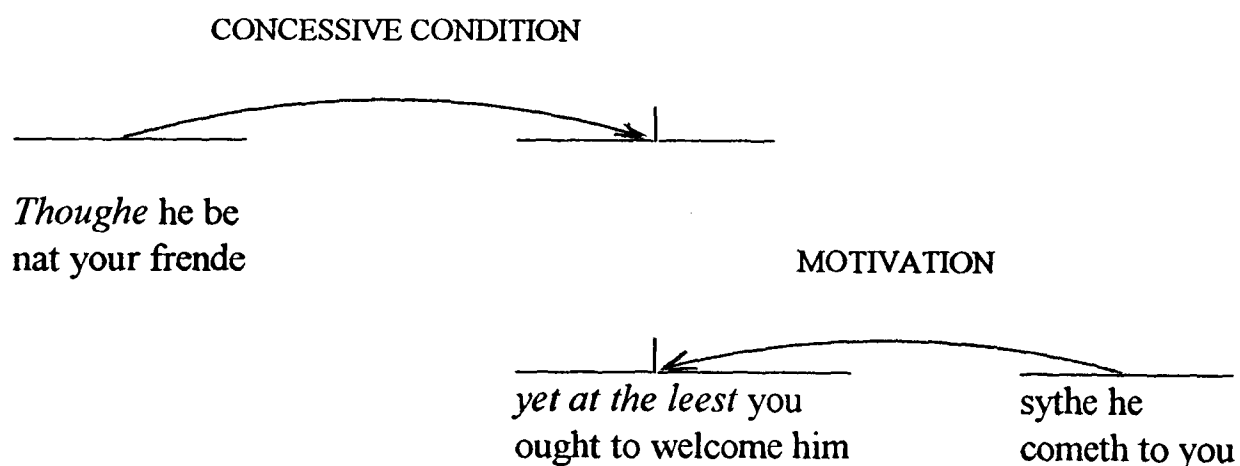


Fig. 7.4. Rhetorical schema for (32c).

The pattern with *if* frequently co-occurs with one or more contrastive conjunctions, such as *yet* (32b,c), and in the presence of a lexical contrast, such as *good vs harme* or *lyfe vs death* (32b,e). The *at least* host is Nuclear. It therefore differs in information structure from the *or at least* pattern, in which the *at least* host was less salient than the other conjunct.

Early contrastive contexts in the corpora are mostly *if*-constructions. Later, *at least* occurs with other contrastive structures (33).

- (33) a. Pensions .. leave but small signs of advantage behind them: | *whereas* ... the method now propounded will *at least* leave the persons ... the advantage of an industrious Education and Profession (1680, Hist)
 b. *Whatever* I have failed in, I shall *at least* let you see with what Obedience I am, Sir, your most humble Servant (1690, Hist)

(33a) shows an ANTITHESIS relation and (33b) a CONCESSIVE-CONDITION relation. Both of these are presentational relations and both involve increasing the reader/hearer's positive regard for the Nucleus (chapter 2.3, see also Appendix A). Examples (32) and (33) illustrate the occurrence of *at least* in the Nucleus of emergent attitudinal schemas, the Nucleus being the idea that the writer favours.

In the examples seen so far, the *at least* host is irrealis (it has not yet taken place and is therefore an unrealized situation) and it is placed on a scale of desirability: the *at least* host would clearly be a 'good thing'. So when *at least* begins to occur with realis hosts (34), it is associated with a favourable evaluation.

- (34) a. I am, if not entirely convinced, *at least* silenced (1710, Hist, Berkeley)
 b. it is certain we here advance a very intelligible proposition *at least*, if not a true one (1750, Hist)
 c. the visits of bees, if not indispensable, are *at least* highly beneficial (1859, Hist)

Examples (34) are interpretable as CONCESSION relations rather than CONDITIONS.¹⁰ The emphasis on the contrast, together with the extension to realis contexts, gives rise to Evaluative *at least*. Pejorative contexts such as (35) are no longer possible in Modern English.

- (35) And *at the leest* he robbeth them or lyeth by them (c1470, CMEPV)

The emergence of Evaluative *at least*, then, occurs in parallel with the shifts in *though* and *if*, which follow the cross-linguistic tendency for concessives to develop out of conditionals (König 1985, Kortmann 1997: 197ff, König and Siemund 2000). As König points out, "given that conditionals ... do not entail their subordinate clauses and given that concessive sentences do, the former can only change to the latter if they lose their hypothetical character" (1985: 273). Examples (32) and (34) above illustrate this loss of hypothetical character and the development of *if* from conditional to concessive conditional and then to concessive. It is therefore plausible, given the association of *at least*

¹⁰ CONDITION relations entail neither the protasis nor the apodosis, while CONCESSION relations entail both. CONCESSIVE-CONDITION relations fall between the two; they often entail just the apodosis (König 1985: 265).

with the *if not p, then q* construction, that the development of the Evaluative sense is partly driven by the conditional > concessive movement. Positive regard characterizes the latter, not the former.¹¹

Evaluative *at least*, then, may originate in patterns such as ‘*if p, at least q*’, ‘*p, but at least q*’ and ‘*[al]though p, at least q*’. It is characterised by the contrast between two events on a scale of ‘benefit’, and the fact that the *at least* host is Nuclear (fig 7.4).

Scalar and Retreat *at least* imply some uncertainty or vagueness: in Scalar *at least* it is the exact degree or quantity involved that is unknown; Alternative and Restrictive *at least* impose scales of quality and applicability; in Epistemic-Retreat *at least* it is not known how certain the event is. But *at least* is unambiguously Evaluative where the event is certain or realized. It contrasts its host with a less favourable state of affairs while suggesting that an even better situation might have been the case. Since in PDE neither of these contrasting points on the scale need be mentioned, Evaluative *at least* is perhaps losing its scalar sense and becoming simply a marker of speaker/writer positive evaluation (as in example 12b; see also section 7.3).

7.2.5 Summary of historical development

Epistemic-Retreat *at least* and Evaluative *at least* have both developed in the Modern period. Epistemic-Retreat developed from earlier Alternative and Restrictive Retreat uses. It implies ‘certainly’, ‘definitely’. When an Alternative construction involves an alternative modality, or when a Restrictive construction involves an epistemic limitation, *at least* no longer applies on the content plane, but emerges on to the attitudinal-epistemic discourse plane. That is, a comment is made on the epistemic status of the previous idea. In rhetorical terms, the *at least* host expands from expressing an OTHERWISE or an ELABORATION Satellite to expressing an EVIDENCE Satellite.

Evaluative *at least* developed from earlier Scalar and Alternative uses at least partly via constructions involving contrast with an unfavourable situation. When an Alternative construction

¹¹ There are interesting parallels between *at least* and *anyway*. Both can be glossed as ‘at any rate’ in many contexts. Both occur in concessive-conditionals, in concessives, as restrictives (Retreat-Restrictive) and as epistemic markers. For analysis and discussion of *anyway*, see König 1985: 278-279, Altenberg 1986, Lenk 1995, Ferrara 1997, Tabor and Traugott 1998: 253-260, Takahara 1998.

involves a scale of desirability, expressed in part by foregrounding the *at least* host, *at least* acquires an approbative sense. It no longer applies on the content plane, but on the attitudinal-evaluative plane to comment on the perceived value of the idea. It presupposes an alternative, less favourable state of affairs, but it is not connective.

7.3. *At least* in PDE

7.3.1. *Distribution of at least*

At least is among the commonest PDE adverbs. Judging by BNC data, it has a frequency per million words (pmw) of about 246.¹² The frequency ranges from 125 pmw in spoken language classified as 'educational', to 335 pmw in written language classified as natural and pure sciences. In the Reviews, Speeches and Conversation corpora, densities of *at least* are as follows: Reviews 238 pmw, Speeches 124 pmw, Conversation 142 pmw. 1,032 occurrences of *at least* were examined, of which 891 were clear enough for analysis. (252 from the Reviews corpus, 172 from the Speeches corpus and 467 from the Conversation corpus). Tables 7.6, 7.7 and 7.8 show the distribution by type and position for each of the three corpora.

Type	Position	Initial	Medial	Final	Total
Scalar		78	1	4	83 (53%)
Retreat					
Alternative		8	5	-	13 (8%)
Restrictive		29	-	15	44 (28%)
Epistemic-Retreat		1	2	-	3 (2%)
Evaluative		7	5	1	13 (8%)
Total		123	13	20	156
Ambiguous = 16					172

Table 7.6. *At least* in Speeches corpus

¹² This figure is calculated from 24,193 occurrences (22,743 written and 1,450 spoken) in 97,603,508 w-units (88,000,017 written and 9,603,491 spoken), that is, the number of w-units in the texts accessible by the types of searches carried out. This excludes a small number of BNC texts for which certain header values are missing and which were therefore not able to be included in some searches. It includes both occurrences of those few texts which erroneously appear twice in the BNC. *At least* is significantly less frequent in spoken language (about 151 pmw) than in written language (about 258 pmw).

Type	Position	Initial	Medial	Final	Total
Scalar		98	-	13	111 (47%)
Retreat					
Alternative		35	-	2	37 (16%)
Restrictive		39	-	10	49 (21%)
Epistemic-Retreat		5	-	-	5 (2%)
Evaluative		21	12	1	34 (14%)
Total		198	12	26	236
Ambiguous = 16					252

Table 7.7. *At least* in Reviews corpus

Type	Position	Initial	Medial	Final	Isolate	Total
Scalar		127	3	34	6	170 (37.1%)
Retreat						
Alternative		9	1	1		11 (2.4%)
Restrictive		1	-	-		1 (0.2%)
Epistemic-Retreat		21	-	1		22 (4.8%)
Evaluative		241	12	1		254 (55.5%)
Total		399	16	37	6	458
Ambiguous = 9						467

Table 7.8. *At least* in Conversation corpus

The most salient difference is that in conversation Evaluative *at least* is dominant, while written language contains proportionately more occurrences of Scalar *at least* and many more of Retreat *at least* than conversation. Evaluative *at least* is perhaps twice as common in conversation as in written reviews and over seven times more common than in speeches (fig. 7.5).

Written language is generally more conservative and lags behind spoken language. So a possible explanation for the greater incidence of Evaluative *at least* in conversation is that it is a more recent sense. This accords with the findings from the historical data (section 7.2). A second plausible

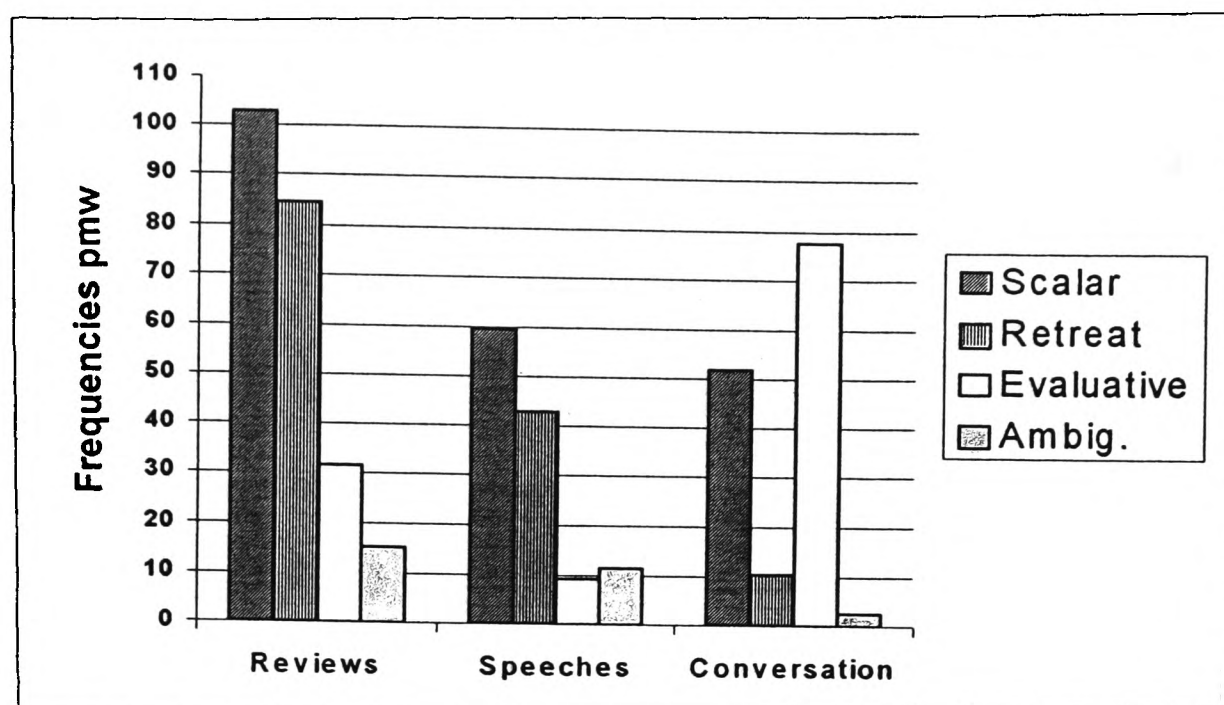


Fig. 7.5. Distribution of *at least* across three genres

explanation is that conversation is almost always more context-bound and expresses speaker subjectivity more than written language.

7.3.2 Rhetorical contexts of *at least*

Retreat *at least* in PDE occurs in Expansion relations; it introduces rhetorical Satellites. The effect of Restrictive *at least*, for instance, is to create a separate intonation unit for its host, which becomes an ELABORATION of the previous idea (36).

- (36) Clearly these are anti-competitive, | *at least* in principle (Sp, Howe)

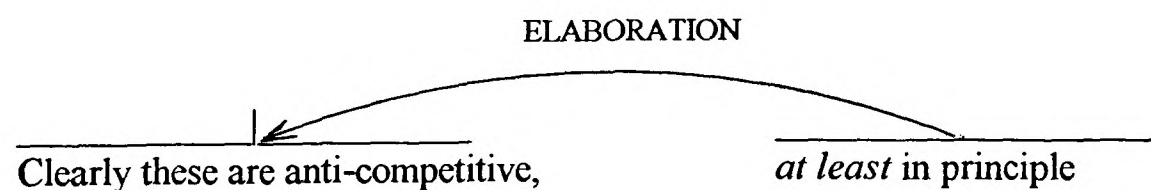


Fig. 7.6. RST schema for (36)

The Alternative patterns *or at least* and *if not p, at least q* are both found in PDE but are relatively less common than in the C18th and C19th data (table 7.9).

	Reviews (n=236)	Speeches (n=156)	Conversation (n=458)
<i>or at least</i>	33 (14%)	11 (7%)	12 (3%)
<i>if (not) p, at least q</i>	7 (3%)	5 (3%)	

Table 7.9. Alternative *at least* patterns in PDE corpora

Although S-N order is more usual, Alternatives can appear in N-S order as in (37).

- (37) a. ... we constantly posit states of completion in our minds, | as solace *at least*, | *if not* delusion. (Rev)
 b. ... with the tenants *at least* supporting an initiative by the local lord | *even if not* instigating reform themselves (Rev)

Epistemic-Retreat *at least* occurs mainly in the EVIDENCE relation and in the following pattern (table 7.10, fig. 7.7). It is discourse connective in that it cannot begin a discourse, but necessarily refers back to the previous idea(s).

Claim (nucleus)	<i>at least</i>	Evidence (satellite)
a. They're alright down here ...	<i>at least</i>	they were the last time we had them. (Conv)
b. The Treasury is, in the old cliché, the Department that "knows the price of everything and the value of nothing"	That <i>at least</i>	is the caricature (Sp)
c. Many years later they become lovers --	<i>at least,</i>	it is dimly possible to construe the text in that way (Rev)

Table 7.10. *At least* in the EVIDENCE relation

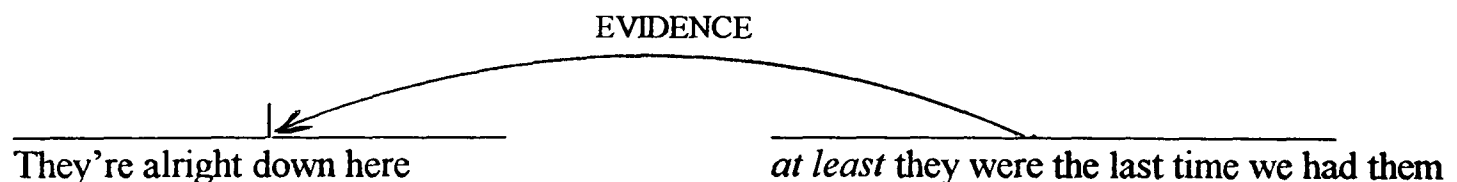


Fig. 7.7. RST schema for table 7.10(a).

At least q implies 'perhaps p', where p is higher up the scale. But in the case of Epistemic-Retreat *at least*, p has just been asserted. The positive 'perhaps p' implicature therefore becomes merely a positive-sounding way of saying 'maybe not quite p', and is interpreted as a correction (cf. *or rather*). That is, the 'perhaps p' implicature is weakened. The use may result from tension between the pragmatic principles of 'say as much as you can' (Horn's Q-principle) and 'do not say that for which you lack evidence' (Grice's Quality maxim). In argumentation terms, a strong assertion (Nucleus) followed by a positively-expressed partial retreat (Satellite) is more persuasive than a weak assertion.

In PDE, Evaluative *at least* occurs with the CONCESSION relation and with the ANTITHESIS relation. Both involve a contrast with a lower point on a scale of benefit. In the CONCESSION relation,

where one idea is more valued than the other, the *at least* idea is the valued idea and is usually second, e.g. *p but at least q* or *although p, q* (table 7.11, fig. 7.8).

Negatively evaluated idea (satellite)	[but] <i>at least</i>	Positively evaluated idea (nucleus)
a. I would be ill advised to compare myself .. with Lord Chancellors of the past	but <i>at least</i>	Sir William could distinguish me from Lord Jowitt (Sp, Irvine)
b. -my son-in-law's .. one of them is not too good	but <i>at least</i>	he does decorate .. he does decorating and things and he helps like that (Conv)
c. Patrick is the most credible of these saints ..., and although not Irish by birth,	he <i>at least</i>	visited and lived in Ireland (Rev)

Table 7.11. *At least* in the CONCESSION relation.

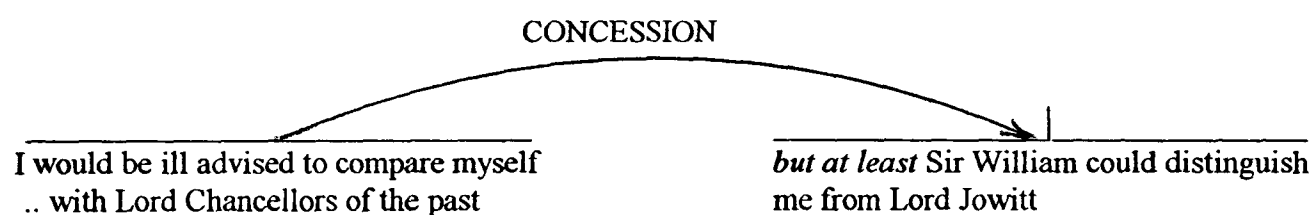


Fig. 7.8. RST schema for table 7.10(a).

The incidence of expressed concession and of collocation with the adversative conjunction *but* is shown in table 7.12.

	Reviews	Speeches	Conversation
Expressed concession idea	23/34 (68%)	11/13 (85%)	60/170 ¹³ (35%)
<i>but + at least</i>	14/34 (41%)	9/13 (69%)	39/254 (15%)

Table 7.12. Concessive contexts of *at least* in three PDE corpora

It is lowest in the Conversation data, where simple non-contrastive Evaluatives are common.

The *at least* host in these examples is nuclear or in focus and usually comes in the S-N order.

But in the ANTITHESIS relation, which is common in the conversation data, the *at least* host often comes first, i.e. *at least q, p* (table 7.13, fig. 7.9).

¹³ There were 170 Evaluative occurrences in the Conversation corpus which could be analysed for contrastive context, due to quality of transcription.

<i>at least</i>	Positively evaluated idea <i>q</i>	Antithetical idea <i>p</i>
a. <i>-at least</i>	you can see the weir today..	we couldn't see it on .. Wednesday could we? (Conv)
b. <i>-at least</i>	it's a white cloud	not a black one (Conv)
c. Besides, as Will says, <i>at least</i>	Charlie is a regular meat-eating beer-drinking cigarette-smoking human being	and not some sanitarium monk. (Rev)
c. <i>At least</i>	the debate is now better informed	than the rumours that we previously had (Sp, Clarke)

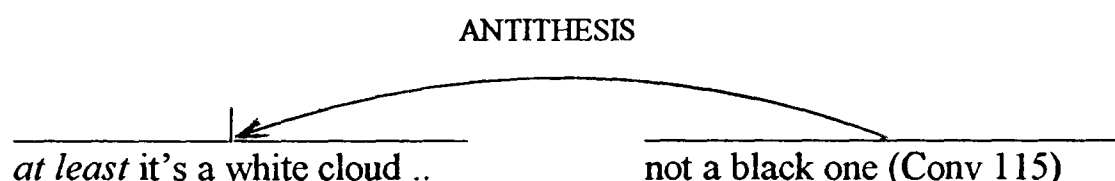
Table 7.13. *At least* in the ANTITHESIS relation

Fig. 7.9. RST schema for table 7.11(b).

The *at least* host is still nuclear. The second idea increases positive regard for the *at least* host.

Thirdly, particularly in conversation, *at least* is often simply a marker of positive evaluation, similar to *fortunately*, without any explicit contrast. Often, it is possible to recover contrasting (concessive and/or antithetical) ideas from the context. In (38) the hearer does just that: B's reply shows that he has understood A to implicate 'you can't cover yours':

- (38) A *-at least* I can cover mine [Concession: 'I have got spots'; Antithesis: 'you can't cover yours']
 B -I don't have to cover mine ...
 (Conv)

But this is not always so obvious: in (39), the *at least* does not link its host to any other idea.

- (39) *-at least* it hasn't come to rain again (Conv)

The basic function of *at least q* is to cancel the implicated upper bound 'at most *q*' and replace it with 'perhaps *p*', where *p* is higher on the scale. But Evaluative *at least* seems to be losing this function and therefore losing its scalar sense.

The Conversation data also show that *at least* with conditions can occur before the *if*-clause

(40).

- (40) -space above the A | 'cos *at least* then if ... it doesn't work it's not gonna delete anything (Conv)

The whole condition is now within the scope of *at least*. Again, this suggests a purely favourable-attitude role for *at least*, with weak scalar implications.

The above examples illustrate the three main rhetorical contexts in which *at least* signals positive evaluation: concessive contexts, antithetical contexts and favourable attitude to the host idea. In all three, the ‘perhaps *p*’ implicature is extremely weak.

Finally, in what appears to be a recent incipient extension of the Evaluative sense, *at least* is sometimes juxtaposed with a causal relation (41).

- (41) a. - I like making me own coffee .. *at least* I know I can drink it (Conv)
 [Implied concession: ‘making coffee is a chore’; Implied antithesis: ‘I don’t know if I can drink coffee I haven’t made myself’]
 b. -I’d rather have a rat in the house than a dog ... *at least* a rat’s small (Conv)
 [Implied concession: ‘rats may not be ideal pets’; Implied antithesis: ‘dogs are large’]

(41a, b) can be glossed as ‘I like making my own coffee [because] I know I can drink it’, ‘I’d rather have a rat in the house than a dog [because] a rat’s small’ (fig. 7.10).

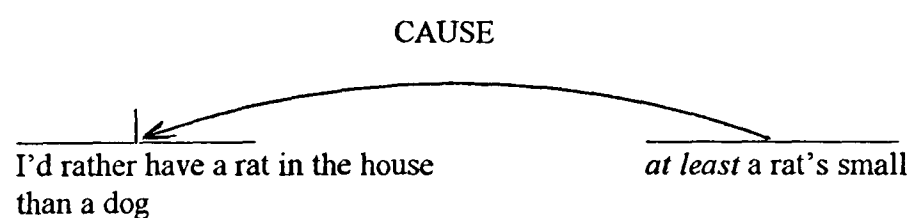


Fig. 7.10. RST schema for example (41b)

The *at least* host gives a reason for the speaker preference expressed in the previous segment, so that the *at least* is perhaps beginning to indicate a causal relationship on the attitudinal-evaluative plane.¹⁴

7.3.3 Layering

As seen in section 7.1.1, all uses of *at least* are scalar in some way and involve a contrast between two values on the relevant scale. Structural ambiguities are possible among the Scalar, Retreat and Evaluative senses. The intended reading is normally clear from the context, and the context can be manipulated to favour other readings. As Lyons points out, “ambiguity which results from

¹⁴ On the relationship between concession and causality, see König and Siemund (2000) and Verhagen (2000).

absolute homonymy cannot be eliminated by manipulating the grammatical environment” (1995: 57).

