Constructional Approaches to Language 33

Discourse Structuring Markers in English

Elizabeth Closs Traugott

John Benjamins Publishing Company

Discourse Structuring Markers in English

Constructional Approaches to Language

ISSN 1573-594X

The *Constructional Approaches to Language* (CAL) series brings together research conducted within different constructional models and makes them available to scholars and students working in this and other, related fields.

The topics range from descriptions of grammatical phenomena in different languages to theoretical issues concerning language acquisition, language change, and language use. The foundation of constructional research is provided by the model known as Construction Grammar (including Frame Semantics), initially developed in the 1980s by Charles Fillmore, a founding member of CAL's Advisory Board. The book series publishes studies in which this model is developed in new directions and extended through alternative approaches. Such approaches include cognitive grammar, conceptual semantics, interaction and discourse, as well as typologically motivated alternatives, with implications both for constructional theories and for their applications in related fields such as communication studies, computational linguistics, AI, neurology, psychology, sociology and sociolinguistics, applied linguistics and anthropology.

This peer-reviewed series is committed to innovative research and will include monographs, thematic collections of articles, and introductory textbooks.

For an overview of all books published in this series, please see
benjamins.com/catalog/cal

Volume 33

Discourse Structuring Markers in English
A historical constructionalist perspective on pragmatics
by Elizabeth Closs Traugott

Discourse Structuring Markers in English

A historical constructionalist perspective on pragmatics

Elizabeth Closs Traugott
Stanford University

John Benjamins Publishing Company
Amsterdam / Philadelphia

The paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

DOI 10.1075/cal.33

Cataloging-in-Publication Data available from Library of Congress:
LCCN 2021059064

ISBN