The readings are therefore polysemies. This is shown in (42a-c).

- (42) a. *At least 25 per cent* of original programmes [and maybe more] will now come from independent companies with a proper proportion of European origin. (Per) (Scalar)
- b. [American films will no longer dominate programming.] *At least, 25 per cent of original programmes will now come from independent companies with a proper proportion of European origin.* (Epistemic-Retreat)
- c. [American films still dominate programming, but] *at least 25 per cent of original programmes now come from independent companies with a proper proportion of European origin.* (Evaluative. Scalar reading still possible)

More interesting, perhaps, are hybrid cases, which seem to straddle boundaries between senses¹⁵. The ‘layering’ of polysemies resulting from diachronic shifts was discussed in chapter 2. Cases of multiple readings show that the layers revealed in synchronic analysis need not be discrete, but can merge into one another via transitional uses. This will now be illustrated with Alternative / Evaluative readings.

As seen in section 7.2, there is a history of *at least* occurring in the apodosis of a conditional construction, in the pattern *if [not] p, [then] at least q*, to indicate an inferior alternative. Kortmann (1997: 203) shows that for adverbial subordinators there is an implicational hierarchy Condition > Concessive condition > Concession. The evidence points to Condition as “the semantic interclausal relation which underlies the central CCC [Causal, Conditional, Concessive] relations” (1997: 204). Expressions of concessive condition tend diachronically to become markers of concession (König 1985, Kortmann 1997: 199, König and Siemund 2000: 343). Since conditions are hypothetical and concession is factual, this shift involves a movement from irrealis to realis. This is the case of English *if*, which in realis contexts has acquired the sense of ‘albeit’, ‘although’. It has already been seen that the semantic shift of *at least* to the Evaluative sense may have been influenced by its link to conditionals. The following discussion aims to show that this shift is apparent in the synchronic

¹⁵ The term ‘hybrid’ is taken from Heine et al (1991): “rather than dealing with a replacement of *X* by *Y*, we observe a continuous decrease of *X* properties and a corresponding increase of *Y* properties. ... we propose to talk of intermediate or hybrid forms. Hybrid forms are part of grammaticalization chains; they are found at the intersection of overlapping stages ... The life span of a hybrid form can be relatively short... However, it may also extend over a period of centuries or even a millennium”. (1991: 231-2)

situation, which reveals the passage of *at least* from Scalar/Alternative to Evaluative.

The *if (not) p, (then) at least q* pattern is still current in PDE with Scalar, Alternative and Evaluative readings depending on context. In (43), for instance, the *at least* host is not realized, *if* is probably conditional, and a predominantly Scalar reading seems appropriate.

- (43) If we cannot have a common defence policy now, | ... then - | so they say | - *at least* the European Council should be able to order WEU to carry out a deployment (Sp, Davis)

But in the next example, there can be a dual interpretation.

- (44) .. if there have to be mega-rich, let them *at least* spend their fortunes with the style and generosity of the Rothschilds. (Rev)

It is not clear to what extent the *if*-clause concedes that there have to be mega-rich, and to what extent it is conditional. The *at least* is correspondingly uncertain between Evaluative and Alternative interpretations. It is similar to the C15th and C16th examples in (32) above, and like them is best seen as Concessive Conditional.

Finally, in (45), the *at least* hosts are clearly realized situations.

- (45) a. John Keay, if not original, is *at least* hugely readable and illuminating (Rev)
b. if her writing is not quite the match of her masters', then *at least* she has chosen well. (Rev)

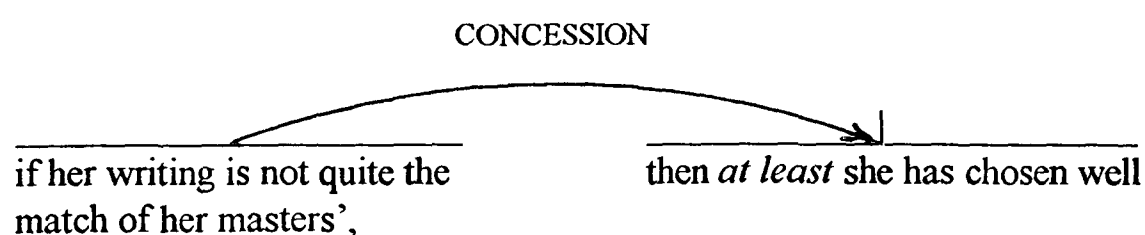


Fig. 7.11 RST schema for (45b)

The *if* seems concessive: the equivalent of 'although'. Both the protasis and the apodosis seem to be entailed, and the Evaluative interpretation is therefore stronger.¹⁶

At least also continues to occur in what König (1985: 264) calls 'irrelevance conditionals' (concessive conditionals of the type *even if p, [then] q*), as in the impersonal examples (46).

¹⁶ Examples (45) sound old-fashioned for the 1990s. What plausibly occurred is that first *though* and then *if* underwent shifts from conditional to concessive; conditional *though* died out and concessive [*al*]*though* has now (almost) eclipsed concessive *if*. Cf. *even though* and *even if* which are still interchangeable in certain contexts.

- (46) a. - .. then you're consider whether being took on | and *even if* you're *not* took on at that place | *at least* you've got something to say you've done | when you're applying for another placement (Conv)
 b. - *even if* it's only ten minutes a day | .. *at least* it's something (Conv)

The relation is largely concessive (*even if* as concessive conditional), and the *at least* predominantly Evaluative.

In general, then, an Alternative *at least* turns into an Evaluative *at least* when the host asserts a realized situation. In the light of the postulated implicational hierarchy mentioned above (condition > concessive condition > concession), the synchronic situation suggests both the probable origin of the Evaluative sense, and the gradualness of its emergence. This is consonant with the historical data (section 7.2). There is no evidence of a sudden metaphorical leap to the attitudinal domain. Rather, the Evaluative sense of *at least* is coerced by its rhetorical contexts.

The PDE data suggest for Alternative-Epistemic-Retreat *at least* and for Scalar-Alternative-Evaluative *at least* what Cruse calls 'sense-spectra', whereby "variants of a single lexical form would seem to be more appropriately visualised as points on a continuum — a seamless fabric of meaning with no clear boundaries" (1986:71). The spectra are best accounted for in terms of reweightings of implicatures. They illustrate what Croft refers to as "the impossibility of separating conventional from nonconventional aspects of meaning in a particular context of use" (2000: 117). The historical data on *at least* examined in section 7.2, together with the observed range of PDE uses within and across text types, suggest that the range of uses observable in the PDE synchronic situation results from gradual, context-induced diachronic meaning shifts. What looks like a 'spectrum' from a synchronic viewpoint can be seen as 'layering' (ch. 2.4) from a diachronic one. The data also highlight the strength and durability of *at least*'s defeasible 'utterance-type meanings' (Levinson 1995).

7.4 Conclusion

In the PDE data, *at least* occurs almost exclusively in the rhetorical contexts illustrated

above.¹⁷ *At least* implies an unknown quantity or degree. It cancels the implicature (from Horn's Q-Principle of 'say as much as you can' (Horn 1984, 1996)) that no more is the case than what is said. There is a transparent semantic relation among the polysemies of *at least*, although the Evaluative and Retreat senses are not predictable from the original scalar sense (cf. French *du moins* vs *au moins* and Spanish *por lo menos* vs *al menos*). The data suggest an underlying, prototypical meaning associated with *at least*, viz. the idea that an acceptable minimum is satisfied on some scale where more is better from the speaker's point of view — for example, more newsworthy. Therefore, *at least* is always scalar and always positive.¹⁸ Moreover, the speaker gauges that the hearer's expectations, as created either by the previous text or by background knowledge, are of a lesser quantity, of a lower point on the scale, so that an explicit or implicit contrast is always present.

What varies among the types of *at least* is the type of scale and the discourse plane.

Diachronically, the changes in scales that have led to Epistemic-Retreat and Evaluative senses are as follows:

Scale of quantity > scale of strength/intensity > scale of epistemic commitment (truth) (Epistemic)
 Scale of quantity > scale of strength > scale of desirability > scale of value (Evaluative)

The range of contexts from typical scalar contexts (quantities) through less typical scalar contexts to typical Retreat and Evaluative contexts points to pragmatic inferencing as the mechanism for the extension of *at least*. Senses extend gradually via intermediate situation types in what Heine et al call 'context-induced reinterpretation' (1991:65-70). In each case there is an eventual shift from the content to the attitudinal plane and a move towards more subjective, speaker-centred meaning. This move results from a gradual reweighting of implicatures. Scalar *at least* means 'minimum n, perhaps more', where n is some degree or quantity. It implicates 'certainly n' and 'more is better' (cf. *at most*).

¹⁷ The patterns illustrated are not claimed to exhaust the syntactic configurations in which *at least* can occur. The aim is not to account for the range of acceptable structures, but to explore what Fillmore and Atkins call "polysemy resulting from a transfer of a semantic frame to a new domain", not "the kind that reflects merely the accommodation of a word to different syntactic patterns" (1992: 101) (see also chapter 1.4.5).

¹⁸ The shared positive nature (the notion that the higher on the scale, the better) of all types of *at least* refers, for Scalar and Retreat uses, to how the writer evaluates the idea as contributing to his or her rhetorical purposes, not to whether, in the socio-physical world, the proposition expressed would be considered a good or a bad thing.

It has been seen that contexts in which the ‘certainly *n*’ implicature is salient lead to an Epistemic-Retreat sense for *at least*, and contexts in which the ‘more is better’ implicature is salient lead to an Evaluative sense. Frequency of occurrence of the relevant rhetorical contexts not only strengthens these implicatures, but also gives *at least* an information structuring role. The patterns involve *at least* in introducing either relatively non-salient material or relatively salient material; that is, in backgrounding a rhetorical retreat, and in focusing a positively evaluated idea. Tables 7.14 and 7.15 summarize these points (*q* = *at least* host; *p* = contrasting, higher point on scale).

	Period 1 & Period 2			Period 2
Type	Scalar	Alternative-Satellite	Restrictive	Epistemic-Retreat
Rhetorical profile				
Rhetorical relation	-	Elaboration	Elaboration	Evidence
Nuclearity	-	<i>at least</i> host = S	<i>at least</i> host = S	<i>at least</i> + host = S
Discourse plane	content	content	content	attitudinal
Interpretation				
Scale	quantity	strength	domain	epistemic commitment
Salient implicature	perhaps <i>p</i>	certainly <i>q</i>	certainly <i>q</i>	not (certainly <i>p</i>)

Table 7.14. Typical characteristics of *at least* in the development of the Epistemic-Retreat sense

	Period 1 & Period 2		Period 2
Type	Scalar	Alternative-Nucleus	Evaluative
Rhetorical profile			
Rhetorical relation	-	Condition	Concession
Nuclearity	-	<i>at least</i> host = N	<i>at least</i> host = N
Discourse plane	content	content	attitudinal
Interpretation			
Scale	quantity	strength/desirability	value
Salient implicature	perhaps <i>p</i> .	more is better	anti- <i>q</i> is a bad thing

Table 7.15. Typical characteristics of *at least* in the development of the Evaluative sense

Morpho-syntactically, *at least* shows increase in scope: from NP modification to VP and clause modification. It undergoes successive category changes: P+NP > P+bare N > VP-Adv > S-Adv. Connective *at least* (Epistemic-Retreat) acquires its own intonation unit (suggesting a possible further category change, but see chapter 1.4). The semantic abstraction and subjectification, together with these morpho-syntactic changes undergone by *at least*, are common to other types of

grammaticalization. Layering is particularly evident in the *at least* development: quasi-evidential and quasi-evaluative uses are still common in the types of constructions in which *at least* has a long history of occurrence. These constructions and the rhetorical patterns they enter into have shown considerable stability over time.

This discussion of *at least* has emphasized the influence of the discourse context in lexical semantic change. By examining the rhetorical relations with which *at least* has habitually occurred, it is possible to appreciate the interdependence of change in this expression with other developments in English, such as *if* conditional > *if* concessive, exclusive *or* > inclusive *or*. It has been shown that *at least* moves from objective meanings towards subjective ones; from reference on the socio-physical plane to reference on the epistemic plane. Its development is partially the result of pressures from other changes taking place in the constructions it regularly occurs in, particularly the condition > concessive condition > concession tendency. Its history therefore supports the claim that internal semantic change is predominantly unidirectional.

CHAPTER 8

OF COURSE

There is a wide difference to be made between Words *of course*, ... and Essential Words

(1680, Lampeter-8)

8.1 Introduction	187
8.2. The history of <i>of course</i>	189
8.2.1 The data	189
8.2.2 The origins of adverbial <i>of course</i>	189
8.2.3 Reanalysis of <i>of course</i> in the C18th	193
8.2.4 Functional split of <i>of course</i>	194
8.2.5 Connective implicatures of <i>of course</i>	198
8.2.6 Summary	200
8.3 <i>Of course</i> in PDE	201
8.3.1 Hedge <i>of course</i>	202
8.3.1.1 Concession / contrast	204
8.3.1.2 Background	207
8.3.1.3 Topic transition	209
8.3.1.4 Topic introduction	211
8.3.1.5 Irony	216
8.3.2 Emphatic <i>of course</i>	217
8.3.3 Distribution of <i>of course</i>	218
8.4 Summary	220

8.1 Introduction

Analysis of corpus data on *of course* in present-day English shows two distinct uses: as a 'relevance' hedge to mean 'naturally', with a flat or rise-fall intonation, and as an emphatic, usually for affirmation, with strong stress on *course*. For example, (1) is ambiguous between a hedge (parenthetical) reading, with comma intonation after *of course*, and an emphatic (integrated) reading, with a single intonation contour for the clause and stress on *course*. In the hedge reading, *of course* introduces a new idea. In the emphatic reading, *of course* contradicts the idea that Asians are not different from Europeans.

(1) *Of course* Asians are different from Europeans ... (Sp)

These two senses appear in the *Collins Dictionary* as follows:

¹Relevance here is used in its common sense, not its Relevance Theory sense.

course

23. **of course.** a. (*adv.*) as expected; naturally. b. (*sentence substitute*) certainly; definitely.

OED2 distinguishes adjectival, adverbial and sentence-adverbial *of course*, this last (37c)

corresponding roughly to the two *Collins Dictionary* senses:

course

37. **of course.**

a. *adjectival.* Belonging to the ordinary procedure, custom, or way of the world; customary; natural, to be expected. Now esp. in *a matter of course*.

b. *adverbial.* In ordinary or due course, according to the customary order, as a natural result. *†of common course:* ordinarily, as an every-day occurrence.

c. Hence, in qualification of the whole clause or sentence: Naturally, as will be expected in the circumstances; for obvious reasons, obviously. (Sometimes used as an emphatic affirmative reply.)

The *Cobuild* corpus-based dictionary lists four uses, plus *of course not*:

course ...

1. You say *of course* 1.1 when you are briefly mentioning something that you expect other people already realize or understand, or when you want to indicate that you think they should realize or understand it. ... 1.2 when you are talking about an event or situation that does not surprise you ... 1.3 as a polite way of saying yes, of giving permission, or of agreeing with someone. ... 1.4 in order to emphasize a statement that you are making, especially when you are agreeing or disagreeing with someone.

2. **Of course not** is an emphatic way of saying no.

The dictionaries describe *S-Adv of course* in terms of more or less fine-grained senses, but all three show roughly the split between the ‘naturally’ use and the emphatic use (table 8.1).

OED2	Collins	Cobuild
naturally, as expected	naturally (23a)	given information (1.1) unsurprising information (1.2)
(emphatic affirmative reply)	definitely (23b)	polite agreement (1.3) emphatic (agreement) (1.4) emphatic negation (2)

Table 8.1. Comparison of dictionary definitions of *of course*

Of course is often discussed as an example of a disjunct or a parenthetical (McCawley 1982: 95-6; Quirk et al. 1985: 609; Espinal 1991: 727; Hoyer 1997: 184; see also chapter 1, note 9). This analysis is plausible for the ‘naturally’ use, where *of course* is backgrounded with regard to its host, but not for emphatic *of course*. This can be seen as a syntactic distinction between the two senses, leading to potential ambiguity as in example (1) above.

Of course is also sometimes described as having a discourse connective function: Quirk et al. classify it as a concessive, contrastive conjunct (1985: 631ff)²; Hoyo describes *of course* as a 'contrastive conjunct' (1997: 154). Finell (1992) classifies *of course* as a topic changer.

Section 8.2 traces the development of adverbial *of course*, and section 8.3 describes the PDE situation of contextual modulation.

8.2 The history of *of course*

8.2.1 The data

Table 8.2 shows the sources of the historical data on *of course*.

Period	No. of occurrences in full-text sources		No. in non-full-text sources OED2
	Hist	Lamp	
1550-1599			2
1600-1649			
1650-1699	4	13	1
1700-1749	2	11	10
1750-1799	20		25
1800-1849	123		23
1850-1899	160		

Table 8.2. Breakdown of historical data on *of course*

The development of *of course* should be seen in the context of other uses of *course*. Data were collected on these, and the following discussion of the origins of the expression draws on many occurrences in the historical data of the precursors of *of course*: *by course*, *by common course*, etc.

8.2.2 The origins of adverbial *of course*

The noun *cours* came from French into Middle English complete with literal and figurative senses of 'running', which involves movement in a particular direction, and already extended to refer to the path or direction of the running. Two of the sub-senses which developed from this were *course* as 'custom, norm' and *course* as 'order, sequence'.

² For Quirk et al (1985: 631), although both disjuncts and conjuncts are syntactically relatively detached, a conjunct has "the function of conjoining independent units rather than one of contributing another facet of information to a single integrated unit".

The sub-sense of *course* as ‘custom, norm’ is illustrated in (2).

- (2) a. Desireth nother Mariage / Ne yit the love of paramours, / Which evere hath be *the comun cours* / Amonges hem that lusti were (CMEPV)
 b. Hit is ordeyned that no maner of Bruer ... sylle noone Ale butt onely bi meser ensealyd aftir *the Course and Custome* of alle Bruers. (c1380, MED)

In this sense it collocated with *common*, in the expression *by common course*, and with the nouns *kynde, nature, law, life* in the expression *by (common) course of N*, to indicate that something occurred normally or naturally (3). Early contexts of *of course* are thus the opposite of those of *after all*, which involve a non-natural sequence of two events (chapter 5.2).

- (3) a. It maketh me backwarde to meve, whan my steppes *by comune course even-forth* pretende (1532, MED)
 b. Soo as syr launcelot had waked as longe as hit had pleasyd hym / thenne *by course of kynde* he slepte (1470-85, OED)
 c. Sicke thinges pat maye not be done *by comun course of nature* .. neipeles they be possibil (c1450 MED)
 d. That dampned was this knyght for to be deed, / *By cours of lawe*, and sholde han lost his heed (1387-1400, Chaucer, Wife of Bath’s Tale)

The adverbial expression *by course* gained wide currency in ME, its meaning extending to include ‘naturally’, ‘in due course’, ‘in order’, ‘as a result’ (4).

- (4) a. arisith in phe Oryent and *bi successife cours* accendith into pe hieste poynt of pe mydday spere. (1450 MED)
 b. Cornelius .. translated it into latyn .. But he shope it so short pat no shalke [‘no man’] might Haue knowlage *by course* how pe case felle (c1450 MED)

Later, however, (C16th - C17th), its meaning narrowed to ‘by turns’ or ‘in order’ (5).

- (5) a. every second yeare *by course* (1587, Th Thomas)
 b. in the after none Iohn Corowe and he did read *by Course* vnto me (1599-1605, Helsinki)
 c. If ther be above two Organistes at once, two shall allwaies attend; if there be but two in all, then they shall wayte *by course*, one after an other (1603-4, MEMEM)
 d. A Kingdome rightly possessed *by course*, Containes more joy than is usurpt by force (1639, MEMEM)

By course overlapped with *of course* in the C17th, then died out. The data sources used here provide no C18th examples of *by course*. The only sense given it by Johnson (1755) is that of ‘by turns’: it is entered under the sense of *course* as ‘order of succession’ and may by 1755 already have

become archaic. It seems to have been entirely replaced by *of course*, which sometimes occurs in the C18th with the ‘by turns’ meaning, as in (6).

- (6) We always succeeded *of course*; no jumping over heads. (1764, OED)

Adverbial use of *of course*, meaning ‘customarily’, ‘ordinarily’, ‘by natural process’, developed in Early Modern English. It overlapped in the C17th with *by course*, and seems to have arisen from the sense of *course* as ‘custom’. The development of *of course* from a Middle English collocation to its Modern English status of SAdv has involved the four overlapping stages outlined in table 8.3.

Stage	Category	Status	Expression
Stage I	P + NP + of + NP	collocation	by/of (the) (common) course (of (the) N)
Stage II	P + bare N	idiom	of course
Stage III	reanalysis as VP-Adv	lexeme	of course
Stage IV	reanalysis as S-Adv	lexeme	of course

Table 8.3. The development of *of course*

As shown in the table, *of course* has undergone two reanalyses — first from P+bare N to VP-Adv, second from VP-Adv to S-Adv — each of which seems to have been at least partially context-driven. These two restructurings are outlined below, together with the likely motivations for them.

OED2's earliest adverbial citations of *of course* are from 1541 (NP modifier) and 1542 (VP modifier), but not till the C17th does it become common (7 and 8). It occurs immediately after the noun it modifies. When it modifies a verb, it also usually occurs in final position, i.e. after the VP, but occasionally before, when the complement of the V is long, as in (8a,c).

- (7) NP modifier
- a. words *of course*, tearmes of disport (1604, MEMEM)
 - b. And this officer *of course* is to perform this ceremony of riding through the city (1662, Hist)
 - c. Hitherto your Objections have been Vulgar and *of course* (1702, OED)
- (8) VP modifier
- a. .. then the court, either by special order or *of course*, grants Publication, or a publishing of the proofs (1656, EMED, Blount)
 - b. 16th. To the Duke where we met *of course*, and talked of our Navy matters (1662, Hist)

- c. When the Lds have Recd & Read the [articles] then *of Course* they will give the Lds Impeached A Certaine day to put In their Answer & then things will goe Currantly on In order to A speedy triall (1678-9, Newdigate, referring to a murder trial)

As a noun modifier, it dwindles in the C19th, occurring mainly with a small set of collocates (especially *words of course* and *matter of course*) and occurs to-day only in fossilized expressions (9).

- (9) those whose ideas are the object of censorship risk death and imprisonment as *a matter of course* (Per)

It is with *of course* as a VP modifier and S-Adv that the rest of this section will be concerned.

8.2.3 Reanalysis of *of course* in the C18th

Towards the end of the C17th, *of course*, which hitherto occurred primarily post-VP, starts to be common in medial (i.e. post-aux, pre-verbal) position (10).

- (10) a. her defects, if any, are *of course* supplied by her Husbands Experience (1676, Lampeter)
 b. The Fire of Youth will *of course* abate (1712, OED)
 c. they learn Love-Songs, ... This *of course* makes them wanton, and so they think of Husbands, before they are capable to choose for themselves (1730, Lampeter)
 d. a Man who reads with any tolerable Understanding would *of course* infer the same thing (1738, Lampeter)

This move may be partly due to information restructuring. By being placed earlier in the clause, the *of course* loses some of its informational salience while the element it modifies gains some. The collocation of *of course* with particular verbs such as *lead*, *follow*, *become*, where the meaning of the verb already implies 'natural consequence', makes *of course* almost redundant (11).

- (11) a. the Enthusiasm or the Pathetique, as Longinus calls it, follows *of course* (1704, Hist)
 b. ... and yet a Divorce follows *of course* (1715, Lampeter)
 c. This leads me *of course* to put your Honors in mind, that your House is almost as much affected by this Clause as the People of Ireland (1720, Lampeter)

Yet final *of course* in examples (11) is in end-focus position. Following the principle of end-focus, *of course* may therefore have moved up the clause (into a pre-verbal slot already open to adverbs), thereby giving the VP greater salience. Compare focus in (12a) and (12b) (= 10d).

- (12) a. If these Places were not frequented, they would fall *of course* (1730, Lampeter)
 b. a Man ... would *of course* infer the same thing (1738, Lampeter)

Of course in the pre-verbal position is not parenthetical; rather, it has reanalysed from a (PP + noun) adverbial to a VP-Adv. The claim here is that the expression does not spontaneously reanalyse, but that the reanalysis is due to the strengthening of the bond between *of* and *course* due to constant collocation, and to the weakening, due to redundancy, of the pragmatic force of *of course*.

Evidence of the gradual reanalysis of *of course* in this period, from a complex structure (PP adverbial) to a simple one (VP-Adv)³, is provided by Johnson's dictionary (1755, 1773). The same citations are used in both the first and the fourth editions, but between the two editions Johnson's analysis changes: in the 4th edition he treats *of course* as a single unit, and his definitions change accordingly (table 8.4).

Johnson's Citation	Johnson's Dictionary entry, I edn 1755	Johnson's Dictionary entry, IV edn. 1773
With a mind unprepossessed by doctors and commentators of any sect, whose reasonings, interpretation and language, which I have been used to, will {I of <i>course</i> /IV of <i>course</i> } make all chime that way; and make another, and perhaps the genuine meaning of the author, seem harsh, strained, and uncouth to me. <i>Locke.</i>	18. <i>course</i> . Series of consequences.	22. <i>Of course</i> . By consequence.
Sense is {I of <i>course</i> /IV of <i>course</i> } annex'd to wealth and power. No muse is proof against a golden show'r. <i>Garth.</i>	18. <i>course</i> . Series of consequences	23. <i>Of course</i> . By settled rule.
Neither shall I be so far wanting to myself, as not to desire a patent, granted {I of <i>course</i> / IV of <i>course</i> } to all useful projectors. <i>Swift.</i>	20. <i>course</i> . Regularity; settled rule.	23. <i>Of course</i> . By settled rule.

Table 8.4. 'Of course' and 'of course' in Johnson's Dictionary, first (I) and fourth (IV) editions.

³ Although these are two discrete analyses, the reanalysis can be said to be gradual in the sense that it is individual occurrences of the expression that are interpreted by hearers as P+noun or VP-Adv, not the abstract form, and the relative proportion of each category shifts gradually (v. ch. 2.4 on gradualness in reanalysis).

By the middle of the C18th, therefore, VP-Adv *of course* is firmly established. The next section shows how *of course* acquired epistemic implicatures with the extension of the now ‘weaker’ expression to new contexts, and how it became a S-Adv.

8.2.4 Functional split of *of course*

Throughout the C18th, there is evidence of a split developing between the ‘normally/usually’ manner sense (Johnson’s “by settled rule”) and the ‘naturally/as a result’ sense (Johnson’s “by consequence”). Contexts suggest which implicatures are likely to have been the stronger⁴: *of course* can be said to be ‘contextually modulated’ (v. chapter 2.4.6) in this period. Thus, the contexts in (13) suggest a ‘normally’ reading, while those in (14) suggest a ‘naturally’ reading.

- (13) a. A Werowance is a Military Officer, who *of Course* takes upon him the Command of all Parties. (1705, OED)
 b. ... the Minister of the Parish, especially in Country Villages, being *of course* resorted to, upon all Occasions of That Nature (1721, Lampeter)
 c. The professors of this pure theology are *of course* easy to be distinguished (1780, 18thb-28)
- (14) a. ... they chop, or pounce with their Hands up and down to cut the Stone or Mineral, going round, which *of course* grinds either of them small (1708, OED)
 b. Our directing the next payment to be made to the Bank would, I believe, possess the Generality of Mankind with an Opinion that we were resolved to abolish the Company..., which would *of course* run the Price of their Stock down .. (1737, OED)
 c. As all energies are attributes, they have reference *of course* to certain energizing substances (1750, OED)

The ‘naturally/as a result’ reading is suggested by the co-occurrence of *of course* with contexts in which the reason for the host idea is expressed in the previous segment. In RST terms, it co-occurs with a CAUSE relation.

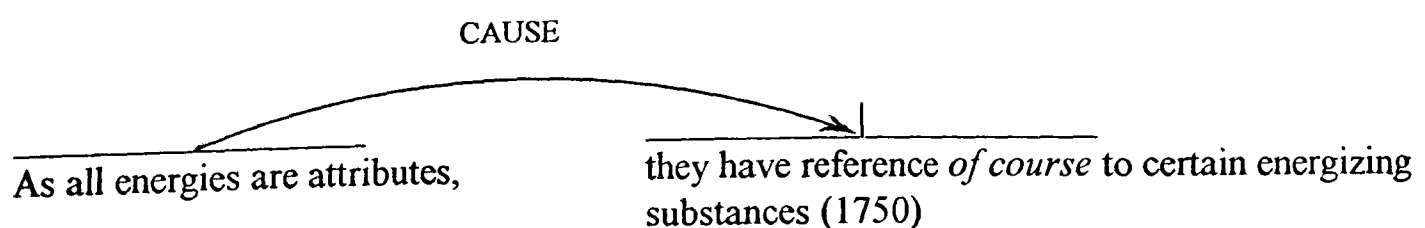


Fig. 8.1. RST schema for (14c)

⁴ Punctuation is too unreliable to be taken as evidence.

In (14a-c) above, there is another mechanism (e.g. *which, as*) to indicate that the second idea results from the previous one(s). Soon, however, *of course* by itself comes to signal ‘as a result’ (15).

- (15) a. The locomotive mania of an Englishman circulates his person, and *of course* his cash, into every quarter of the kingdom. (1786, OED)
 b. Girls ... are often cruelly left by their parents without any provision, and, *of course*, are dependent on not only the reason, but the bounty of their brothers (1792, Hist)
 c. But he loved the Stuart family, and his history is, *of course*, their apology. (1799, Hist)

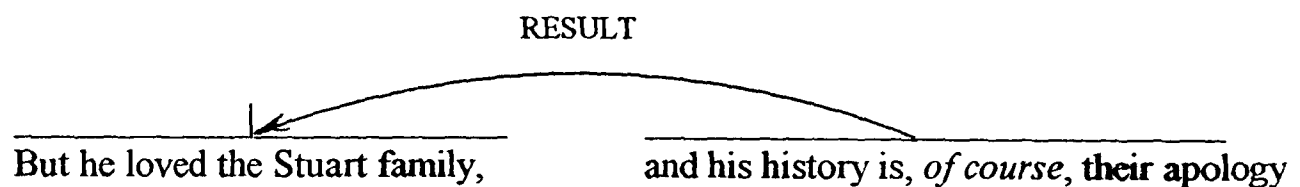


Fig. 8.2 RST schema for (15c)

Of course in this sense of ‘as a result’ acts as a connective on the content plane, linking its host to the previous idea.

The *of course* host idea is presented as resulting from the previous idea. Where the host idea is an irrealis context, however, it tends to express a belief rather than a fact. *Of course* then signals not that a real-world event results from the previous idea, but that a belief results, i.e. an epistemic commitment. This can be seen in (16), where the co-occurrence of *of course* with *must* suggests that epistemic inferences were already being drawn from the early C18th.

- (16) a. As Homer is the Author nearest to those, his Style *must of course* bear a greater Resemblance to the sacred Books than that of any other Writer (1715, Hist)
 b. If this necessary Being hath no change or succession in his nature, his existence *must of course* be unsuccessive (1737, OED)
 c. If a poor child is to be whipped equally for telling a lie, or for a snotty nose, he *must of course* think them equally criminal. (1752, OED)

In these examples, *must* and *of course* are both conveying the idea of ‘natural result or conclusion’. But they suggest a judgement by the speaker, based on evidence from the previous clause, that the conclusion can be drawn. *Of course* co-occurs here with the EVIDENCE relation (fig. 8.3), so it is likely to acquire evidential implicatures.

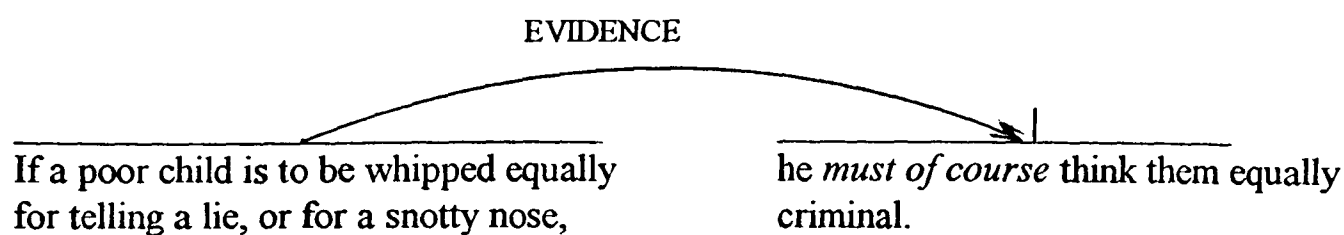


Fig. 8.3. RST schema for (16c).

It seems, therefore, that a re-weighting of implicatures was taking place, with *of course* acquiring, in some contexts, the subjective implicature that the event was natural *in the speaker's judgement*. It is thus ready to become an epistemic speaker-oriented sentence adverb⁵, with the type of context exemplified in (16) acting as a 'bridging context' (v. chapter 2.4.3).

By the C19th, *of course* is very frequent⁶. Most uses are clearly interpretable as epistemic speaker-oriented S-Adv. About half the occurrences with full clauses from this set of data are in medial position and a quarter already in initial position (table 8.5)

Position of clausal SAdv (n=197)	
initial	27%
medial (pre-verbal)	51%
final (post-verbal)	22%

Table 8.5. *Of course* in the C19th

(17) illustrates epistemic senses for *of course* in initial, medial and post-verbal positions.

- (17) a. I find that black pantaloons are considered by them as necessary, and *of course* one would not have them made uncomfortable by the want of what is usual on such occasions (c1815, Hist, Austen, Letters)
- b. upon this, and upon all other matters connected with it, I shall *of course* be most anxious to hear your opinion in detail (1828, Hist)
- c. Mrs C. invented it all *of course*. (c1815, Hist, Austen, Letters)

This shift to epistemic status is a shift in discourse plane (v. chapter 2.3). *Of course* is no longer a part of the representational idea expressed by the clause i.e. indicating that something is natural in the external world, but is instead a speaker comment, i.e. indicating the speaker's opinion of

⁵ As seen above (example 8c), *of course* sometimes occurs in initial position from the C17th, meaning 'naturally', 'usually'. But these occurrences are neither parenthetical nor SAdv by the definitions given in ch. 1.

⁶ Very high frequency can be observed in less formal text from the turn of the C19th. In Jane Austen's letters, for instance, the frequency of *of course* extrapolates to 490 pmw. (cf. frequency in spoken PDE: the BNC gives 284 pmw in the demographically-sampled spoken part and 577 pmw in the context-governed spoken part).

the status of the idea. Contexts such as those in (17) are ‘switch contexts’ (v. chapter 2.4.3), in which it is hard to get a non-epistemic reading.

Emphatic *of course* also appears, in response to questions, in the first half of the C19th (18).

- (18) a. And you were present? *Of course*; why not? (1838, OED)
 b. Do you pretend you can discriminate the wheat from the tares? *of course* not (1840, Hist)

Due to the lack of dialogue in the data, it is not well represented. But it is consonant with the role of *of course* as an emphasizer: other such epistemic markers that can respond to yes/no questions are *clearly, naturally, evidently, indeed*.

In summary, one of the implicatures that developed from ME *cours* was the **idea** of naturalness, ordinariness, custom or default. The collocation with *of* became idiomatised and then reanalysed as a single unit, first as a VP adverb of manner, then as a S-Adv. The S-Adv was first a conjunctive meaning ‘as a result’, and later an epistemic speaker-oriented S-Adv.

Since they overlap in time, in dialect and in register, and since dual analyses must be allowed for (the surface ambiguity allows for the representational and epistemic-modal analyses to co-occur), the different stages should not be thought of as discrete. Based on the available data⁷, the development occurred as plotted in fig. 8.4. The figure shows the long periods of latent ambiguity and the long periods of parallel uses (layering).

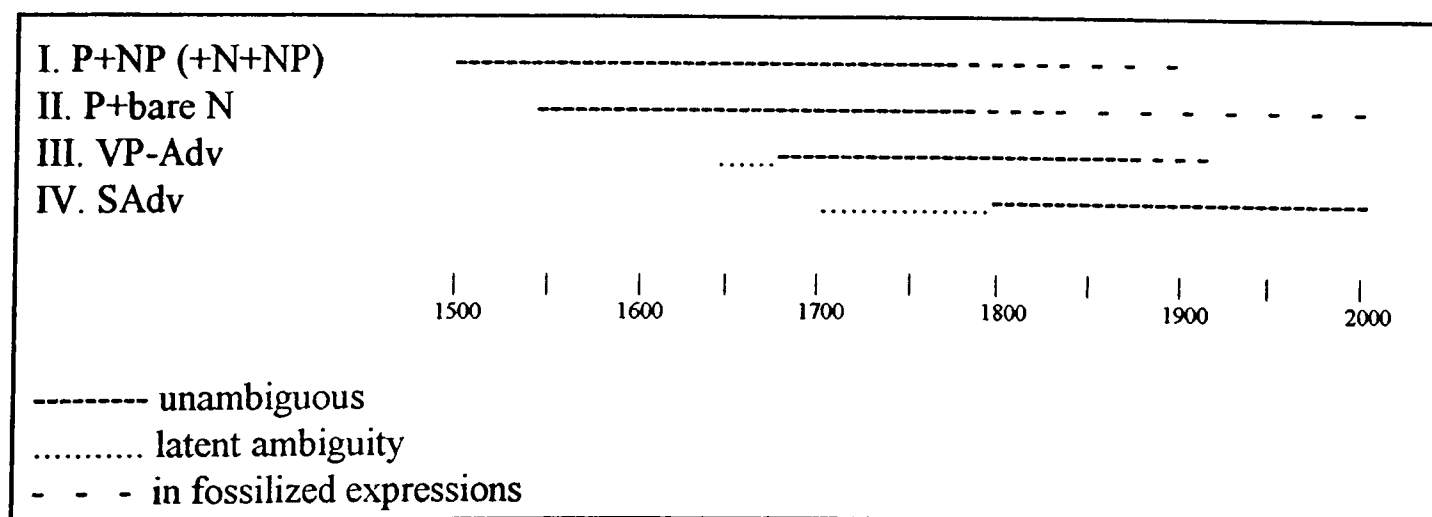


Fig. 8.4 Development of *of course*

⁷ As already mentioned, the available data are likely to be conservative, being written language.

At Stage II (P+bare N), the meaning of *of course* is relatively transparent and the sequence is probably compositional (cf. *out of necessity* in PDE). At Stage III (VP-Adv), it may be starting to become opaque. Routinized expressions often diverge in meaning from their source lexeme, adding a new polysemy. In this case, the demise of the PP *of course* seems to have come at around the same time as the loss of the sense of ‘custom, habitual practice’ for the noun *course*. If so, the resulting opacity of *of course* may have accelerated the reanalysis. At Stage IV (SAdv), the core notion of naturalness in *of course* remains, but becomes gradually more subjectivized. With widening scope, *of course* occurs more frequently in initial and medial position. And there is no longer any pragmatic constraint on the types of events, states or actions with which it can occur. This makes way for a further increase in frequency of use.

8.2.5 Connective implicatures of *of course*

In a further development, *of course* acquires new connective implicatures on the attitudinal plane. Although weakened, the epistemic sense is not lost as a result. Already in its manner VP-Adv sense of ‘normally’, ‘in the normal course of things’, *of course* often occurs in a contrastive context (19).

- (19) a. The town-proper was *of course* the collection of dwellings; *but*, in the vulgar acceptance the same word embraced the entire district or township. (1809, OED)

Since *of course* indicates normality, which is less newsworthy than abnormality, it comes to function as a backgrounder. From the end of the C18th, *of course* beings to co-occur with concession relations in argumentation contexts (20). (In (20a), the presence of *naturally* suggests an epistemic reading for *of course*.)

- (20) a. In the infancy of society, ... chiefs and priests ... must have had unbounded sway. An aristocracy, *of course*, is naturally the first form of government. *But*, clashing interests soon losing their equipoise, a monarchy and hierarchy break out of the confusion .. (1792, Hist)
- b. This demand would *of course* raise the price of labour, *but* if the yearly stock of provisions .. was not increasing, this rise would soon turn out to be merely nominal (1798, Hist)

This tendency continues in the C19th. The typical pattern is an *of course* clause containing backgrounded material, followed by a clause beginning with *but* containing foregrounded material (table 8.6):

Backgrounded segment	<i>but</i>	Foregrounded segment
a. He is a rogue <i>of course</i> ,	but	a civil one (c1805, Hist, Austen)
b. These divisions are, <i>of course</i> , subject to considerable variations, ...	But	if they merely approximate towards the truth, a rise in the price of corn must be both slow and partial (1814, Hist, Malthus)
c. You are <i>of course</i> well aware of all that is stated,	but	Mr G. wished that it should be sent to you (1828, Hist, Peel)
d. The writing is, <i>of course</i> , reversed by this process;	but	the paper to which it is transferred being thin, the characters are seen through it on the other side, in their proper position (1832, Hist, Babbage)
e. <i>Of course</i> it is impossible, if no one attempts to restore it;	but	if all willed it, how would it be impossible (1840, Hist, Newman)

Table 8.6. Backgrounding *of course* in first half of C19th

Of course co-occurs here with the CONCESSION relation (fig. 8.5).

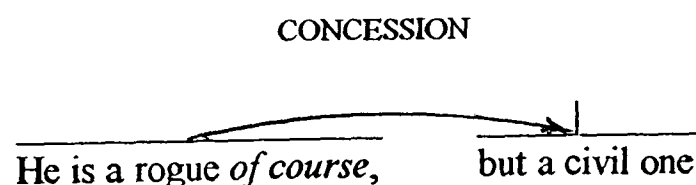


Fig. 8.5. The RST schema for example 8.6a

The pattern becomes relatively commoner in the second half of the C19th (21), when we also find likely parenthetical occurrences (21b,d).

- (21)
- a. *Of course* the success of any species of bee may be ... altogether independent of the quantity of honey which the bees could collect. But let us suppose that this latter circumstance determined ... the numbers of a humble-bee .. (1859, Hist)
 - b. He gave no theory, *of course*, ... But the mere mention of an annual period is interesting in the history of tidal theory (1882, Hist)
 - c. He has no visionary schemes of revolution and transformation, though *of course* he would like his class to rule (1886, Hist)
 - d. *Of course*, I am naturally a partial judge of my father's character; but this I may say, that during my experience of over seventy years I have never known a more incessantly industrious man (1885, Hist)

For *of course* in initial position it is particularly difficult to judge whether it is parenthetical, but again the presence of *naturally* in (21d) makes a non-parenthetical reading difficult (cf. 20a). What the data strongly suggest is that parenthetical status arose in contrastive contexts, where it regularly had contrastive/concessive implicatures.

Of course is clearly associated with backgrounding. In contrastive and concessive relations between discourse segments, the usual order is S-N, in accordance with end-focus. Where this order is reversed, as in (21c), *of course* emphasizes the backgrounded status of *he would like his class to rule*, despite that segment's final position. The co-occurrence of *of course* with a concessive or contrastive relation is well established in the second half of the C19th (table 8.7).

S-Adv <i>of course</i> 1800-1849 (n=116)	S-Adv <i>of course</i> 1850-1899 (n=90)
In contrastive contexts	In contrastive contexts
followed by 'but' 18%	followed by 'but' etc 39%
preceded by 'though' 2%	preceded by 'though' 7%

Table 8.7. 19th *of course* in contrastive contexts

Of course does not signal a concessive relation; it merely co-occurs with such a relation. The relation is signalled by another DC such as *but* or *though*. As pointed out in ch. 1.4, parentheticals that signal relations only occur in the second segment of the relation.

8.2.6 Summary

In the C18th, the two main *of course* uses, 'normally' and 'naturally', bifurcate. The latter *of course* co-occurs with the RESULT relation, and comes to mean 'as a result'. It also co-occurs with the EVIDENCE relation and comes to signal subjective certainty (i.e. speaker commitment). During the C19th, this epistemic *of course* presumably split into an 'emphatic' sense, where epistemic commitment was salient, and a 'hedge' sense where the 'obviously' meaning was salient. The 'obviously' use co-occurs first with the EVIDENCE then with the CONCESSION relation. It indicates that its host idea is backgrounded in relation to the adjacent segment (fig. 8.6).

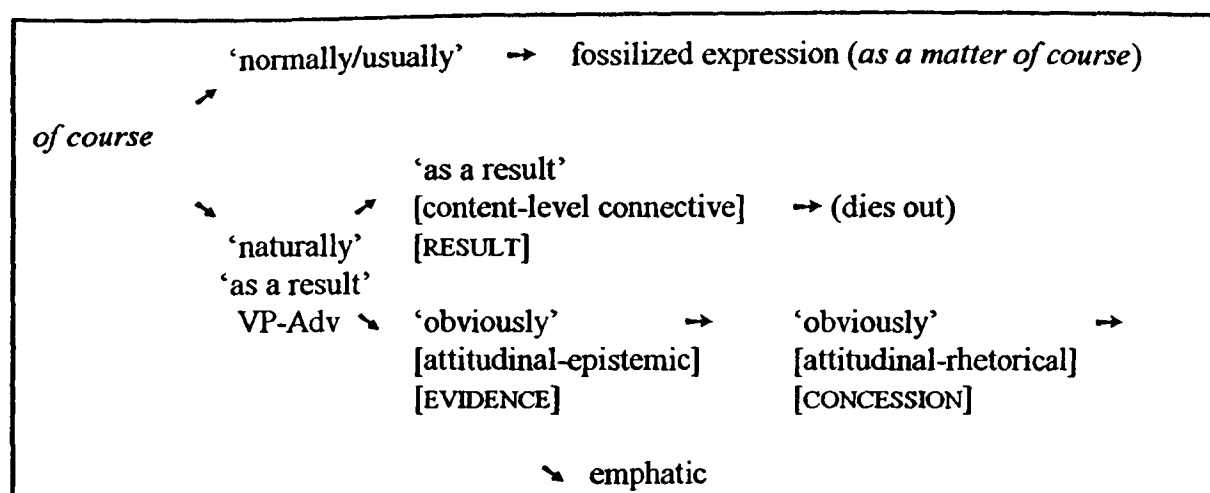


Fig. 8.6. Development of *of course* in C18th and C19th

According to König and Siemund, “causal connectives may develop into concessive ones” (2000: 346). The present data, however, do not support the development of concessive *of course* directly from causal (‘as a result’) *of course*. Rather, the concessive seems to derive from the modal adverb of certainty. So between the causal connective (content plane) and the concessive (attitudinal-rhetorical plane) comes the relevance hedge use (attitudinal-epistemic plane). The development parallels that of other modal expressions of certainty, such as *certainly*, *obviously*, which are also used in concessive rhetorical structures.

The development of the speaker commitment sense enabled *of course* to occur with almost any type of proposition. *Of course* in PDE occurs almost exclusively as a speaker-oriented S-Adv functioning on the attitudinal discourse plane. It will be argued in the next section that *of course* is continuing to diverge in new directions in PDE.

8.3 *Of course* in PDE

The underlying sense of *of course* is that of naturalness or givenness; it reduces the implied ‘newsworthiness’ of an idea. As mentioned above (8.1), PDE corpus data on *of course* reveal two distinct uses: first, and more commonly, as a relevance hedge suggesting ‘naturally’; second as an emphatic (affirmative or negative), found mainly in spoken dialogue.

8.3.1 Hedge of course

“Hedges measure the word or idea against what is expected” (Tannen 1993: 43). *Of course* counters the expectation that the next idea will be consonant with the previous idea(s) and with the overall context of the discourse. But in certain rhetorical contexts it takes on more specific functions than simply to hedge relevance.

Of course acts as a relevance hedge by reducing the status of the constituent it modifies, and so presenting it as ‘given’. The writer/speaker guards against the addressee’s querying the inclusion of an idea which seems redundant, or misinterpreting an idea as implying some novelty. Every instance of hedging *of course* can be seen to convey reduction of informational status. In (22), this is on coherence grounds.

- (22) The Government has a duty to the taxpayer to ensure that any public money is spent wisely and well. So the Government’s contribution to any project must be soundly based.
 But we are also looking to share risks with the private sector. The PFI will add value when its money will lead to better management. So this requires project risk to be transferred to and controlled by private sector managers.
 Keeping all risks in the public sector will imply no benefits from private finance. But *of course*, we are not looking to transfer all risks. It will not be costeffective [sic] to do so. So we are in the business of risk sharing.
 In a nutshell the private sector must genuinely assume risk ...

(Sp, Young)

With its successive shifts of focus framed by ‘but ...so’, this passage exemplifies what Altenberg calls “contrastive zig-zagging” with “an undercurrent of defensive hedging” (1986: 31). The information that the private sector will not be expected to assume all the risk needs to be presented as non-salient because the default starting-point adopted by the speaker is the expectation that it will assume no risk (the speaker’s message of ‘more risk than expected’ involves contrasting ‘some risk’ with ‘no risk’). The *of course* host is therefore potentially incoherent because there is no ‘all risks’ assumption to counter. *Of course* removes that incoherence by hedging relevance.

Within the general relevance hedge use of *of course*, five common contexts of usage can be identified. They will be termed: concession, background, topic shift, new point (including ‘new item’) and ironical *of course*.

Concession: *Of course* introduces an idea which acts as a counterpoise to an adjacent, contrasting, focused idea. In argumentation contexts, the *of course* host is a conceded argument which is followed by a preferred, counter-argument. Contrast/concession of *of course* performs a backgrounding role too: the dispreferred argument is backgrounded:

- (23) *Of course*, the Sereny-Bell version may be the truth, but we cannot, on the basis of this book, make that assumption (Rev)

Background: The term ‘background’ refers here to the RST relation (see Appendix A) rather than to the information structure sense, though the two are related⁸. The *of course* host is, or begins, a digression from the current topic or line of narrative to provide some background information that aids understanding of something in the main storyline. It often shows a change of tense. The discourse then resumes where it left off:

- (24) A: ... and she wanted a dress like that.^
 B: ^oh yes!^
 A: ... this girl had got the material ... erm ... a wild silk cream ... and er ... Elizabeth made up this dress for her ... *of course* er ... Michael Caine’s daughter i- ... is quite a busty girl, you know
 B: oh I see yeah
 A: and this bride wasn’t
 B: [laugh]
 A: so it didn’t qui- look quite the same .. you know

(Conv)

Topic transition: The *of course* signals that its host is a departure from the current topic. It therefore marks a rhetorical contrast. The *of course* host is a transition idea signalling the end of a topic:

- (25) [current topic= defence diplomacy] Defence is, *of course*, just one component of overall security. NATO must work within the wider framework that includes the United Nations, the OSCE, ... [new topic=non-defence organizations]
 (Sp, Robertson)

Topic introduction: the *of course* host presents a new idea (new to the discourse at that point, that is: it frequently contains known information). This may be a new topic, a new point or a

⁸ RST-BACKGROUND material is always backgrounded (it is expressed in a satellite) but by no means all backgrounded material involves an RST BACKGROUND relation.

new item. It may occupy a long clause complex, or, at a more local level, the *of course* host may be a single NP or AdjP that concludes a list of items in a set.

- (26) I am concerned that such sentences may in some cases be ineffective
 [topic=effectiveness of imprisonment] And then, *of course*, there is the question of cost. The cost of imprisoning defendants is enormous ... [topic = cost]

(Sp, Bingham)

Ironical *of course*. A further feature of *of course* is that it often introduces bad news or a report of some inconvenience, with the sense of irony that comes from confidence in the triumph of adversity over good fortune.

- (27) so I .. I made him swop places as .. er .. with Bryony and *of course* he wou- he wasn't sitting there! .. he was gonna sit in his own seat .. I said you can't cos you keep hitting Richard in the face for nothing

(Conv)

These five contexts of use are not mutually exclusive; they do not represent five identifiable senses of *of course*. Frequently, for example, a digression (background) can also be seen as a topic shift, or a new point may involve a contrast. *Of course* spans perhaps the widest range of usage of the four adverbials in this study. Each context of occurrence highlights particular combinations of implicatures (see table 8.13 below).

8.3.1.1 Concession/Contrast

As seen in section 8.1, OED, Collins and Cobuild dictionaries make no mention of a connective meaning of *of course*. But Quirk et al. (1985) include *of course* among concessive or resultive conjuncts (1985: 635-6); *of course* “commonly expresses superficial agreement with what has preceded while at the same time hinting at a more fundamental disagreement” (Quirk et al 1985: 1469). Concessive *of course* “makes a conjoin resemble a subordinate clause” (1985: 639).

A common pattern is for *of course* to introduce an argument or statement which is then countered by a stronger argument or statement. The *of course* clause concedes a point and sets up an expectation of a stronger, contrasting point. That is, it regularly occurs in the CONCESSION relation with the implicature ‘granted’ or ‘I concede that’. Typically, its host expresses backgrounded

<i>Of course</i>	Concession	[Elaboration/Evidence]	but	Counter-assertion
<i>Much of course</i>	remains to be done.		But	let us not forget how much has already been achieved. (Sp)
<i>Of course,</i>	we must expect to experience for some time the aftershocks of the Soviet empire.	Chechnya is an example.	But nonetheless,	I believe that reform and democracy are becoming entrenched. (Sp)
<i>It is, of course,</i>	good that the Russians are ready to work with the OSCE in Chechnya.		But	that does not erase what has already happened. (Sp)
<i>Of course,</i>	decisions on enlargement are ultimately political.		But,	as the Prime Minister said at Madrid, NATO is a military alliance, it is not a political club. (Sp)
<i>Of course</i>	last year saw the Mexican economic crisis and a slow down in economic growth in some Latin American countries.		But	many Latin American economies were hardly affected at all. (Sp)
<i>Bellow, of course</i>	has nothing to prove as a writer,		but	it would be idle to pretend that this is one of his best books (Rev)
<i>There are, of course</i>	regular haggis-munchers,	just as there are Francis fanatics,	but	such people must be regarded with a raised eyebrow (Rev)
<i>Of course,</i>	nobody is ever ashamed of Bohemian-style poverty;	that is the point of it.	Still,	you would not want to dine with her crowd. (Rev)
<i>[of course]</i>	I mean if they sold their house quickly []	then, you know they'd come down here obviously ...	but	it's not likely that they'd sell their house that quickly (Conv)
<i>Of course</i>	The Boss didn't force immediate change,	either in Fianna Fail or in public life.	But	it raised hares; and some of them are running still. (Rev)

Table 8.8a. Examples of *of course* in the CONCESSION relation with preferred S-N order.

Preferred assertion	Contrastive conjunction	<i>of course</i>	Conceded assertion
This government is about more than redistribution of cash and directly provided services	- although	[<i>of course</i>]	that will [] remain part of our mission. (Sp, Field)

Table 8.8b. Example of *of course* in the CONCESSION relation with N-S order

	DemSam (n=331)	BNCPer (n=491)	Reviews (n=263)	Speeches (n=662)
CONCESSION	3%	24%	37%	37%

Table 8.9. Proportions of *of course* in CONCESSION relations

8.3.1.2 Background

Particularly in spoken conversation, *of course* often introduces a BACKGROUND Satellite. This Satellite increases the ability of the addressee to comprehend an element in the Nucleus. Typically, the Satellite precedes the Nucleus in the following pattern:

Topic A	+ <i>of course</i> + Background [+ Elaboration]	+ Topic A resumed
	S _____	N _____

Topic A	<i>of course</i>	Background	[Elaboration]	Topic A resumed
a. anyway [laugh] Phyllis arrived and afterwards it was .. by then he had calmed down and he was fine and I said wanna read the story now	{ <i>of course</i> }	cos he missed it []	when he decided he couldn't do without his car [laugh] ..	so I said next week perhaps come without your car .. I think I'd won him over by the end but .. it was a bit hairy (Conv)
b. A: and she got on my desk ... and then I said .. oh I'll help you down .. so she sat on my desk to get off ...	and <i>of course</i>	it's not attached B: oh! .. ^no^		A: ^so^ she leant forward and whole thing went flying forward (Conv)
c. And it remains inconceivable to me to consider constructing the kind of international order we want on anything other than the nation state as the basic building block. ...	<i>of course</i>	Defence cooperation has, [], been a notable feature of post-War Europe.	We have pooled our resources and internationalised our policies, to a degree that is without historical precedent. We have, after all, an integrated command structure in NATO. Through NATO, through the new multinational force structures, and in our approach to European defence, we in Britain intend to deepen our cooperation in the defence field.	But the defence of its citizens remains, for all that, the first and most fundamental responsibility of a national government. (Sp, Rifkind)
d. Our commitment will be demonstrated clearly when we host the Second Asia-Europe Meeting (ASEM) in London next April ASEM will be a centrepiece of our Presidency.	<i>Of course,</i>	the first ASEM came as a result of an idea put forward three years ago by Singapore.	Singapore went on to play a prominent role ... We found ourselves working closely with Singapore	Many key initiatives for next April's meeting have come from Singapore ... (Sp, Fatchet)

Table 8.10. *Of course* in the BACKGROUND relation

The host idea may be given or new information, so that a main implicature is either ‘recall that’ or ‘bear in mind that’. In spontaneous conversation, *of course* repeatedly occurs in the context of a narrative, where it introduces an aside or comment, indicating that its host idea is not part of the main storyline (table 8.10). RST schemas for two of the above examples, a and b, are given in figs. 8.8 and 8.9.

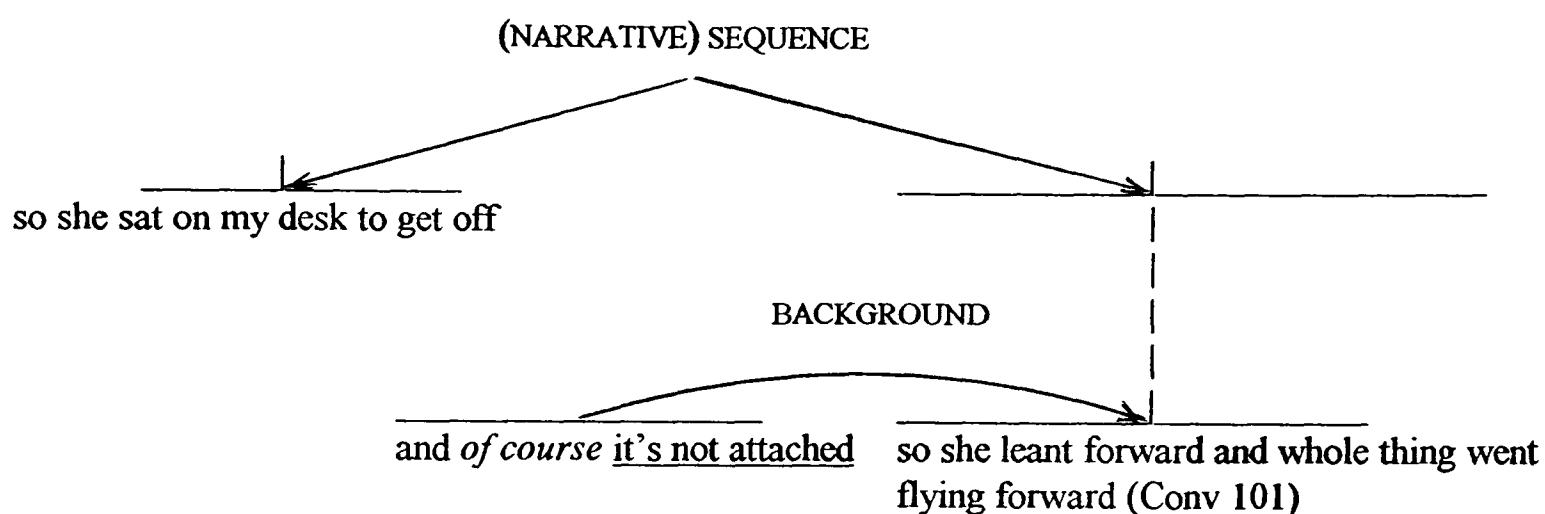


Fig 8.8. *Of course* in BACKGROUND relation: RST schema for example 8.10b

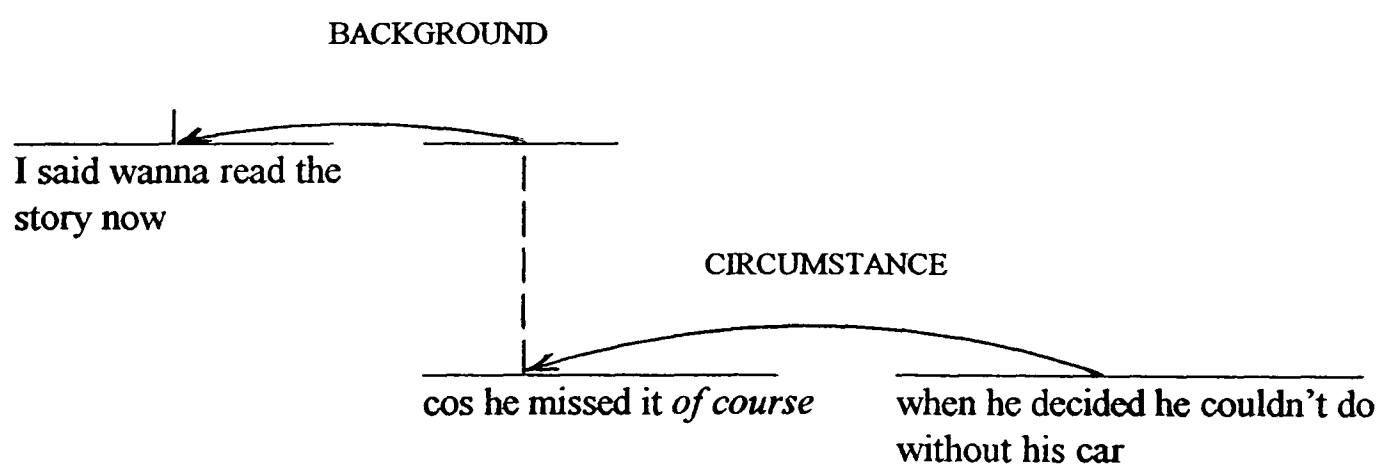


Fig. 8.9. *Of course* in BACKGROUND relation: RST schema for example 8.10a

The *of course* host is a piece of inserted background information (e.g. the relevance to ASEM of Singapore, in example (8.10d)) to assist interpretation of what follows when the main topic is resumed. The interpolated background comment very often involves a change of tense (examples b, c, d in table 8.10). The *of course* refers back: it indicates that a shift of discourse frame is taking place, then the main narrative line or argumentation line resumes. Other implicatures are also present: example (8.10d) involves the introduction of the sub-topic of Singapore, example (8.10c) involves a concession

to defence cooperation. These *of course* tokens therefore also carry topic introduction and concessive implicatures.

8.3.1.3 Topic transition

Of course co-occurs with change in subject-matter at both more global ('new topic') and more local ('new point' and 'new item') levels.⁹ The difference between a new topic and a new point or item is partly a matter of degree: 'new topic' usually refers to subject-matter developed over more than one clause or clause complex. More importantly, a new topic involves counter-expectation (cf. Chafe 1994: 122). Addressees by default interpret a new idea as consonant with the previous one (i.e. 'more of the same') in both subject-matter and stance, unless otherwise marked (v. chapter 2.3). An attitudinal adverbial often provides such a counter-expectation marker, and is therefore rhetorically contrastive. The adverbial can usually be replaced by *but*, but not by *and*. This is the case for topic transition *of course*. Often, especially in speeches, the host of *of course* is a transition unit which signals the end of the current topic and the advent of a new one. A typical form of this pattern is

Topic A N——	+ <i>of course</i> + EVALUATION of topic A S —————	Topic B
----------------	---	---------

Transitions in political speeches and book reviews are exemplified in (28).

- (28) a. ... *Of course* these are not the only areas where the EU is being developed. We are working to reform [X], ... We have to resolve issues over [Y]. And ... we will have important negotiations over [Z]. (Sp, Major)
- b. [current topic = NATO]... because creating a relationship of trust with those outside NATO is just as important as enlarging NATO. THE EU DIMENSION. *Of course*, change in NATO is only part of the story of change in Europe. Every major European and trans-atlantic institution ... is changing too ... I would like to mention one institution in particular. The European Union ... [new topic = the EU] (Sp, Rifkind)
- c. [current topic = population issues] *Of course*, population geography is only one of many valuable approaches to the study of population ... Indeed, multi-method research approaches ... (Rev)

⁹ As pointed out in ch 2.1, 'topic' is a pre-theoretical notion (Brown & Yule 1983: 69-73): topics are identified by addressees (and analysts) intuitively. Topic boundaries are often said to be marked by adverbials in initial position (ch. 2.1). But *of course* occurs just as often in medial position at a topic boundary (55% medial, 42% initial in Speeches, 47% medial, 42% initial in Reviews).

In political speeches *of course* often occurs immediately after a new topic heading (28b). But the *of course* host refers back to the previous topic, and the salient implicature of the *of course* is a rhetorical contrast. In these examples it can be replaced by *but* or *however*. Expressions like *not the only, not just, only part of* background the old topic, by conceding its partial occupancy of a conceptual space, and prepare the ground for the new one. The *of course* is thus both backgrounding and contrastive (cf. *but*): it signals that its host idea is not a further elaboration of the previous topic, but is shifting the discourse up a level in the rhetorical hierarchy (it is contrastive on the speech act plane). Fig. 8.10 shows the rhetorical structure of the passage from which example (28b) is taken, in which *of course* signals the close of the topic NATO..

1. NATO's doctrine and structure are evolving rapidly to meet the changing requirements
2. There should be increasing emphasis
3. on operations like SFOR in Bosnia
4. on the European Security and Defence Identity in NATO
5. and on even more transparency with our partners
6. because creating a relationship of trust with those outside NATO is just as important as enlarging NATO
- THE EU DIMENSION
7. *Of course*, change in NATO is only part of the story of change in Europe. (Sp, Rifkind)

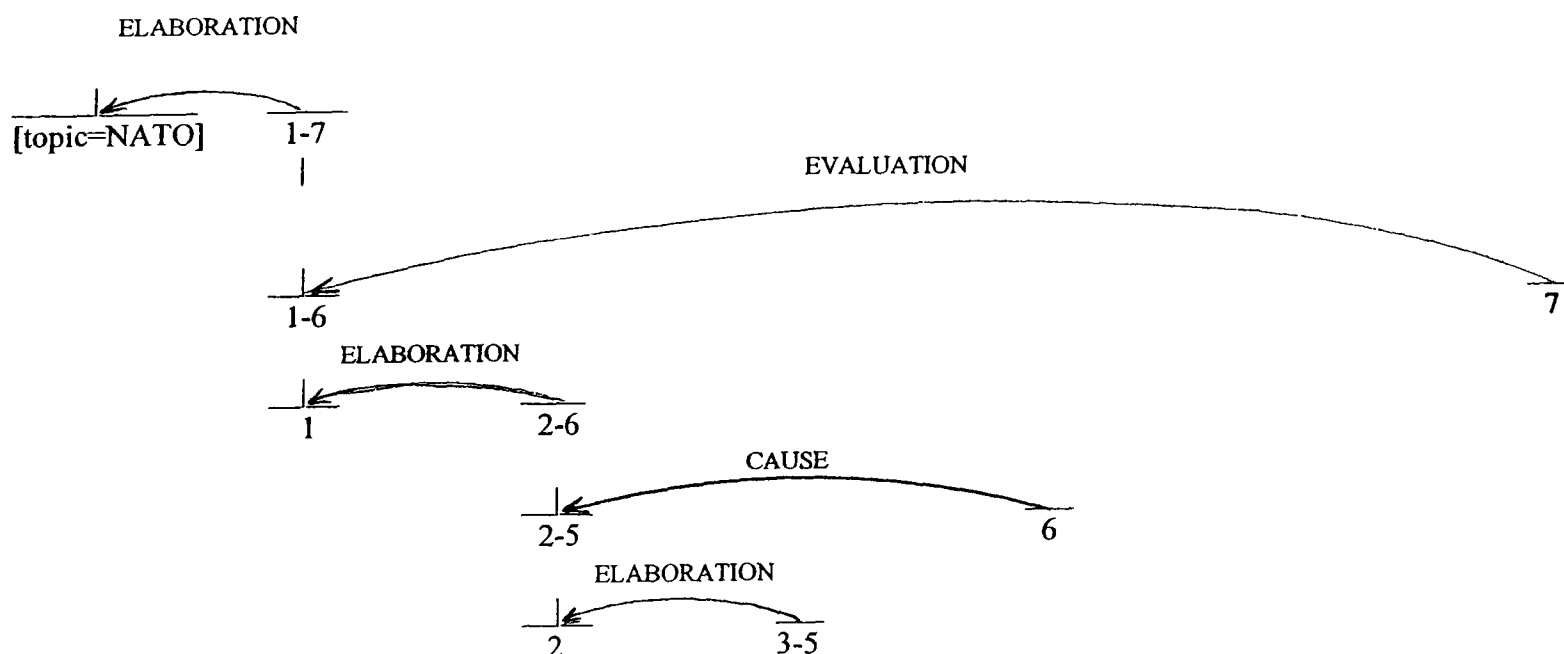


Fig. 8.10. Rhetorical structure of context of example (28b)

Of course in topic-transition may have developed by analogy with contrastive/concessive *of course*. In examples (28) above, a clear contrast is drawn between the old and new topics. So these examples of *of course* arguably fall into the contrastive category. Particularly in argumentation discourse, new topics tend to be contrastive, the topic change being part of a higher-level argumentation strategy that zig-zags between pros and cons. As pointed out in chapter 2.2.1, while

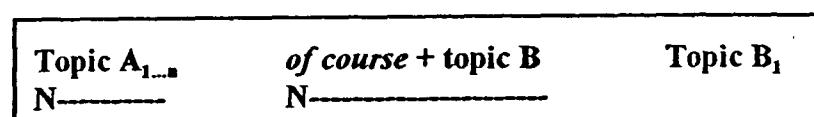
narrative often segments into episodes based on time and place, argument tends to segment into standpoints or claims, with their reasons or grounds.

The examples in (29) below show *of course* with simultaneous contrast and new topic. The rhetorical structure of these passages is interpreted here as SEQUENCE of topics (ultimately ELABORATIONS of a global topic) rather than as CONCESSION or CONTRAST, but this is not the only analysis possible.

- (29) a. Recent academic work .. concludes that aid has become more effective over time. It shows a positive impact on ... The impact has been particularly significant ...
<p> *Of course* we need to be realistic. There will not be a dramatic rise in development assistance ... There will also be setbacks ...
(Sp, Chalker)
- b. ... how little, in practice, international co-operation has developed. It is not particularly surprising that ... There is discussion at the moment on this remains simply an area for discussion. <p> But, *of course*, representatives of competition authorities do talk to each other. I might term this 'informal co-operation' Many useful discussions take place in the margins ...
(Sp, Bloom)

8.3.1.4 Topic introduction

A rather different pattern occurs where some (sub-)topic has been exhausted and the *of course* host is a direct introduction of a new one. It differs from the transition case in that the *of course* host is the new topic and is therefore at the same level in the topic hierarchy as the topic being replaced:



The new topic tends to be in end-focus position, and is then developed in the following clauses.

Examples are:

- (30) a. [current topic=financial legislation] All of these measures ... help considerably in creating the necessary climate ... <p> *Of course*, running alongside all of that, will be the continued efforts of the Industrial Development Board to attract new and inward investment [new topic=inward investment]
(Sp, Mowlam)
- b. [current topic=tax questions] These are the tax questions which impinge on capital availability to business. But there are *of course* other important questions which must be answered if we are to amend our structures to reinforce strengths and address weaknesses. They relate directly to the structure of our capital markets. First, .. [new topic=structure of capital markets]
(Sp, Young)

c. And, *of course*, this catalogue raises an unspoken question. Islamic art scholarship is caught in a quandary. Is it right for ..? Can we say that ...?

(Rev)

d. The authors take the reader ... Other chapters discuss ...and expressions of power in medieval times. <p> There are, *of course*, some difficulties. Although it calls itself an introduction, there is little concession to ... Numerous maps and plans, otherwise well produced, have ...

(Rev)

In examples (30), *of course* occurs at major turns in the discourse. In each case the new topic is put into end-focus position by one means or another (e.g. existentials, inversion). In (30b) the *but* and the *of course* together signal a contrastive topic shift, the *but* providing the contrast and *of course* the topic shift. And in all these cases, *of course* is rhetorically contrastive: it marks a break in discourse continuity, a shift of perspective in the discourse. It indicates that the current context (the ‘topic framework’ or ‘presuppositional situation’) is no longer valid for interpretation of the next idea(s). That is, without *of course*, the text would seem incoherent.

In coherent discourse, a new idea is interpreted as a continuation of the previous discourse in the absence of a marker to the contrary. The following ‘topic shift’ example illustrates the coherence-building function *of course*:

- (31) I am very conscious of the debt which we all owe to the work of the advisory committees. Their untiring efforts ... are neither as well known nor as well appreciated as they should be. I was therefore particularly pleased to see their work so positively endorsed by the ... Committee on judicial appointments. [^]That committee saw some scope for improvement, for example in the arrangement for the selection of members. Its views will be reflected in the revised Directions to Advisory Committees ... (Sp, Irvine)

It is at once apparent that the insertion of *of course* at [^] would remove a non sequitur and restore coherence. The *of course* is clearly connective. It is closer to *but* than it is to *naturally*: if the advisory committees work well, it is surprising, rather than natural, that the Committee should find fault with them. The problem is that *but* overtly focuses and contrasts, whereas *of course* has a concessive and backgrounding flavour, which enables an unexpected or unpopular idea to be presented as obvious and given and coherent with what has gone before. The rest of this speech reveals that Irvine is dissatisfied with the committees, to whom he is speaking. There therefore seems to be a politeness motivation

behind the choice of *of course*. Eventually, however, the contrastive implicatures generated willy-nilly by the context may begin to stick to the *of course*.

A similar type of shift indicated by *of course* is the introduction of a new point or a new item which continues the current discourse perspective. ‘New point’ refers to a topic developed over one discourse segment (usually clause or clause complex), and ‘new item’ refers to the last of two or more ideas that are expressed within one segment (e.g. by NPs or PPs). The pattern is the same as for topic introduction, but lower in the topic hierarchy.

- (32) a. A: and Edmund does a lot of grass cutting and er
 B: (***) tree cutting and
 A: all manner of things ^with^
 B: ^mm^
 A: George .. they’ve got a great com- a great relationship ^going^
 B: ^good .. good^
 A: ^and *of course* he spends Christmas^ (Conv)
- b. A: they’ve [tennis club] got twelve thirteen lawns there
 B: gosh
 A: they haven’t got the membership
 B: shame
 A: and *of course* the upkeep these days is quite high
 B: yes indeed (Conv)
- c. A: ^ooh it was a sight^
 B: ^and the greenhouses^
 A: the trees were racing down the river .. the river was running oh they were huge .. then *of course* when it subsides you’ve got a build up of trees .. that’s left .. you know left on the .. where they’ve blocked
 B: there’s a big tree was wedged on er .. I don’t know if it’s still there .. on Friday (Conv)

In examples (32) above, *of course* co-occurs with an ELABORATION relation. That is, within an overall topic, one or more points are made about that topic, and each of those points may have further elaborations embedded within it. This type of schema is exemplified in fig. 8.11, which shows the rhetorical structure of example (32c) above, where the overall topic of conversation is a recent flood.

ELABORATIONS are like lists of ideas: in the ELABORATION examples in the corpus, *of course* frequently occurs after *and* or *then*. Where it’s the interlocutor (i.e. the person taking up the floor) who adds the point, *of course* also acts as a politeness marker. Example (33) shows *of course* first with a

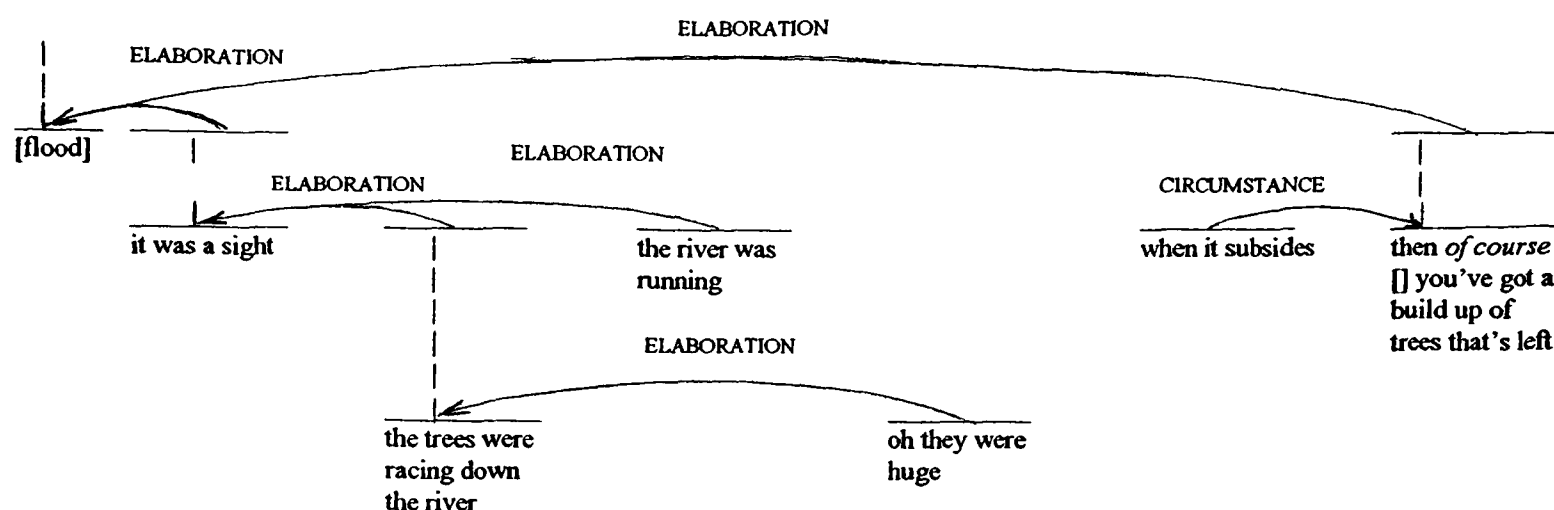


Fig. 8.11. RST diagram of *of course* in ELABORATION relation, example (32c)

topic change (ELABORATION) and second with a new point (embedded ELABORATION) introduced by the interlocutor.

- (33) A: that must've been some time ago
 B: yes
 A: nineteen fifty something I think ...
 B: erm .. *of course* the economic out in Zambia the economic situation is disastrous now ... (***)
 A: and the drought *of course* has made it worse
 B: pardon?
 A: the ^drought^
 B: ^the drought^ yeah .. yeah ... just now the drought seems to be much more widespread now doesn't it (Conv)

(34) shows the introduction of a new point by the interlocutor.

- (34) A: we're their customers
 B: ah .. that's right
 A: that's their living innit?
 B: that's right
 A: without- .. they need er er anyone like ourselves .. well .. er er they'd have nothing to do!
 B: that's right .. absolutely!
 A: ^(***)^
 B: ^and *of course*^ they're not so hard pressed as people are in these big cities!
 A: oh no!
 B: you know .. they have a nicer time of it here ^altogether don't they?
 A: ^oh undoubtedly!^ ^undoubtedly!^ (Conv)

The introduction of a new, final point under a continuing topic is especially common in political speeches, where ideas are frequently presented in threes in traditional hortatory fashion.

- (35) a. Our links across the Atlantic are strong and deep. We have, too, long-standing ties ... with the nations of the Commonwealth, ... And *of course* we are in Europe, ...
(Sp, Lang)
- b. We are already working with you ... on measures to counter the drugs threat. We also want to work towards a satisfactory banana regime ... And *of course* we are at one with you in resisting the objectionable extra-territorial effects of the Helms-Burton legislation.
(Sp, Symons)
- c. There have been too concerns about the profit levels ... Concerns have also been voiced about executive rewards ... And, *of course*, there have been some questions and concerns about the role of the regulator ...
(Sp, Beckett)
- d. a single currency could improve the efficiency ..., a single currency could lead to stronger trade ... Monetary union could ... secure low inflation ... And, *of course*, a single currency within Europe would reduce the costs that businesses and tourists face ...
(Sp, Clarke)

This rhetorical pattern is much rarer in the Reviews corpus. An example is:

- (36) we see her in the plain unaffected descriptions of her lifelong friend Ellen Nussey, ... We see her in the shrewd and lively prose of Harriet Martineau, ... And, *of course*, we have Elizabeth Gaskell's vivid, affectionate accounts ...
(Rev)

Where *of course* introduces the last element of a list, the focusing function seems especially strong. The salience of the element that is placed last is marked by *of course* (37).

- (37) a. ... not only the 15 Members of the EU, but also the many European countries that do not belong to the Union. And, *of course*, the US and Canada.
(Sp, Major)
- b. We are already working with:
*women in local government - in conjunction with Hilary Armstrong;
*women in every region in the country - working closely with the Scottish and Welsh women MPs; and - *of course*,
*women in the trade unions.
WORKING WITH WOMEN IN TRADE UNIONS.
For years we have worked together ...
(Sp, Harman)
- c. Earlier he has dutifully recounted the Scottish walking holiday, the house moves, the hospital, the friends, the adventures of brother George in America, and, *of course*, the politics.
(Rev)

(37b) shows that 'new item' is simultaneously a topic change: the third item in the list becomes the next main topic.

The *of course* in ‘topic introduction’, ‘new point’ and ‘new item’ contexts is associated with the promotion of an idea, whereas *of course* in background and concession contexts is associated with the demotion of its host idea. When *of course* introduces a new topic, point or item, it has a focusing function: the kind of “mild promoting function” that Miller (1984: 35) describes for *like*, *ken* and *see* in a sample of spoken Scottish English, and which draws attention to what the speaker is going to say. The salient implicatures of *of course* in the ‘new topic/point/item’ contexts are to do with discourse organization. *Of course* can also draw attention to the fact that the speaker is going to say something. That is, it can function also on the speech-act plane. This *of course* seems to be becoming an Additive of the same kind as *in fact* (chapter 6.3), meaning ‘and another thing ...’. It functions as a presentational device, providing a pointer to the forthcoming speech-act.

8.3.1.5 Irony

Of course sometimes suggests irony, hyperbole or humour. Although it does not enter into a rhetorical pattern, the use is salient enough to warrant mention. It signals that what is meant is not quite the face value of the utterance, or points to a shared presupposition that ‘Sod’s law’ operates by default. It signals a reversal, and the speaker’s humorous, often self-mocking distancing from its consequences. It can be replaced by *unfortunately* in the examples below, so the negative evaluation is clear: this *of course* operates on the attitudinal-evaluative plane.

Nine per cent of *of course* tokens in the Conversation corpus fit this description, as in (38).

- (38) a. I said oh well .. I .. I’ve been cooking breakfast for Neil .. *of course* by the operation of sod’s law .. Neil suddenly goes off the idea of having breakfast .. so I find myself cooking breakfast for this other lump
(Conv)
- b. A: fancy her saying it ... she was my grammar school friend .. you see
B: mm
A: you brews the ginger ... so I thought well I’m blessed
B: [laugh]
A: and *of course* it was ages before I knew she meant ‘bruise’
B: mm
A: once I started to make it I realized .. ^what it was^
B: ^(***)^ .. yes
A: b-r-u-i-s-e
(Conv)

The foregoing discussion and examples of hedge *of course* show that its usage is more complex than its usual characterization as a relevance hedge suggests. The five patterns account for more than two-thirds of *of course* tokens in Speeches and Reviews (table 8.13 below). Many of the other hedge occurrences simply indicate that a situation or event is natural in the speaker's opinion. But others indicate a further degree of subjectification: they suggest that the host idea itself is the speaker's own viewpoint or interpretation of another event (39).

- (39) a. T.M., like the other English gays he met along the way, was, *of course*, exploiting his relative wealth and European sophistication to prey on the young Arabs who caught his eye. (Rev 246)
 b. ... and he's paying over the odds for this house .. she said but *of course* he's paying for the area (Conv 174)

Of course suggests that *exploiting ...* in (a) and *paying for ...* in (b) are evaluations by the writer or speaker rather than actions of the subjects 'T.M.' and 'he'.

8.3.2 *Emphatic of course*

Emphatic *of course* strengthens illocutionary force. It also occurs in answers to questions, to express agreement and to reinforce an idea (40).

- (40) a. *Of course* bullying the BBC works. That's why we do it. (Rev)
 b. A: he must have alcohol in the body ... ^continuously^
 B: ^*of course*^ he has ... he was drinking Saturday afternoon to Saturday night. Sunday afternoon and the Sunday night ... (Conv)
 c. A: I mean they all say they understand what it's like being on the dole and that .. but they don't
 B: *of course* they don't ... nobody knows (Conv)

Like *yes* and *no*, emphatic *of course* (*not*) emphasizes polarity and is rhematic (is the comment).

Emphatic *of course* is usually connective. Other modal adverbs of certainty such as *naturally*, *certainly*, *obviously*, are connective in the same way when they bear primary stress. They refer back to a previous idea and therefore cannot usually begin a discourse. It is likely that these connective senses evolved directly from answers to questions (cf chapter 2.2 on dialogic residue in written argumentation) because they 'answer' an expressed or implicit epistemic doubt.

Emphatic *of course (not)* expresses contrast with doubt: ‘how could you doubt it?’. In (40c) above, B restates A’s idea with increased force. In (40b), the fact that *has* indicates stronger epistemic commitment than *must have* works in the same direction.

In each case, *of course* is prosodically tonic. With its host it forms a single intonation unit in which it bears the primary stress and its host idea is not asserted but presupposed. There is one assertion, an epistemic one. This is consistent with Chafe’s ‘one new idea per intonation unit’ constraint (Chafe 1980, 1992). In the case of the hedge *of course*, by contrast, the host is asserted. Thus, there are two assertions, a content level one and an attitudinal level one.

In each case of emphatic *of course*, then, there is a restatement (or ellipsis) of the previous idea. The *of course* and its host together form the satellite of the RESTATEMENT relation. As noted above, this use accounts for about a third of the conversation tokens.

8.3.3 Distribution of *of course*

The distribution of *of course* across three text types is shown in fig. 8.12. The frequencies are calculated from the PDE data consisting of 1,436 tokens of *of course*: 254 from Reviews, 680 from Speeches and 502 from Conversation.

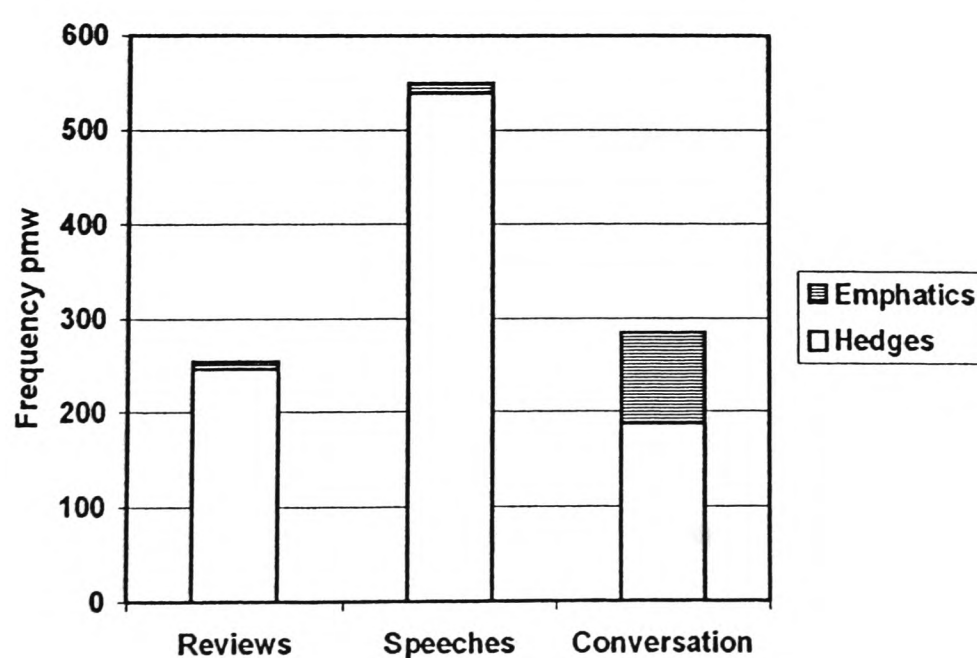


Fig. 8.12. Frequency per million words of *of course* in PDE data

The relative proportions of hedge and emphatic are shown in table 8.11.

	Conv (n=502)	Sp (n=680)	Rev (n=254)
Hedge	66%	98%	97%
Emphatic	34%	2%	3%

Table 8.11. Hedge and emphatic *of course* in PDE data

One main difference between hedge and emphatic is the absence of emphatic occurrences in monologue, since their function is to answer a question or to agree with an interlocutor. The other salient feature is the high density of *of course* in political speeches: 550 occurrences per million words. This compares with only 125 pmw calculated over 30 million words from the periodicals section of the BNC and is the highest frequency of any of the adverbs in this study. *Of course* in concessive and topic-shift contexts is characteristic of argumentative discourse.

The positions of *of course* are shown in table 8.12.

	Initial	Medial	Final
Conversation (n=331)	242 (73%)	14 (4%)	75 (23%)
Speeches (n=668)	297 (44%)	324 (49%)	47 (7%)
Reviews (n= 246)	107 (43%)	88 (36%)	51 (21%)

Table 8.12. Position of hedge *of course* with regard to its host in the three corpora

The preference for initial position in Conversation may indicate that the expression is moving in this direction. On the other hand, planned monologue has its own information structuring techniques. Initial position is overwhelmingly preferred where *of course* introduces a new point. There is no evidence that position correlates with other patterns of use.

The distribution of the contextual uses of hedge *of course* across the four corpora is shown in table 8.13. The high proportion of concessive *of course* in Speeches and Reviews reflects the persuasive purposes of these registers, with their preference for repeated use of a few argumentation patterns. Around a third of occurrences of *of course* are in the context of a concession, diminishing the importance of one idea in favour of another idea that the speaker/writer wishes to promote. In the

	Conv (n=331)	Sp (n=668)	Rev (n=246)
Basic relevance hedge	201 (61%) ¹⁰	176 (26%)	79 (32%)
+ Contrast/concession	9 (3%)	247 (37%)	90 (37%)
+ Background	43 (13%)	41 (6%)	6 (2%)
+ Topic shift	21 (6%)	141 (21%)	40 (16%)
+ New point/end of list	28 (8%)	63 (9%)	26 (11%)
+ Irony/misfortune	29 (9%)	0	5 (2%)

Table 8.13. Distribution of hedge *of course* in PDE data

Periodicals section of the BNC, by comparison, 24% of occurrences of *of course* are in a concession relation.

8.4 Summary

As the dictionaries agree, *of course* is unified in its core sense of ‘naturally’ or ‘as expected’. But the word *naturally* would be a poor substitute for *of course* in many of the above examples. This discussion of *of course* has emphasized the degree to which it is diverging in PDE.

Hedge and emphatic *of course* can be regarded as polysemous. They do not overlap; they have different intonation contours. As shown by example (1), they can give rise to an ambiguity but not to a merged reading.

The various hedge types identified do not amount to polysemy; they are better described as contextual ‘modulation’ (chapter 1.4.5), where different implicatures become more salient depending on the context. *Of course* is perhaps the least polysemous of the four modal adverbials in this study — it is at the lower end of the polysemy spectrum.

Changes in the balance of implicatures due to contextual modulation can lead to stable and established implicatures that may eventually develop into new senses. In the case of *of course*, the extension of meaning seems not to be the result of ad hoc, one-off extensions to new contexts, as is

¹⁰ Includes all those instances where, due to transcription problems, the wider context is not clear. The figures for the Conversation data are therefore not directly comparable with those for the other two text types.

often suggested happens in semantic change (e.g. Taylor, 1995: 285-6) but rather the gradual merging of the interpretations of the expression with the interpretation of the regular contexts.

The data suggest a strong but defeasible connective function for *of course*. The *of course* that marks a topic boundary and the *of course* that introduces a concession in an argumentation have not lost the idea of givenness or expectedness. But they are undoubtedly often connective in that they signal a contrast with the previous discourse and, in the case of argumentation, create an expectancy of contrast with the subsequent discourse. They therefore serve to establish particular relations between two ideas in a text, on the attitudinal-rhetorical plane in the case of concession, on the speech-act plane in the case of new topic/point/item.

Many implicatures are associated with *of course*. This section has suggested that the contexts in which hedge *of course* is found in PDE can be grouped according to which implicatures are the most salient. Table 8.14 summarizes the five clusters.

Contextual modulation	Salient implicatures	Discourse plane	Typical co-occurring RR	Typical status of host	Often with
Concession	I concede that Granted	Attitudinal-rhetorical	CONCESSION	S	contrastive conjunction
Background	Recall that Bear in mind that	Speech-act	BACKGROUND	S	change of tense
New topic -transition	But However	Speech-act	EVALUATION	S	
New topic -introduction; new point/item	And another thing	Speech-act	ELABORATION SEQUENCE	S N	
Irony	Unfortunately	Attitudinal-evaluative	-	-	

Table 8.14. Five common contexts of hedge *of course* in PDE

In every case the core meaning of naturalness is present in some degree. The additional, not mutually exclusive, implicatures are entirely context-dependent but may crystallize eventually.

In the salient implicatures, we have a clue to why *of course* is so popular with politicians: it not only acts as a hedge on illocutionary force (Brown and Levinson 1978: 146ff), but allows the speaker to let through evaluative implicatures while appearing to be objective. Keller (1994: 76-77) recounts

how German *Weib* and *Frau* acquired pejorative connotations, not because they were used in pejorative contexts, but because the more polite *Dame* came to be used for safety, even where politeness was redundant. This process is familiar from the euphemism cycle. In PDE *of course* something similar may be happening. Better-established, overt expressions of contrast, or direct appeals to the addressee, are avoided as being more difficult to retreat from or more face-threatening. With *of course*, one is safe: disagreement or offence are preempted. *Of course* implies shared values between speaker and addressee. The *of course* host is barely asserted, and what is taken for granted is not open to question.

It has been shown that to appreciate the role of *of course* in discourse organization it is essential to look well beyond the host clause; to examine the discourse context at topic and macro-topic levels. This analysis of PDE *of course* has found that it can serve a very similar pragmatic function whether it has a string of ideas in its scope (e.g. new macro-topic), or just one idea embodied in a phrase (e.g. list or background uses).

It has also been seen that *of course* first split into ‘normally’ and ‘naturally’ senses. Then from the ‘naturally’ sense a causal sense split off, which has not survived. Another split into hedging and emphatic senses took place at some point. The PDE data suggest potential for a further split in the hedge sense between the backgrounding/concession use and the topic change/new point use. The former are demoting uses and tend to relate the host to the previous discourse. The latter are promoting uses, and look forward in the discourse, very much like Additive *in fact* (chapter 6.3). Therefore the information structure of the *of course* host in relation to adjacent discourse segments is instrumental in the reweighting of implicatures that leads eventually to new polysemies.

CHAPTER 9

FROM RHETORICAL PATTERN TO POLYSEMY

This study of some English adverbials has traced their slow semantic shifts from content-level adverbials of manner modifying verbs or verb phrases to sentence adverbials expressing speaker attitude. These expressions now have epistemic, evaluative and rhetorical functions. It has been argued that these functions arose not by ad hoc, novel uses of the expressions gaining currency, but by slow, gradual reweighting of implicatures in their existing, typical contexts of use. The new functions are acquired by 'contagion' with other implicatures generated in those contexts.

To some extent it is a lack of close 'fit' between concept and expression that allows for these semantic shifts. Concepts being naturally vastly more fine-grained than linguistic expressions, and pragmatic implicatures having to make up that wide gap, there is perhaps more scope than has traditionally been recognized for the semi-conventionalized meanings that are 'utterance-type meanings'. This study has tried to show, by examining the discourse contexts of a few English modal adverbials, that these semi-conventionalized meanings can be mapped to context types. It has argued that a better understanding of synchronic polysemy can be reached by paying attention to the role of discourse context types in shaping meaning.

The adverbs examined show synchronic polysemy and have developed historically from more concrete senses towards more abstract and subjective ones. From quite different origins, they have developed uses that have converged towards epistemic meanings. They have then picked up connective functions. This type of development is sometimes described as movement along a unidirectional path of semantic change. A more fitting metaphor for these adverbs is perhaps that of a fan opening out, for they evidence layering. Older uses carry on, sometimes for many centuries, and it is possible that the persistence of original features in the new uses is in part due to the continuing influence of these co-existing uses of the same form. The fact that one sense is non-predictable from another, or that a form satisfies linguistic tests for homonymy (ambiguity, zeugma, etc.), does not entail that the senses are represented in the mental

lexicon as unrelated. Conversely, apparent unity of sense does not preclude that the sense is stored redundantly as part of several constructions in which it participates.

The PDE situation of the four adverbs described in chapters 5-8 is summarized and simplified in table 9.1.

Plane Adverb	Content	Attitudinal Epistemic	Evaluative	Rhetorical	Speech Act
<i>after all</i>	<i>after all this</i>	Counter-expectation		Justificative	Justificative
<i>in fact</i>	<i>true in fact</i>	Emphasizer		Antithetical Elaborative	Additive
<i>at least</i>	<i>at least ten</i>	Emphasizer (almost obsolete)	Evaluative	Epistemic-Retreat	
<i>of course</i>	<i>a matter of course</i>	Emphasizer	Irony	Concessive	Continuative Topic shifter

Table 9.1. Summary of PDE uses of *after all*, *in fact*, *at least* and *of course*.

In principle, we may be able to equate a particular reading with a particular domain. In context, language is not used so neatly. For instance, justificative *after all* can often be read as both justifying making an assertion and implying a reason for its content.

The development evidenced for English modal adverbs such as *after all*, *in fact*, etc. is not of the same type as those described in the grammaticalization literature for ‘full verb>aspect,tense, mood marker’ or for ‘lexical noun>case marker’, etc. It is not a typical instance of grammaticalization as the term is traditionally understood. It meets few of the criteria suggested by Lehmann’s model of grammaticalization (table 9.2).:

	syntagmatic		paradigmatic	
phonological	coalescence	?	erosion/loss	×
morphosyntactic	fixedness/ condensation	×	closed-class/ obligatorification	?
functional	idiomaticization	✓	desemanticization, expansion, simplification, loss	✓

Table 9.2. Extent to which the development of *after all*, *in fact*, *at least*, *of course* meets Lehmann’s criteria of grammaticalization

The development does nonetheless conform to the unidirectional movements, observed in other instances of grammaticalization, from more lexical towards more grammatical and from more context-independent towards more context-dependent. The main characteristics are increased subjectification, increased rhetorical function, and in most cases increased frequency. It fits in with three of Traugott's 'correlated diachronic continua' of subjectification in grammaticalization (1995):

Propositional function	→	Discourse function
Objective meaning	→	Subjective meaning
Syntactic subject	→	Speaking subject

It may be that grammaticalization is not an entirely satisfactory category. It is based on observations that the following types of change are unidirectional: semantic change, whereby meanings become more abstract; syntactic change involving structural reanalysis and/or decategorialization; phonological coalescence and erosion. The fact that these three types of change co-occur in the emergence of grammatical formatives, such as the development of tense, aspect and case markers, has led to their status as criteria for a category of linguistic change: grammaticalization. But these types of change also co-occur in other contexts. The formation of affixes from originally lexical material is one step in a cycle of language change that is arguably not driven by grammatical (morpho-syntactic) reanalysis. By overlapping with other types of language change, grammaticalization may obscure these and prevent a maximally parsimonious account of internal language change. Therefore either the criteria of grammaticalization need to be refined to characterize a truly discrete type of change, or they need to be relaxed so that grammaticalization becomes synonymous with internal language change — pragmatic, semantic and morphosyntactic.

Recent emphasis on constructions as the objects of grammaticalization, rather than individual lexemes, together with work in Construction Grammar, which proposes a new, wider definition of grammar, suggest a new, broader definition of grammaticalization. In such a wider definition, the 'discourse constructions' in which modal adverbials occur might perhaps be said to be grammaticalizing. A label is needed to refer to the range of interdependent, co-occurring movements that characterize internal language change. It remains to be seen whether 'grammaticalization' will fill that role.

Meaning changes gradually. The stages identified in the history of these expressions and labelled in the foregoing chapters are therefore a descriptive convenience. It is possible, however, to trace clusters of uses that map to context types, and it is, we have argued, these clusters that reveal the ongoing changes, rather than isolated or special uses.

We cannot account for discourse connective expressions without reference to high-level rhetorical structure; they are not easy to explain if looked at through a sentential lens. We therefore need to have a means of identifying and describing the types of (albeit loose) structure that exist beyond sentence level. The data on these four adverbials points to their PDE uses resulting from semi-conventional rhetorical patterns. These patterns may encourage more holistic, less compositional mental representations, so that the sense of one part of a sequence permeates other parts, such as the transfer of contrastive sense to *in fact* in antithetical relations, or of an evaluative sense to *at least* in contrastive contexts. Rhetorical strategies for achieving communicative goals therefore naturally motivate change. One important strategy is that of allowing something to be understood rather than expressing it overtly; of underspecifying our intent. Examples are the use of *of course* where a concession is made, or the use of *after all* with a truism as a justificative.

The co-occurring semantic, pragmatic and structural changes which these modal adverbs have undergone, and which we have labelled grammaticalization, are part of a unidirectional tendency towards greater subjectification of meaning. It has been argued that this is the result of the expressions absorbing and assimilating some of the contextual meaning regularly generated in their normal usage.

APPENDIX A

Definitions of some rhetorical relations

(from Mann and Thompson 1987)

N = Nucleus; S = Satellite; W = writer; R = reader

Relation name	ANTITHESIS
Type	Presentational
Constraints on N	W has positive regard for the situation presented in N
Constraints on S	none
Constraints on N + S combination:	the situations presented in N and S are in contrast; because of an incompatibility that arises from the contrast, one cannot have positive regard for both the situations presented in N and S; comprehending S and the incompatibility between the situations presented in N and S increases R's positive regard for the situation presented in N
Effect	R's positive regard for N is increased
locus of the effect	N
Relation name	BACKGROUND
Type	Presentational
Constraints on N	R won't comprehend N sufficiently before reading text of S
Constraints on S	none
Constraints on N + S combination:	S increases the ability of R to comprehend an element in N
Effect	R's ability to comprehend N increases
locus of the effect	N
Relation name	CIRCUMSTANCE
Type	Subject-matter
Constraints on N	none
Constraints on S	S presents a situation (not unrealized)
Constraints on N + S combination:	S sets a framework in the subject matter within which R is intended to interpret the situation presented in N
Effect	R recognizes that the situation presented in S provides a framework for interpreting N
locus of the effect	N and S
Relation name	CONCESSION
Type	Presentational
Constraints on N	W has positive regard for the situation presented in N
Constraints on S	W is not claiming that the situation in S doesn't hold
Constraints on N + S combination:	W acknowledges a potential or apparent incompatibility between the situations presented in N and S; W regards the situations presented in N and

	S as compatible; recognizing the compatibility increases R's positive regard for the situation presented in N
Effect	R's positive regard for the situation presented in N is increased
locus of the effect	N and S
Relation name	CONDITION
Type	Subject-matter
Constraints on N	none
Constraints on S	S presents a hypothetical, future or otherwise unrealized situation (relative to the situational context of S)
Constraints on N + S combination:	Realization of the situation presented in N depends on realization of that presented in S
Effect	R recognizes how the realization of the situation presented in N depends on the realization of the situation presented in S
locus of the effect	N and S
Relation name	CONTRAST
Type	Subject-matter
Constraints on N	multi-nuclear
Constraints on the combination of nuclei:	no more than two nuclei; the situations presented in these two nuclei are (a) comprehended as the same in many respects (b) comprehended as differing in a few respects and (c) compared with respect to one or more of these differences
Effect	R recognizes the comparability and the difference(s) yielded by the comparison that is being made
locus of the effect	multiple nuclei
Relation name	ELABORATION
Type	Subject-matter
Constraints on N	none
Constraints on S	none
Constraints on N + S combination:	S presents additional detail about the situation or some element of subject matter which is presented in N or inferentially accessible in N in one or more of the ways listed below. In the list, if N presents the first member of any pair, then S includes the second.
	1. set : member
	2. abstract : instance
	3. whole : part
	4. process : step
	5. object : attribute
	6. generalization : specific
Effect	R recognizes the situation presented in S as providing additional detail for N. R identifies the element of subject matter for which detail is provided.
locus of the effect	N and S
Relation name	EVALUATION
Type	Subject-matter
Constraints on N	none
Constraints on S	none

Constraints on N + S combination:

S relates the situation in N to degree of W's positive regard toward the situation presented in N
 Effect R recognizes that the situation presented in S assesses the situation presented in N and recognizes the value it assigns
 locus of the effect N and S

Relation name EVIDENCE

Type Presentational

Constraints on N R might not believe N to a degree satisfactory to W

Constraints on S The reader believes S or will find it credible

Constraints on N + S combination:

R's comprehending S increases R's belief of N

Effect R's belief of N is increased

locus of the effect N

Relation name INTERPRETATION

Type Subject-matter

Constraints on N none

Constraints on S none

Constraints on N + S combination:

S relates the situation presented in N to a framework of ideas not involved in N itself and not concerned with W's positive regard

Effect R recognizes that S relates the situation presented in N to a framework of ideas not involved in the knowledge presented in N itself

locus of the effect N and S

Relation name JUSTIFY

Type Presentational

Constraints on N none

Constraints on S none

Constraints on N + S combination:

R's comprehending S increases R's readiness to accept W's right to present N

Effect R's readiness to accept W's right to present N is increased

locus of the effect N

Relation name NON-VOLITIONAL CAUSE

Type subject matter

Constraints on N presents a situation that is not a volitional action

Constraints on S none

Constraints on N + S combination

S presents a situation that, by means other than motivating a volitional action caused the situation presented in N; without the presentation of S, R might not know the particular cause of the situation; a presentation of N is more central than S to W's purposes in putting forth the N-S combination

Effect R recognizes the situation presented in S as a cause of the situation presented in N

locus of the effect N and S

Relation name NON-VOLITIONAL RESULT

Type subject matter

Constraints on N none

Constraints on S	presents a situation that is not a volitional action
Constraints on N + S combination	N presents a situation that caused the situation presented in S; presentation of N is more central to W's purposes in putting forth the N-S combination than is the presentation of S.
Effect	R recognizes that the situation presented in N could have caused the situation presented in S
locus of the effect	N and S
Relation name	OTHERWISE
Type	Subject-matter
Constraints on N	presents an unrealized situation
Constraints on S	presents an unrealized situation
Constraints on N + S combination:	realization of the situation presented in N prevents realization of the situation presented in S
Effect	R recognizes the dependency relation of prevention between the realization of the situation presented in N and the realization of the situation presented in S
locus of the effect	N and S
Relation name	RESTATEMENT
Type	Subject-matter
Constraints on N	none
Constraints on S	none
Constraints on N + S combination:	S restates N, where S and N are of comparable bulk
Effect	R recognizes S as a restatement of N
locus of the effect	N and S
Relation name	SEQUENCE
Type	Subject-matter
Constraints on N	multi-nuclear
Constraints on the combination of nuclei:	A succession relationship between the situations is presented in the nuclei
Effect	R recognizes the succession relationship among the nuclei
locus of the effect	multiple nuclei
Relation name	SUMMARY
Type	Subject-matter
Constraints on N	N must be more than one unit
Constraints on S	none
Constraints on N + S combination:	S presents a restatement of the content of N, that is shorter in bulk
Effect	R recognizes S as a shorter restatement of N
locus of the effect	N and S

APPENDIX B

Data sources

A. Synchronic data

Conversation corpus

4.2 million words of demographically-sampled spoken dialogue from the British National Corpus. This corpus contains language from a representative sample of British English speakers selected by region, age, gender and social class.

Speeches corpus

1.4 million words of political speeches given in the British Isles between 1995 and 1999. The texts are the scripts made available to the press via government departments, and do not necessarily exactly match the actual delivery. Most of the speakers are British or Irish government ministers and politicians. The speeches can be assumed to have been written by a mixture of speech-writers and the speakers themselves.

Reviews corpus

1.1 million words of book reviews published during the period 1996-1998, mostly from the following periodicals: The Sunday Times, The Times, The Electronic Telegraph, The Irish Times, British Archaeology, The Independent, The Times Higher Education Supplement.

Periodicals corpus

This is the Periodicals section of the British National Corpus: 27.9 million words of texts from newspapers, journals and magazines.

B. Diachronic data

Full-text sources

The diachronic part of the Helsinki Corpus of English Texts, University of Helsinki. Middle English and Early Modern English sections. 1150-1710

The Corpus of Early Modern English Correspondence sampler, University of Helsinki. 1418-1680.

The Lampeter Corpus, University of Wales at Lampeter. 1641-1739.

Historical texts (listed below)

Non-full-text sources

The Corpus of Middle English Prose and Verse, University of Toronto

Michigan Early Modern English materials, University of Michigan

The Newdigate Letters, transcribed and edited by Philip Hines Jr.

The Early Modern English Dictionaries Database, University of Toronto

Johnson's Dictionary of the English Language, 1st and 4th editions

Oxford English Dictionary, 2nd edition

Middle English Dictionary

List of texts in the Historical Texts database (dates are of publication)

- 1310 The Harley Lyrics. Middle English fiction and poetry
- 1377-1379 Langland, *Piers Plowman*
- 1384-1462 An anthology of Chancery English
- 1385 Chaucer. *Troilus and Criseyde*
- 1387-1394 Chaucer. *The Canterbury Tales*. Middle English
- 1400 Anon. *Sir Gawain and the Green Knight*
- 1400 *Pearl*
- 1425-1496 The Paston letters
- 1440 The York plays
- 1440 The alliterative *Morte Arthure*
- 1460 The Townley plays
- 1485 *Everyman*
- 1523-1525 John Bouchier, Lord Berner. Translation of *The Chronicles of Froissart*
- 1547 The Elizabethan Homilies vol. 1
- 1563-1571 Roger Edgeworth. *Sermons very fruitfull, godly and learned*
- 1563-71 Elizabethan homilies, vol. 2
- 1575 Gammer Gvton's Nedle
- 1580 William Bullokar. *Three pamphlets on grammar*
- 1590 Richard Hooker. A sermon.
- 1591 Anon. *Troublesome reign of King John*
- 1595 Philip Sidney. *The defence of Poesie*
- 1608 Cyril Tournier. *The Revenger's Tragedy*
- 1614 Richard Carew. *The excellency of the english Tongue*
- 1640 Thomas Hobbes. *The elements of law natural and politic*
- 1640 Ben Jonson. *Timber*
- 1659-1667 Samuel Pepys. *The concise Pepys (diaries)*
- 1668 Josiah Childs. *Brief observations concerning trade and interest of money*
- 1689 John Locke. *A letter concerning toleration*
- 1690 John Locke. *Two treatises of Government. Of civil government. Book II*
- 1691 John Locke. *Short observations on a Printed Paper entitled "For encouraging the coining .. money in England"*
- 1691 John Locke. *Some considerations of the consequences of the lowering of interest and the raising of the value of money*
- 1691 Dudley North. *Discourses upon trade*
- 1697 Charles D'Avenant. *An essay on the East India Trade*
- 1704 John Dennis. *The grounds of criticism in poetry*
- 1710 Berkeley. *Principles of human knowledge*
- 1713 Berkeley. *Three dialogues between Hylas and Philonous*
- 1715 Alexander Pope. *Preface to a translation of Homer's Iliad*
- 1739 Review of Hume's Treatise in 'History of the Works of the Learned'
- 1740 Anon. Letter to 'Commonsense' on Hume's *Account of Necessity*
- 1741 David Hume. *On the delicacy of taste and passion*
- 1741 David Hume. *Of superstition and enthusiasm*
- 1741 David Hume. *Of the liberty of the press*
- 1742 David Hume. *Of the rise and progress of the arts and sciences*
- 1742 David Hume. *Of essay writing*
- 1745 A letter from a gentleman to his friend in Edinburgh
- 1749 William Warburton. *Hume on Miracles*

- c1751 David Hume. *An enquiry concerning human understanding*
- 1752 Anthony Ellys, *An Examination of Mr Hume's arguments in his essay on miracles*
- 1752 Thomas Rutherford, *The Credibility of Miracles Defended*
- 1752-1784 Reviews of Hume's writings
- 1753 Anon. *Some late opinions concerning the foundations of morality*
- 1757 David Hume. *Of tragedy*
- 1757 David Hume. *Of the standard of taste*
- 1757 William Warburton, *Remarks on .. the Natural History of Religion*
- 1757 Caleb Fleming, *Postscript on Mr Hume's Natural History of Religion*
- 1758 Thomas Stona, *Remarks on .. the Natural History of Religion*
- 1759 Adam Smith. *The theory of moral sentiments*
- 1759 Edward Young. *Conjectures on original composition*
- 1767 Adam Ferguson. *An essay on the history of civil society*
- 1776 WS Powell, *The insufficiency of Mr Hume's objection to .. Miracles,*
- 1776 Anon. *An account of the life and writings of David Hume*
- 1776 Adam Smith. *The wealth of nations*
- 1777 'Tobias Simple', *Strictures on the 'Account ... Hume'*,
- 1777 Adam Smith. Letter to W. Strahan Esq
- 1777 George Home. Letter to Adam Smith
- 1780 Thomas Hayter. *Remarks on .. Dialogues concerning natural religion*
- 1781 Jeremy Bentham. *An Introduction to the Principles of Morals and Legislation*
- 1792 Mary Wollstonecraft. *A vindication of the rights of women*
- 1793 Anon. *An examinat'n of Hume's Essay on Justice*
- 1794 William Paley. *Hume on miracles*
- 1795-1817 Jane Austen. *Letters.*
- 1797 H. Richter. *On the origin .. of necessary connection*
- c1797 Benjamin Thompson, Count Rumford. *Essay 2 - Of the fundamental principles on which general establishments for the relief of the poor may be f'd in all countries*
- 1798 Thomas Malthus. *An essay on the principle of population*
- 1799 James Caulfield. *David Hume*
- 1800 William Smellie, *The Life of David Hume, Esq*
- 1814 TR Malthus. *The Corn Laws*
- 1815 TR Malthus. *The nature of rent*
- 1817 Samuel Taylor Coleridge. *Biographia Literaria*, vol. 1
- 1828 Robert Peel. *Memoirs*
- 1832 Charles Babbage. *On the economy of machinery and manufactures*
- 1840 JH Newman. *Tract 1: Thoughts on the ministerial commission. Tract 2: The Catholic Church.*
- 1840 JH Newman. *Thoughts .. on alterations to the liturgy*
- 1850 Thomas Carlyle. *Latter-day Pamphlets*
- 1857 Elizabeth Gaskell. *The life of Charlotte Brontë*
- 1859 JS Mill. *On Liberty*
- 1859 Charles Darwin. *The origin of species*
- 1861 Henry Maine. *Ancient law.*
- 1873 W Bagehot. *Lombard Street A description of the money market*
- 1875 James McCosh. *The Scottish Philosophy*
- 1885 James Nasmyth. *Autobiography*

References

- Aarts, B. (1996). The rhetorical adverb *simply* in present-day British English. C. E. Percy, C. F. Meyer, & I. Lancashire, Synchronic corpus linguistics. Papers from the Sixteenth International Conference on English Language Research on Computerized Corpora (ICAME 16). Amsterdam and Atlanta: Rodopi.
- Abraham, W. (1981). Modal particle research. The state of the art. Multilingua Vol. 10, 9-15.
- Abraham, W. (1991a). Discourse particles in German: how does their illocutive force come about? W. Abraham, Discourse particles, 203-252. Amsterdam: John Benjamins.
- Abraham, W. (1991b). The grammaticalization of the German modal particles. E. C. Traugott, & B. Heine, Approaches to grammaticalization Vol. 2, 331-380. Amsterdam: John Benjamins.
- Aijmer, K. (1986). Why is *actually* so popular in spoken English? G. Tottie, & I. Backlund, English in speech and writing, 119-129. Uppsala: Acta Universitatis Upsaliensis.
- Aijmer, K. (1996a). Conversational routines in English. Convention and creativity. London: Longman.
- Aijmer, K. (1996b). Swedish modal particles in a contrastive perspective. Language Sciences Vol. 18 (1-2), 393-427.
- Aijmer, K. (1997). *I think* -- an English modal particle. T. Swan, & O. J. Westvik, Modality in Germanic languages: historical and comparative perspectives, 1-47. Berlin: Mouton de Gruyter.
- Aitchison, J. (2001). Language change: Progress or decay? (3rd ed.) Cambridge: Cambridge University Press.
- Akinnaso, F. N. (1982). On the differences between spoken and written language. Language and Speech Vol. 25, 97-125.
- Algeo, J. (1998). Vocabulary. S. Romaine, The Cambridge history of the English language Vol. IV 1776-1997, 57-91. Cambridge: Cambridge University Press.
- Allerton, D. J., & Cruttenden, A. (1974). English sentence adverbials: their syntax and their intonation in British English. Lingua Vol. 34, 1-29.
- Altenberg, B. (1986). Contrastive linking in spoken and written English. G. Tottie, & I. Bäcklund, English in Speech and Writing. Stockholm: Almqvist and Wiksell.
- Anderson, L. B. (1986). Evidentials, paths of change, and mental maps: typologically regular asymmetries. W. Chafe, & J. Nichols, Evidentiality: The linguistic coding of epistemology, 273-312. Norwood, NJ: Ablex.
- Anscombe, J.-C., & Ducrot, O. (1983). L'argumentation dans la langue. Bruxelles: Mardaga.
- Anttila, R. (1972/1989). Historical and comparative linguistics (2nd ed.). Amsterdam: John Benjamins.
- Ariel, M. (1994). Pragmatic operators. R. E. Asher, & J. M. Y. Simpson, The encyclopedia of language and linguistics, 3250-3253. Oxford: Pergamon.

- Ariel, M. (1998). Discourse markers and form-function correlations. A. H. Jucker, & Y. Ziv, Discourse markers: Description and theory, 223-259. Amsterdam: John Benjamins.
- Arnovick, L. K. (1999). Diachronic pragmatics. Seven case studies in English illocutionary development. Amsterdam: John Benjamins.
- Barlow, M., & Kemmer, S. (1994). A schema-based approach to grammatical description. S. D. Lima, R. Corrigan, & G. K. Iverson, The reality of linguistic rules, 19-42. Amsterdam: John Benjamins.
- Bartsch, R. (1976). The grammar of adverbials: A study in the semantics and syntax of adverbial constructions. Amsterdam: North-Holland.
- Bateman, J. A., & Rondhuis, K. J. (1997). 'Coherence relations': towards a general specification. Discourse Processes Vol. 24, 3-49.
- Bazzanella, C. (1990). Phatic connectives. Journal of Pragmatics Vol. 14, 629-647.
- Beaman, K. (1984). Coordination and subordination revisited: syntactic complexity in spoken and written narrative discourse. D. Tannen, Coherence in Spoken and Written Discourse, 45-80. Norwood, N.J.: Ablex.
- Bellert, I. (1977). On semantic and distributional properties of sentential adverbs. Linguistic Inquiry Vol. 8(2), 337-351.
- Benveniste, É. (1966/1974). Problèmes de linguistique générale. Paris: Gallimard.
- Biber, D. (1988). Variation across Speech and Writing. Cambridge: Cambridge University Press.
- Biber, D., & Finegan, E. (1988). Adverbial stance types in English. Discourse Processes Vol. 11, 1-34.
- Biber, D., & Finegan, E. (1989). Styles of stance in English: lexical and grammatical marking of evidentiality and affect. Text Vol. 9(1), 93-124.
- Bisang, W. (1998). Adverbiality: the view from the Far East. J. van der Auwera, Adverbial constructions in the languages of Europe, 641-812. Berlin: Mouton de Gruyter.
- Blakemore, D. (1987). Semantic constraints on relevance. Oxford: Blackwell.
- Blakemore, D. (1988a). *So* as a constraint on relevance. R. M. Kempson, Mental representations: the interface between language and reality, 183-195. Cambridge: Cambridge University Press.
- Blakemore, D. (1988b). The organization of discourse. F. J. Newmeyer, Linguistics: The Cambridge survey Vol. IV: Language: the socio-cultural context, 229-250. Cambridge: Cambridge University Press.
- Blakemore, D. (1992). Understanding utterances. Oxford: Blackwell.
- Blakemore, D. (1996). Are apposition markers discourse markers? Journal of Linguistics Vol. 32(2), 325-347.
- Bolinger, D. (1972). Degree words. The Hague: Mouton.
- Bréal, M. (1897). Essai de sémantique. Science des significations. Paris: Hachette.
- Brinton, L. (1990). The development of discourse markers in English. J. Fisiak, Historical linguistics and philology, 45-71. Berlin and New York: Mouton de Gruyter.

- Brinton, L. (1996). Pragmatic markers in English: Grammaticalization and discourse functions. Berlin: Mouton de Gruyter.
- Brown, G. (1977). Listening to spoken English. London: Longman.
- Brown, G., & Yule, G. (1983). Discourse analysis. Cambridge: Cambridge University Press.
- Brown, P., & Levinson, S. C. (1978). Politeness. Some universals in language usage. Cambridge: Cambridge University Press.
- Brugman, C. (1986). Sisterhood is more powerful than you thought: scopal adverb placement and illocutionary force. Chicago Linguistic Society Vol. 22 (Papers from the parasession on pragmatics and grammatical theory), 40-53.
- Bühler, K. (1934). Theory of Language. Amsterdam: Benjamins.
- Bybee, J. (1988). The diachronic dimension in explanation. J.A. Hawkins, Explaining language universals, 350-379. Oxford: Blackwell.
- Bybee, J. (1998). Irrealis as a grammatical category. Anthropological Linguistics Vol. 40(2), 257-271.
- Bybee, J. (in press). Mechanisms of change in grammaticalization: the role of frequency. R.D. Janda & B.D. Joseph, Handbook of historical linguistics. Oxford: Blackwell.
- Bybee, J. L., & Dahl, Ö. (1989). The creation of tense and aspect systems in the languages of the world. Studies in Language Vol. 13, 51-103.
- Bybee, J., Perkins, R., & Pagliuca, W. (1994). The evolution of grammar. Tense, aspect and modality in the languages of the world. Chicago and London: University of Chicago Press.
- Cann, R. (1993). Formal semantics. An introduction. Cambridge : Cambridge University Press.
- Chafe, W. (1980). The deployment of consciousness in the production of a narrative. W. Chafe, The pear stories: Cognitive, cultural and linguistic aspects of narrative production. Norwood, N.J.: Ablex.
- Chafe, W. (1982). Integration and involvement in speaking, writing and oral literature. Deborah Tannen, Spoken and written language: exploring orality and literacy, 35-53. Norwood, N.J.: Ablex.
- Chafe, W. (1988). Punctuation and the prosody of written language. Written Communication Vol. 5(4), 395-426.
- Chafe, W. (1990). Looking ahead. Text Vol. 10(1/2), 19-22.
- Chafe, W. (1992). The flow of ideas in a sample of written language. W. C. Mann and S. A. Thompson, Discourse description: Diverse linguistic analyses of a fund-raising text, 267-294. Amsterdam/Philadelphia: John Benjamins.
- Chafe, W. (1994). Discourse, consciousness, and time. Chicago: University of Chicago Press.
- Chafe, W. (1996). Beyond beads on a string and branches in a tree. A. E. Goldberg, Conceptual structure, discourse and language, 49-65. Stanford, CA: CSLI.
- Chafe, W., & Nichols, J. (eds). (1986). Evidentiality: The linguistic coding of epistemology. Norwood, N.J.: Ablex.

- Chierchia, G., & McConnell-Ginet, S. (1992). Meaning and grammar: An introduction to semantics. Cambridge, MA: MIT Press.
- Cinque, G. (1999). Adverbs and functional heads. A cross-linguistic perspective. Oxford: Oxford University Press.
- Claudi, U., & Heine, B. (1986). On the metaphorical base of grammar. Studies in Language Vol. 10(2), 297-335.
- Comrie, B. (1989). Language universals and linguistic typology. Syntax and morphology (2nd ed.). Oxford: Blackwell.
- Conte, M.-E., Petöfi, J. S., & Sözer, E. (1989). Foreward. M.-E. Conte, J. S. Petöfi, & E. Sözer, Text and discourse connectedness: Proceedings of the Conference on Connectivity and Coherence, Urbino, July 16-21 1984, xi-xix. Amsterdam: John Benjamins.
- Corston-Oliver, S. H. (1998). Identifying the linguistic correlates of rhetorical relations. M. Stede, L. Wanner, & E. Hovy, Discourse relations and discourse markers: Proceedings of the Workshop, 15th August 1998, Université de Montréal, 8-14. Montréal: Université de Montréal.
- Cresswell, M. J. (1985). Adverbial modification: interval semantics and its rivals. Dordrecht: Reidel.
- Cristofaro, S. (1998). Grammaticalization and clause linkage strategies: a typological approach with particular reference to Ancient Greek. A. G. Ramat, & P. J. Hopper, The limits of grammaticalization, 59-88. Amsterdam: John Benjamins.
- Croft, W. (1991). Syntactic categories and grammatical relations: The cognitive organization of information. Chicago: University of Chicago Press.
- Croft, W. (1995). Intonation units and grammatical structure. Linguistics Vol. 33, 839-882.
- Croft, W. (1998). Linguistic evidence and mental representations. Cognitive Linguistics Vol. 9(2), 151-173.
- Croft, W. (2000). Explaining language change: An evolutionary approach. London: Longman.
- Cruse, A. (1986). Lexical semantics. Cambridge: Cambridge University Press.
- Cruttenden, A. (1986). Intonation. Cambridge: Cambridge University Press.
- Crystal, D. (1991). A dictionary of linguistics and phonetics (3rd ed.). Oxford: Blackwell.
- Davidson, D. (1982). Moods and performances. A. Margalit, Meaning and use, 1-20. Dordrecht: Reidel.
- Davison, A. (1975). Indirect speech acts and what to do with them. P. Cole, & J. L. Morgan, Syntax and semantics. Vol 3: Speech acts, 143-185. New York: Academic Press.
- de Beaugrande, R. (1980). Text, discourse and process. London: Longman.
- DeLancey, S. (1997). Mirativity: the grammatical marking of unexpected information. Linguistic Typology Vol. 1(1), 33-52.
- Di Eugenio, B., Moore, J., & Paolucci, M. (1997). Learning features that predict cue usage. Proceedings of the 35th Annual Meeting of the ACL and 8th Conference of the European Chapter of the ACL, 80-87.

- Dik, S. C., Hengeveld, K., Vester, E., & Vet, C. (1990). The hierarchical structure of the clause and the typology of adverbial satellites. A. M. Bolkestein, J. Nuyts, & C. Vet, Layers and levels of representation in language theory: A functional view, 25-70. Amsterdam: John Benjamins.
- Dorgeloh, H. (1997). Inversion in modern English. Form and function. Amsterdam: John Benjamins.
- Dry, H. A. (1983). The movement of narrative time. Journal of Literary Semantics Vol. 12, 19-53.
- Du Bois, J. W. (1993). Meaning without intention: lessons from divination. J. H. Hill, & J. T. Irvine, Responsibility and evidence in oral discourse, 48-71. Cambridge: Cambridge University Press.
- Du Bois, J. W. (1994). Pragmatic universals. R. E. Asher, & J. M. Y. Simpson, The encyclopedia of language and linguistics, 3257-3260. Oxford: Pergamon.
- Ducrot, O. (1980). Dire et ne pas dire. Paris: Hermann.
- Duranti, A. (1993). Intentions, self, and responsibility: an essay in Samoan ethnopragmatics. J. H. Hill, & J. T. Irvine, Responsibility and evidence in oral discourse, 24-47. Cambridge: Cambridge University Press.
- Enkvist, N. E. (1972). Old English adverbial *þa* - an action marker? Neuphilologische Mitteilungen Vol. 73, 90-96.
- Ernst, T. B. (1984). Towards an integrated theory of adverb position in English. Bloomington, Indiana: Indiana University Linguistics Club.
- Espinal, M. T. (1991). The representation of disjunct constituents. Language Vol. 67(4), 726-762.
- Ferrara, K. W. (1997). Form and function of the discourse marker *anyway*: implications for discourse analysis. Linguistics Vol. 35, 343-378.
- Fillmore, C. J. (1976). Frame semantics and the nature of language. S.R. Harnad et al., Origins and evolution of language and speech, 20-32. New York: Academy of Sciences.
- Fillmore, C. J., & Atkins, B. T. S. (1992). Toward a frame-based lexicon: The semantics of RISK and its neighbors. A. Lehrer and E. Kittay, Frames, fields and contrasts. New essays in semantic and lexical organization, 75-102. Hillsdale, N.J.: Lawrence Erlbaum.
- Fillmore, C. J., & Atkins, B. T. S. (1994). Starting where the dictionaries stop: the challenge of corpus lexicography. B.T.S. Atkins and A. Zampolli, Computational approaches to the lexicon, 349-393. Oxford: Oxford University Press.
- Fillmore, C. J., Kay, P., & O'Connor, M. C. (1988). Regularity and idiomaticity in grammatical constructions: the case of *let alone*. Language Vol. 64(3), 501-538.
- Finnell, A. (1989). 'Well now and then'. Journal of Pragmatics Vol. 13, 653-656.
- Finnell, A. (1992). The repertoire of topic changers in personal, intimate letters: a diachronic study of Osborne and Woolf. M. Rissanen, O. Ihalainen, T. Nevalainen, & I. Taavitsainen, History of Englishes. New methods and interpretations in historical linguistics, 720-735. Berlin: Mouton de Gruyter.
- Fischer, O. (1997). On the status of grammaticalization and the diachronic dimension in explanation. Transactions of the Philological Society Vol. 95(2), 149-187.

- Fitzmaurice, S. M. (2000). Some remarks on the rhetoric of historical pragmatics. Journal of Historical Pragmatics Vol. 1(1), 1-6.
- Ford, C. E. (2000). The treatment of contrasts in interaction. E. Couper-Kuhlen, & B. Kortmann, Cause - Condition - Concession - Contrast. Cognitive and discourse perspectives, 283-311. Berlin: Mouton de Gruyter.
- Forston, B.W. (in press). An approach to semantic change. R.D. Janda & B.D. Joseph, Handbook of historical linguistics. Oxford: Blackwell.
- Fox, B. A. (1987). Discourse structure and anaphora. Written and conversational English. Cambridge: Cambridge University Press.
- Fraser, B. (1988). Types of English discourse markers. Acta Linguistica Hungarica Vol. 38, 19-33.
- Fraser, B. (1990). An approach to discourse markers. Journal of Pragmatics Vol. 14(3), 383-395.
- Fraser, B. (1996). Pragmatic markers. Pragmatics Vol. 6(2), 167-190.
- Fraser, B. (1998). Discourse markers. J. L. Mey, Concise Encyclopedia of Pragmatics, 256-257. Amsterdam: Elsevier.
- Geeraerts, D. (1983). Reclassifying semantic change. Quaderni di Semantica Vol. 4, 217-240.
- Geertz, C. (1973). The interpretation of cultures. Selected essays. New York: Basic Books.
- Geis, M. L., & Zwicky, A. M. (1971). On invited inferences. Linguistic Inquiry Vol. 2, 561-566.
- Givón, T. (1979). On understanding grammar. New York: Academic Press.
- Givón, T. (1987). Beyond foreground and background. R. S. Tomlin, Coherence and grounding in discourse, 175-188. Amsterdam: John Benjamins.
- Givón, T. (1989). Mind, code and context: Essays in pragmatics. Hillsdale, N.J.: Lawrence Erlbaum.
- Givón, T. (1990). Syntax: A functional typological introduction Vol. 2. Amsterdam: John Benjamins.
- Givón, T. (1995). Functionalism and grammar. Amsterdam: John Benjamins.
- Goldberg, A. E. (1995). Constructions: A construction grammar approach to argument structure. Chicago: University of Chicago Press.
- González-Álvarez, M. D. (1996). Epistemic disjuncts in Early Modern English. International Journal of Corpus Linguistics Vol. 1(2), 219-256.
- Greenbaum, S. (1969). Studies in English adverbial usage. London: Longman.
- Grice, H. P. (1989). Studies in the way of words. Cambridge, MA and London: Harvard University Press.
- Grimes, J. E. (1975). The thread of discourse. The Hague: Mouton.
- Gronemeyer, C. (1999). On deriving complex polysemy: the grammaticalization of *get*. English Language and Linguistics Vol. 3(1), 1-39.
- Grosz, B. J., & Sidner, C. L. (1986). Attention, intentions and the structure of discourse. Computational Linguistics Vol. 12(3), 175-204.

- Haegeman, L. (1985). Subordinating conjunctions and x^l-syntax. Ghent: Studia Germanica Gandensia.
- Haiman, J. (1985). Natural syntax: Iconicity and erosion. Cambridge: Cambridge University Press.
- Halliday, M.A.K. (1991). Corpus studies and probabilistic grammar. K. Aijmer & B. Altenberg, English corpus linguistics. Studies in honour of Jan Svartvik, 30-43. London: Longman.
- Halliday, M. A. K. (1994). An Introduction to Functional Grammar (2nd ed.). London: Edward Arnold.
- Halliday, M. A. K., & Hasan, R. (1976). Cohesion in English. London: Longman.
- Hankamer, J. (1977). Multiple analyses. Charles N. Li, Mechanisms of syntactic change, 583-607. Austin and London: University of Texas Press.
- Hansen, M.-B. M. (1995). Metadiscursive markers in spoken French - the words *bon* and *ben* as examples. Français Moderne Vol. 63(1), 20-41.
- Hansen, M.-B. M. (1997). *Alors* and *donc* in spoken French: a reanalysis. Journal of Pragmatics Vol. 28(2), 153-187.
- Hansen, M.-B. M. (1998a). The function of discourse particles. A study with special reference to spoken standard French. Amsterdam: John Benjamins.
- Hansen, M.-B. M. (1998b). The semantic status of discourse markers. Lingua Vol. 104, 235-260.
- Hanson, K. (1987). On subjectivity and the history of epistemic expressions in English. Chicago Linguistic Society Vol. 23, 132-147.
- Harabagiu, S. M. (1999). From lexical cohesion to textual coherence: a data driven perspective. Journal of Pattern Recognition and Artificial Intelligence Vol. 13(2), 247-265.
- Harris, A. C., & Campbell, L. (1995). Historical syntax in cross-linguistic perspective. Cambridge: Cambridge University Press.
- Hartmann, D. (1998). Particles. J. L. Mey, Concise Encyclopedia of Pragmatics, 657-663. Amsterdam: Elsevier.
- Haspelmath, M. (1998). Does grammaticalization need reanalysis? Studies in Language Vol. 22, 49-85.
- Haspelmath, M. (2000). Why can't we talk to each other? A review article of Newmeyer (1998). Lingua Vol. 110, 235-255.
- Haspelmath, M., & König, E. (1998). Concessive conditionals in the languages of Europe. J. van der Auwera, Adverbial constructions in the languages of Europe, 563-640. Berlin: Mouton de Gruyter.
- Heine, B. (1992). Grammaticalization chains. Studies in Language Vol. 16(2), 335-368.
- Heine, B. (1993) Auxiliaries. Cognitive forces and grammaticalization. Oxford: Oxford University Press.
- Heine, B. (1999). On the role of context in grammaticalization. Paper presented at the International Symposium 'New Reflections in Grammaticalization', Potsdam, June 1999.
- Heine, B. (in press). Grammaticalization. R.D. Janda & B.D. Joseph, Handbook of historical linguistics. Oxford: Blackwell.

- Heine, B., Claudi, U., & Hünnemeyer, F. (1991). Grammaticalization: A conceptual framework. Chicago and London: University of Chicago Press.
- Heine, B., & Reh, M. (1984). Grammaticalization and reanalysis in African languages. Hamburg: Helmut Buske.
- Hellman, C. (1995). The notion of coherence in discourse. Gert Rickheit and Christopher Habel, Focus and coherence in discourse processing, 190-202. Berlin and New York: Walter de Gruyter.
- Heny, F. W. (1973). Sentence and predicate modifiers in English. J. Kimball, Syntax and semantics 2, 217-245. New York: Academic Press.
- Herring, S. (1991). The grammaticalization of rhetorical questions in Tamil. E. Traugott & B. Heine, Approaches to grammaticalization, Vol. 1, 253-285. Amsterdam: John Benjamins.
- Higginbotham, J. (1995). Adverbs. Oxford: University of Oxford, ms.
- Hobbs, J. R. (1985). On the coherence and structure of discourse. Stanford, CA: Centre for the Study of Language and Information. *Technical Report CSLI-85-37*
- Hock, H. H., & Joseph, B. D. (1996). Language history, language change and language relationship. An introduction to historical and comparative linguistics. Berlin: Mouton de Gruyter.
- Hooper, J. B., & Thompson, S. A. (1973). On the applicability of root transformations. Linguistic Inquiry Vol. 4, 465-497.
- Hopper, P. J. (1979). Aspect and foregrounding in discourse. T. Givón, Syntax and semantics Vol. 12: Discourse and syntax, 213-241. New York: Academic Press.
- Hopper, P. J. (1990). Principles of grammaticalization: towards a diachronic typology. W. P. Lehmann, Language typology 1987: Systematic balance in language. Papers from the Linguistic Typology Symposium, Berkeley, 1-3 December 1987., 157-170. Amsterdam: John Benjamins.
- Hopper, P. J. (1991). On some principles of grammaticization. E. C. Traugott and B. Heine, Approaches to grammaticalization Vol. 1, 17-35. Amsterdam and Philadelphia: John Benjamins.
- Hopper, P. J. (1996). Some recent trends in grammaticalization. Annual Review of Anthropology Vol. 25, 217-236.
- Hopper, P. J. (1998). The paradigm at the end of the universe. A. G. Ramat, & P. J. Hopper, The limits of grammaticalization, 147-158. Amsterdam: John Benjamins.
- Hopper, P. J., & Thompson, S. A. (1980). Transitivity in grammar and discourse. Language Vol. 56(1), 251-299.
- Hopper, P. J., & Thompson, S. A. (1984). The discourse basis for lexical categories in universal grammar. Language Vol. 60, 703-752.
- Hopper, P. J., & Traugott, E. C. (1993). Grammaticalization. Cambridge: Cambridge University Press.
- Horn, L. R. (1984). Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. Deborah Schiffrin, Meaning, form and use in context: Linguistic applications (GURT '84), 11-42. Washington: Georgetown University Press.
- Horn, L. R. (1996). Presupposition and implicature. Shalom Lappin, The handbook of contemporary semantic theory, 299-319. Oxford: Blackwell.

- Hovy, E. H. (1993). Automated discourse generation using discourse structure relations. Artificial Intelligence Vol. 63, 341-385.
- Hovy, E. H., & Maier, E. (1995). Parsimonious or profligate? How many and which discourse structure relations? ms.
- Hoye, L. (1997). Adverbs and modality in English. London: Longman.
- Huang, S.-F. (1975). A study of adverbs. The Hague: Mouton.
- Hunston, S. (1993). Projecting a sub-culture: the construction of shared worlds by projecting clauses in two registers. D. Graddol, L. Thompson, & M. Byram, Language and Culture, 98-112. Clevedon: Multilingual Matters.
- Jackendoff, R. (1972). Semantic interpretation in generative grammar. Cambridge, Mass.: MIT Press.
- Jackendoff, R. (1996). Conceptual semantics and cognitive linguistics. Cognitive Linguistics Vol. 7(1), 93-129.
- Jakobson, R. (1956/1990). On language (edited by Linda R. Waugh and Monique Monville-Burston). Cambridge, MA: Harvard University Press.
- Jespersen, O. (1924). The philosophy of grammar. London: Allen & Unwin.
- Johansson, S., & Oksefjell, S. (1996). Towards a unified account of the syntax and semantics of *get*. Jenny Thomas and Mick Short, Using corpora for language research, 57-75. London: Longman.
- Johnson-Laird, P. N. (1983). Mental models. Cambridge: Cambridge University Press.
- Jones, L. B., & Jones, L. K. (1979). Multiple levels of information in discourse. Linda K. Jones, Discourse studies in Mesoamerican languages Vol. 1, 3-28. Callas: Summer Institute of Linguistics.
- Jucker, A. H. (1993). The discourse marker *well*: a relevance-theoretical account. Journal of Pragmatics Vol. 19, 435-452.
- Jucker, A. H. (ed). (1996). Historical pragmatics: Pragmatic developments in the history of English. Amsterdam: John Benjamins.
- Jucker, A. H. (1997). The discourse marker *well* in the history of English. English Language and Linguistics Vol. 1, 91-110.
- Jucker, A. H., & Smith, S. W. (1998). And people just you know like 'wow': discourse markers as negotiating strategies. A. H. Jucker, & Y. Ziv, Discourse markers. Descriptions and theory, 171-202. Amsterdam: John Benjamins.
- Jucker, A. H., & Ziv, Y. (1998). Introduction. A. H. Jucker, & Y. Ziv, Discourse markers. Descriptions and theory, 1-12. Amsterdam: John Benjamins.
- Jucker, A. H., & Ziv, Y. (eds.). (1998). Discourse markers. Descriptions and theory. Amsterdam: John Benjamins.
- Kalokerinos, A. (1995). *Even* - how to make theories with a word. Journal of Pragmatics Vol. 24(1-2), 77-98.
- Kamp, H., & Reyle, U. (1993). From discourse to logic. Dordrecht: Kluwer.

- Kato, K. (1993). Sentence adverbs: their syntax and discourse functions. Hirakata City: Intercultural Research Institute, Kansai Gaido University.
- Kay, P. (1990). Even. Linguistics and Philosophy Vol. 13, 59-111.
- Kay, P. (1992). At least. A. Lehrer and E.F. Kittay, Frames, fields and contrasts: New essays in semantic and lexical organization, 309-331. Hillsdale, N.J.: Lawrence Erlbaum.
- Kay, P. (1997). Construction grammar. P. Kay, Words and the grammar of context, 123-131. Stanford, CA: CSLI.
- Keller, R. (1994). On language change. The invisible hand in language. London: Routledge.
- Kennedy, G. (1998). An introduction to corpus linguistics. London: Longman.
- Kienpointner, M. (1995). On the art of finding arguments: what ancient and modern masters of invention have to tell us about the 'ars inveniendi'. F. H. van Eemeren, R. Grootendorst, J. A. Blair, & C. A. Willard, Perspectives and approaches. Proceedings of the Third ISSA Conference on Argumentation, University of Amsterdam, June 21-24 1994 Vol. 1, 328-337. Amsterdam: International Centre for the Study of Argumentation.
- Knott, A. (1995). An empirical methodology for determining a set of coherence relations. Proceedings of the AAAI Spring Symposium on Empirical Methods in Discourse Interpretation and Generation, Stanford, CA, 1995.
- Knott, A., & Dale, R. (1992). Using linguistic phenomena to motivate a set of rhetorical relations. Edinburgh: University of Edinburgh, Human Communication Research Centre.
- Knott, A., & Mellish, C. (1996). A feature-based account of the relations signalled by sentence and clause connectives. Language and Speech Vol. 39(2-3), 143-183.
- Knott, A., & Sanders, T. (1998). The classification of coherence relations and their linguistic markers: an exploration of two languages. Journal of Pragmatics Vol. 30(2), 135-175.
- Koch, P. (1999). Frame and contiguity. On the cognitive bases of metonymy and certain types of word formation. K.-U. Panther, & G. Radden, Metonymy in language and thought, 139-167. Amsterdam: John Benjamins.
- Koktova, E. (1986). Sentence adverbials in a functional description. Amsterdam: John Benjamins.
- König, E. (1985). On the history of concessive connectives in English: diachronic and synchronic evidence. Lingua Vol. 66, 363-381.
- König, E. (1988). Concessive connectives and concessive sentences: cross-linguistic regularities and pragmatic principles. J.A. Hawkins, Explaining language universals, 145-166. Oxford: Blackwell.
- König, E. (1991). The meaning of focus particles: A comparative perspective. London and New York: Routledge.
- König, E., & Siemund, P. (2000). Causal and concessive clauses: formal and semantic relations. E. Couper-Kuhlen, & B. Kortmann, Cause - condition - concession - contrast. Cognitive and discourse perspectives, 341-360. Berlin: Mouton de Gruyter.
- König, E., & Traugott, E. C. (1988). Pragmatic strengthening and semantic change: the conventionalizing of conversational implicature. W. Hüllen, & R. Schulze, Understanding the lexicon: meaning, sense and world knowledge in lexical semantics, 110-124. Tübingen: Max Niemeyer Verlag.

- Kortmann, B. (1997). Adverbial subordination. A typology and history of adverbial subordinators based on European languages. Berlin: Mouton de Gruyter.
- Kroch, A. S. (1989). Reflexes of grammar in patterns of language change. Language Variation and Change Vol. 1, 199-244.
- Kroon, C. (1996). Discourse markers, discourse structure and functional grammar. J. H. Connolly, Discourse and pragmatics in functional grammar, 17-32. Berlin: Mouton de Gruyter.
- Kurylowicz, J. (1965). The evolution of grammatical categories. Diogenes Vol. 51, 55-71.
- Kytö, M. (1993). Manual to the Diachronic Part of the Helsinki Corpus of English Texts. Coding Conventions and Lists of Source Texts (2nd ed.). Helsinki: University of Helsinki Department of English.
- Labov, W. (1972). The study of language in its social context. P. P. Giglioli, Language and social context. Selected readings, 283-307. Harmondsworth: Penguin.
- Labov, W. (1984). Intensity. D. Schiffrin, Meaning, form and use in context: Linguistic applications, 43-70. Washington, DC: Georgetown University Press.
- Lambrecht, K. (1994). Information structure and sentence form: Topic, focus and the mental representation of discourse referents. Cambridge: Cambridge University Press.
- Langacker, R. W. (1977). Syntactic reanalysis. Charles N. Li, Mechanisms of syntactic change, 57-139. Austin: University of Texas Press.
- Langslow, D. (1998). Latin discourse particles, 'medical Latin' and 'classical Latin'. University of Oxford Working Papers in Linguistics Vol. 3, 20-36.
- Lass, R. (in press). Remarks on (uni)directionality. O. Fischer, A. Rosenback & D. Stein, Pathways of change. Grammaticalization in English. Amsterdam: John Benjamins.
- Leech, G. (1992). Corpora and theories of linguistic performance (with Comments by Bengt Sigurd). Jan Svartvik, Directions in corpus linguistics, 105-125. Berlin: Mouton de Gruyter.
- Lehmann, C. (1982/1995). Thoughts on grammaticalization: A programmatic sketch Vol. 1. Cologne: University of Cologne. Revised version published as Thoughts on grammaticalization. Munich: Lincom Europa.
- Lehmann, C. (1991). Grammaticalization and related changes in contemporary German. Elizabeth Closs Traugott and Bernd Heine, Approaches to grammaticalization Vol. 2, 493-535. Amsterdam and Philadelphia: John Benjamins.
- Lehmann, C. (1992). Word order change by grammaticalization. M. Gerritsen, & D. Stein, Internal and external factors in syntactic change, 395-416. Berlin and New York: Mouton de Gruyter.
- Lenk, U. (1995). Discourse markers and conversational coherence. B. Wårvik, S.-K. Tanskanen, & R. Hiltunen, Organization in discourse. Proceedings from the Turku Conference, 341-352. Turku: University of Turku.
- Levin, B. (1993). English verb classes and alternations. Chicago: Chicago University Press.
- Levinson, S. C. (1983). Pragmatics. Cambridge: Cambridge University Press.

- Levinson, S. C. (1995). Three levels of meaning. F. R. Palmer, Grammar and meaning. Essays in honour of Sir John Lyons, 90-115. Cambridge: Cambridge University Press.
- Levinson, S. C. (2000). Presumptive meanings. The theory of generalized conversational implicature. Cambridge, MA: MIT Press.
- Lewis, D. M. (1999). From modal adverbial to discourse connective: some rhetorical effects in present-day English. J. Verschueren, Pragmatics in 1998. Selected papers from the 6th International Pragmatics Conference Vol. 2, 363-375. Antwerp: International Pragmatics Association.
- Lichtenberk, F. (1991). Semantic change and heterosemy in grammaticalization. Language, Vol. 67, 475-509.
- Locke, J. (1692/1975). An essay concerning human understanding. Oxford: Clarendon Press.
- Longacre, R. E. (1996). The grammar of discourse (2nd ed.). New York: Plenum.
- Lundquist, L. (1989). Modality and text constitution. M.-E. Conte, J. S. Petöfi, & E. Sözer, Text and discourse connectedness: Proceedings of the Conference on Connectivity and Coherence, Urbino, July 16-21 1984, 103-117. Amsterdam: John Benjamins.
- Lyons, J. (1977). Semantics. Vols. 1 & 2. Cambridge: Cambridge University Press.
- Lyons, J. (1995). Linguistic semantics: An introduction. Cambridge: Cambridge University Press.
- Mann, W. C. (1999). An introduction to Rhetorical Structure Theory (RST). Paper available at: <http://www.sil.org/linguistics/RST/rintro99.htm>.
- Mann, W. C., Matthiessen, C. M. I. M., & Thompson, S. A. (1992). Rhetorical structure theory and text analysis. William C. Mann and Sandra A. Thompson, Discourse description: Diverse linguistic analyses of a fund-raising text, 39-78. Amsterdam: John Benjamins.
- Mann, W. C., & Thompson, S. A. (1986). Relational propositions in discourse. Discourse Processes Vol. 9(1), 57-90.
- Mann, W. C., & Thompson, S. A. (1987). Rhetorical structure theory: A theory of text organization ISI Report RS-87-190. Marina del Rey, CA: Information Sciences Institute, University of Southern California.
- Mann, W. C., & Thompson, S. A. (1988). Rhetorical structure theory: toward a functional theory of text organization. Text Vol. 8(3), 243-281.
- Manning, C., & Schütze, H. (1999). Foundations of statistical natural language processing. Cambridge, MA: MIT Press.
- Marcu, D. (1999). Discourse trees are good indicators of importance in text. I. Mani, & M. Maybury, Advances in automatic text summarization, 123-136. Cambridge, MA: MIT Press.
- Martin, J. R. (1992). English Text. System and Structure. Amsterdam/Philadelphia: John Benjamins.
- Maschler, Y. (n.d.). Metalanguaging and discourse markers in bilingual conversations.
- Matthiessen, C. M. I. M., & Thompson, S. A. (1988). The structure of discourse and 'subordination'. John Haiman and Sandra A. Thompson, Clause combining in grammar and discourse, 275-329. Amsterdam: John Benjamins.

- Maynard, S. K. (1993). Discourse modality. Subjectivity, emotion and voice in the Japanese language. Amsterdam: John Benjamins.
- McCawley, J. (1982). Parentheticals and discontinuous constituent structure. Linguistic Inquiry Vol. 13, 91-106.
- McCawley, J. D. (1983). The syntax of some English adverbs. Chicago Linguistic Society Vol. 19, 262-282.
- McCawley, J. D. (1988). The syntactic phenomena of English Vol. 2. Chicago and London: University of Chicago Press.
- McConnell-Ginet, S. (1982). Adverbs and logical form: a linguistically realistic theory. Language Vol. 58, 144-184.
- Meillet, A. (1912/1958). Linguistique historique et linguistique générale. Paris: Champion.
- Michael, I. (1970). English grammatical categories and the tradition to 1800. Cambridge: Cambridge University Press.
- Miller, J. (1984). Discourse patterns in spoken English. Sheffield Working Papers in Language and Linguistics Vol. 1, 10-39.
- Miller, J., & Weinert, R. (1998). Spontaneous spoken language. Syntax and discourse. Oxford: Clarendon Press.
- Moeschler, J. (1989). Pragmatic connectives, argumentative coherence and relevance. Argumentation Vol. 3, 321-339.
- Moore, J. D., & Paris, C. L. (1992). Planning text for advisory dialogues: capturing intentional and rhetorical information. Marina del Rey, CA: Information Sciences Institute Technical Report no. 92-22.
- Moore, J. D., & Pollack, M. E. (1992). A problem for RST: the need for multi-level discourse analyses. Computational Linguistics Vol. 18(4), 537-544.
- Moser, M. G., & Moore, J. D. (1996). Toward a synthesis of two accounts of discourse structure. Computational Linguistics Vol. 22(3), 409-419.
- Moser, M., Moore, J. D., & Glendening, E. (1996). Instructions for coding explanations: identifying segments, relations and minimal units. Pittsburgh, PA: University of Pittsburgh Dept. of Computer Science.
- Nakamura, W. (1997). A cognitive approach to English adverbs. Linguistics Vol. 35, 247-287.
- Nevalainen, T. (1991). But, only, just. Focusing adverbial change in Modern English 1500-1900. Helsinki: Société Néophilologique de Helsinki.
- Nevalainen, T. (1994). Diachronic issues in English adverb derivation. U. Fries, G. Tottie, & P. Schneider, Creating and using English language corpora, 139-147. Amsterdam: Rodopi.
- Nevalainen, T. (1997). The processes of adverb derivation in Late Middle and Early Modern English. M. Rissanen, M. Kytö & K. Heikkonen, Grammaticalization at work. Studies of long-term developments in English, 145-189. Berlin: Mouton de Gruyter.
- Newmeyer, F. J. (1998). Language form and language function. Cambridge, MA: MIT Press.

- Nichols, J., & Timberlake, A. (1991). Grammaticalization as retextualization. E. C. Traugott, & B. Heine, Approaches to grammaticalization Vol. I, 129-146. Amsterdam: John Benjamins.
- Nølke, H. (1989). Modality and polyphony: a study of some French adverbials. M. Herslund, On modality. Papers from meetings and discussions in the Linguistic Circle of Copenhagen. Travaux du Cercle Linguistique de Copenhague Vol.23, 45-63. Copenhagen: The Linguistic Circle of Copenhagen.
- Nuyts, J. (1993a). Epistemic modal adverbs and adjectives and the layered representation of conceptual and linguistic structure. Linguistics Vol. 31, 933-969.
- Nuyts, J. (1993b). From language to conceptualization: the case of epistemic modality. Chicago Linguistic Society Vol. 29 (What we think, what we mean, and how we say it: Papers from the parasession on the correspondence of conceptual, semantic and grammatical representations), 271-286.
- Oates, S. L. (1999). State of the art report on discourse markers and relations. Brighton: Information Technology Research Institute, University of Brighton.
- Oberlander, J., & Moore, J. D. (1999). Cue phrases in discourse: further evidence for the core:contributor distinction. Paper presented at the workshop Levels of Representation in Discourse, University of Edinburgh, July 7-9, 1999.
- Östman, J.-O. (1982). The symbiotic relationship between pragmatic particles and imprompt speech. N. E. Enkvist, Impromptu speech: A symposium. Åbo: Åbo Akademi.
- Östman, J.-O. (1995). Pragmatic particles twenty years after. B. Wårvik, S.-K. Tanskanen, & R. Hiltunen, Organization in discourse. Proceedings from the Turku Conference. Anglicana Turkuensia, 14. Turku: University of Turku.
- Owen, M. (1981). Conversational units and the use of 'Well ...'. P. Werth, Conversation and discourse: Structure and interpretation, 99-116. London: Croom Helm.
- Pagliuca, W. (ed.). (1994). Perspectives on grammaticalization. Amsterdam: John Benjamins.
- Pallander-Collin, M. (1996). The rise and fall of *methinks*. T. Navalainen, & H. Raumoulin-Brunberg, Sociolinguistics and language history: Studies based on the corpus of early English correspondence, 131-149. Amsterdam: Rodopi.
- Pallander-Collin, M. (1997). A medieval case of grammaticalization - *methinks*. M. Rissanen, M. Kytö, & K. Heikkonen, Grammaticalization at work: Studies of long-term developments in English. Berlin: Mouton de Gruyter.
- Palmer, F. R. (1986). Mood and modality. Cambridge: Cambridge University Press.
- Palsgrave, J. (1530/1969). Lesclarcissement de la langue francoyse. Menston, Yorks.: Scholar Press. Facsimile reprint.
- Parsons, T. (1990). Events in the semantics of English: a study in subatomic semantics. Cambridge, MA: MITPress.
- Perelman, C., & Olbrechts-Tyteca, L. (1969). The new rhetoric: A treatise on argumentation. London: University of Notre Dame Press.
- Peterson, P. (1998). On the boundaries of syntax: non-syntagmatic relations. P. Collins, & D. A. Lee, The clause in English. In honour of Rodney Huddleston, 229-250. Amsterdam: John Benjamins.
- Petöfi, J. S. (1983). Introduction. Text Vol. 3(1), 3-9.

- Petöfi, J. S. (ed.) (1988). Text and discourse constitution: Empirical aspects, theoretical approaches. Berlin: Mouton de Gruyter.
- Pin-Ngern Conlon, S. & Evens M. (1994). An adverbial lexicon for natural language processing systems. International Journal of Lexicography Vol. 7(3), 197-221.
- Polanyi, L. (1988). A formal model of discourse structure. Journal of Pragmatics Vol. 12, 601-638.
- Polanyi, L. (1995). The linguistic structure of discourse. Technical Report no. CSLI-96-200. Stanford, CA: Centre for the Study of Language and Information.
- Powell, M. J. (1992). The systematic development of correlated interpersonal and metalinguistic uses in stance adverbs. Cognitive Linguistics Vol. 3, 75-110.
- Power, R., Scott, D., & Doran, C. (n.d.). Generating embedded discourse markers from rhetorical structure. Brighton: University of Brighton.
- Pustejovsky, J. (1995). The generative lexicon. Cambridge, MA: MIT Press.
- Pustejovsky, J., Bergler, S., & Anick, P. (1993). Lexical semantic techniques for corpus analysis. Computational Linguistics Vol. 19(2), 331-358.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). A comprehensive grammar of the English language. London: Longman.
- Radford, A. (1988). Transformational grammar: a first course. Cambridge: Cambridge University Press.
- Radford, A. (1997). Syntactic theory and the structure of English. Cambridge: Cambridge University Press.
- Ramat, A. G., & Hopper, P. J. (1998). Introduction. A. G. Ramat, & P. J. Hopper, The limits of grammaticalization, 1-11. Amsterdam: John Benjamins.
- Ramat, A. G., & Hopper, P. J. (eds) (1998). The limits of grammaticalization. Amsterdam: John Benjamins.
- Ramat, P., & Ricca, D. (1998). Sentence adverbs in the languages of Europe. J. van der Auwera, Adverbial constructions in the languages of Europe, 187-275. Berlin: Mouton de Gruyter.
- Redeker, G. (1990). Ideational and pragmatic markers of discourse structure. Journal of Pragmatics Vol. 14(3), 367-381.
- Redeker, G. (1991). Linguistic markers of discourse structure. [Review of Schiffrin 1987.] Linguistics Vol. 29(6), 1139-1172.
- Reinhart, T. (1983). Point of view in language: the use of parentheticals. G. Rauh, Essays on deixis, 169-194. Tübingen: Günter Narr.
- Rickheit, G., & Habel, C. (eds.) (1995). Focus and coherence in discourse processing. Berlin and New York: Walter de Gruyter.
- Ricoeur, P. (1971). The model of the text: meaningful action considered as a text. Social Research Vol. 38(3), 529-562.
- Risselada, R. (1998). The discourse functions of *sane*: Latin marker of agreement in description, interaction and concession. Journal of Pragmatics Vol. 30(2), 225-244.

- Rizzi, L. (1997). The fine structure of the left periphery. L. Haegeman, Elements of grammar, 281-337. Dordrecht: Kluwer.
- Roberts, I. G. (1993). A formal account of grammaticalization in the history of Romance futures. Folia Linguistica Historica Vol. 13, 219-258.
- Roberts, I. G. (1997). Comparative syntax. London: Arnold.
- Romaine, S., & Lange, D. (1991). The use of *like* as a marker of reported speech and thought: a case of grammaticalization in progress. American Speech Vol. 66, 227-229.
- Rouchota, V. (1998). Procedural meaning and parenthetical discourse markers. A.H. Jucker, & Y. Ziv, Discourse markers. Description and theory, 97-126. Amsterdam: John Benjamins.
- Sanders, T. J. M., & Spooren, W. P. M. (1999). Communicative intentions and coherence relations. W. Bublitz, U. Lenk, & E. Ventola, Coherence in spoken and written discourse: How to create it and how to describe it. Selected papers from the International Workshop on Coherence, Augsburg 24-27 April 1997, 235-250. Amsterdam: John Benjamins.
- Sanders, T. J. M., Spooren, W. P. M., & Noordman, L. G. M. (1992). Toward a taxonomy of coherence relations. Discourse Processes Vol. 15(1), 1-35.
- Sanders, T. J. M., Spooren, W. P. M., & Noordman, L. G. M. (1993). Coherence relations in a cognitive theory of discourse representation. Cognitive Linguistics Vol. 4, 93-133.
- Schachter, P. (1983). Part-of-speech systems. T. Shopen, Language typology and syntactic description. Vol. 1, 3-61. Cambridge: Cambridge University Press.
- Schiffrin, D. (1987). Discourse Markers. Cambridge: Cambridge University Press.
- Schiffrin, D. (1988). Conversation analysis. Frederick J. Newmeyer, Linguistics: The Cambridge survey, 251-276. Cambridge: Cambridge University Press.
- Schiffrin, D. (1992). Anaphoric *then*: aspectual, textual and epistemic meaning. Linguistics Vol. 30(4), 753-792.
- Schiffrin, D. (1994). Approaches to Discourse. Oxford: Blackwell.
- Schourup, L. C. (1983). Common discourse particles in English conversation. Ohio: Ohio State University Working Papers in Linguistics.
- Schreiber, P. A. (1971). Some constraints on the formation of English sentence adverbs. Linguistic Inquiry Vol. 2, 83-101.
- Schwenter, S. (2000). Viewpoints and polysemy: linking adversative and causal meanings of discourse markers. E. Couper-Kuhlen, & B. Kortmann, Cause - condition - concession - contrast. Cognitive and discourse perspectives, 257-281. Berlin: Mouton de Gruyter.
- Schwenter, S., & Traugott, E. C. (1995). The semantic and pragmatic development of substitutive complex prepositions in English. A. Jucker, Historical pragmatics: Pragmatic developments in the history of English, 243-273. Amsterdam: John Benjamins.
- Schwenter, S., & Traugott, E. C. (2000). Invoking scalarity. The development of *in fact*. Journal of Historical Pragmatics Vol. 1(1), 7-25.

- Scott, D. R., & de Souza, C. S. (1990). Getting the message across in RST-based text generation. R. Dale, C. Mellish, & M. Zock, Current research in natural language generation. London: Academic Press.
- Searle, J. R. (1969). Speech acts. An essay in the philosophy of language. Cambridge: Cambridge University Press.
- Siewierska, A. (1991). Functional Grammar. London: Routledge.
- Spanoghe, A.-M. (1996). Mais ou et donc or ni car? Projet pour un dictionnaire des connecteurs discursifs en français, en espagnol et en portugais. Studia Neophilologica Vol. 68, 107-121.
- Sperber, D., & Wilson, D. (1995). Relevance: Communication and cognition (2nd. ed.). Oxford: Blackwell.
- Stein, D. (1985). Discourse markers in Early Modern English. R. Eaton, O. Fischer, W. Koopman, & F. van den Leek, Papers from the 4th International Conference on English Historical Linguistics, 283-302. Amsterdam: John Benjamins.
- Stenström, A.-B. (1990). Lexical items peculiar to spoken discourse. J. Svartvik, The London-Lund corpus of spoken English: Description and research. Lund: Lund University Press.
- Stenström, A.-B. (1995). Some remarks on comment clauses. B. Aarts, & C. F. Meyer, The verb in contemporary English: theory and description, 290-299. Cambridge: Cambridge University Press.
- Stenström, A.-B. (1998). From sentence to discourse: *cos* (*because*) in teenage talk. A. H. Jucker, & Y. Ziv, Discourse markers. Descriptions and theory, 127-146. Amsterdam: John Benjamins.
- Stenström, A.-B., & Andersen, G. (1996). More trends in teenage talk: a corpus-based investigation of the discourse items *cos* and *innit*. C. E. Percy, C. F. Meyer, & I. Lancashire, Synchronic corpus linguistics. Papers from the sixteenth International Conference on English Language Research on Computerized Corpora (ICAME 16), 189-203. Amsterdam and Atlanta: Rodopi.
- Stubbs, M., & Gerbig, A. (1993). Human and inhuman geography: on the computer-assisted analysis of long texts. M. Hoey, Data, description, discourse. Papers in honour of John McH. Sinclair on his sixtieth birthday, 64-85. London: HarperCollins.
- Swan, T. (1988). Sentence adverbials in English: A synchronic and diachronic investigation. Oslo: Novus Forlag.
- Swan, T. (1991). Adverbial shifts: evidence from Norwegian and English. Dieter Kastovsky, Historical English syntax, 409-439. Berlin and New York: Mouton de Gruyter.
- Sweet, H. (1892). A new English grammar, logical and historical. Oxford: Clarendon Press.
- Sweetser, E. (1990). From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure. Cambridge: Cambridge University Press.
- Tabor, W. (1993). The gradual development of degree modifier *sort of* and *kind of*: a corpus proximity model. Chicago Linguistic Society Vol. 29, 451-465.
- Tabor, W., & Traugott, E.C. (1998). Structural scope expansion and grammaticalization. A.G. Ramat, & P.J. Hopper, The limits of grammaticalization, 229-272. Amsterdam: John Benjamins.
- Takahara, P. O. (1998). Pragmatic functions of discourse markers in English and Japanese. J. Verschueren, Pragmatics in 1998. Selected papers from the 6th International Pragmatics Conference Vol. 2, 547-576. Antwerp: International Pragmatics Association.

- Tanaka, K. (1997). The Japanese adverbial *yahari* and *yappari*. R. Carston, & S. Uchida, Relevance theory: applications and implications. Amsterdam: John Benjamins.
- Tannen, D. (ed.) (1982). Spoken and written language: exploring orality and literacy. Norwood, NJ: Ablex.
- Tannen, D. (1993). What's in a frame? Surface evidence for underlying expectations. D. Tannen, Framing in discourse, 14-56. Oxford: Oxford University Press.
- Taylor, J. R. (1995). Linguistic categorization: Prototypes in linguistic theory (2nd ed.). Oxford: Clarendon Press.
- Thomas, T. (1587/1972). Dictionarium Linguae Latinae et Anglicanae. Cambridge/London: T. Thomasii/R. Boyle. Menston, Yorks.: Scholar Press. Facsimile reprint. Electronic version March 1996 by Ian Lancashire, University of Toronto.
- Thomason, R. H., & Stalnaker, R. C. (1973). A semantic theory of adverbs. Linguistic Inquiry Vol. 2, 45-59.
- Thompson, S. A., & Longacre, R. (1985). Adverbial clauses. Timothy Shopen, Language typology and syntactic description. Cambridge: Cambridge University Press.
- Thompson, S. A., & Mulac, A. (1991). A quantitative perspective on the grammaticization of epistemic parentheticals in English. Elizabeth Closs Traugott and Bernd Heine, Approaches to grammaticalization Vol. 2, 313-329. Amsterdam and Philadelphia: John Benjamins.
- Tomlin, R. S., Forrest, L., Pu, M. M., & Kim, M. H. (1997). Discourse semantics. T. A. van Dijk, Discourse as structure and process, 63-111. London: Sage.
- Trask, R. L. (1993). A dictionary of grammatical terms in linguistics. London: Routledge.
- Traugott, E.C. (1986). From polysemy to internal semantic reconstruction. Berkeley Linguistic Society 12, 539-550.
- Traugott, E. C. (1988). Pragmatic strengthening and grammaticalization,. Berkeley Linguistics Society Vol. 14, 406-416.
- Traugott, E. C. (1995a). Subjectification in grammaticalisation. D. Stein and S. Wright, Subjectivity and subjectivisation, 31-54. Cambridge: Cambridge University Press.
- Traugott, E.C. (1995b). The role of the development of discourse markers in a theory of grammaticalization. Paper presented at the Twelfth International Conference on Historical Linguistics, Manchester, August 1995.
- Traugott, E. C. (1997). The discourse connective *after all*: a historical pragmatic account. Paper presented at the Sixteenth International Congress of Linguists, Paris, July 1997.
- Traugott, E. C. (1998). Historical pragmatics and the development of scalar focus particles. Paper presented at the Tenth International Conference on English Historical Linguistics, Manchester 1998.
- Traugott, E.C. (1999). The rhetoric of counter-expectation in semantic change: a study in subjectification. P. Koch & A. Blank, Historical semantics and cognition, 177-196. Berlin: Mouton de Gruyter.
- Traugott, E.C. (in press). Constructions in grammaticalization. R.D. Janda & B.D. Joseph, Handbook of historical linguistics. Oxford: Blackwell.

- Traugott, E. C., & Heine, B. (eds.) (1991). Approaches to grammaticalization Vol. 1: Focus on theoretical and methodological issues. Vol. 2: Focus on types of grammatical markers. Amsterdam and Philadelphia: John Benjamins.
- Traugott, E. C., & Heine, B. (1991). Introduction. Elizabeth Closs Traugott and Bernd Heine, Approaches to grammaticalization, 1-14. Amsterdam and Philadelphia: John Benjamins.
- Traugott, E. C., & König, E. (1991). The semantics-pragmatics of grammaticalization revisited. Elizabeth Closs Traugott and Bernd Heine, Approaches to grammaticalization Vol. 1, 189-218. Amsterdam and Philadelphia: John Benjamins.
- Ullmann, S. (1962). Semantics. An introduction to the science of meaning. Oxford: Basil Blackwell.
- Ullmann, S. (1964). Language and style. Oxford: Basil Blackwell.
- van Dijk, T. A. (1977). Text and context: Explorations in the semantics and pragmatics of discourse. London: Longman.
- van Dijk, T. A. (1979). Pragmatic connectives. Journal of Pragmatics Vol. 3, 447-456.
- van Eemeren, F. H. (1990). The study of argumentation as normative pragmatics. Text Vol. 10(1/2), 37-44.
- van Valin, R. D., & LaPolla, R. J. (1997). Syntax. Structure, meaning and function. Cambridge: Cambridge University Press.
- Vennemann, T. (1975). An explanation of drift. C. N. Li, Word order and word order change, 269-306. Austin, TX: University of Texas Press.
- Verhagen, A. (2000). Concession implies causality, though in some other space. E. Couper-Kuhlen, & B. Kortmann, Cause - condition - concession - contrast. Cognitive and discourse perspectives, 361-380. Berlin: Mouton de Gruyter.
- Warner, R. G. (1985). Discourse connectives in English. New York: Garland Press.
- Watts, R. J. (1988). A relevance-theoretic approach to commentary pragmatic markers: the case of *actually, really* and *basically*. Acta Linguistica Hungarica Vol. 38, 235-260.
- Watts, R. J. (1989). Taking the pitcher to the 'well': native speakers' perception of their use of discourse markers in conversation. Journal of Pragmatics Vol. 13, 210-226.
- Werlich, E. (1976). A Text Grammar of English. Heidelberg: Quelle & Meyer.
- Werth, P. (1984). Focus, coherence and emphasis. London: Croom Helm.
- Willett, T. (1988). A cross-linguistic survey of the grammaticalization of evidentiality. Studies in Language Vol. 12(1), 51-97.
- Wilson, D., & Sperber, D. (1993). Linguistic form and relevance. Lingua Vol. 90, 5-25.
- Ziv, Y. (1985). Parentheticals and functional grammar. A. M. Bolkestein, C. de Groot, & J. L. Mackenzie, Syntax and pragmatics in functional grammar, 181-199. Dordrecht: Foris.

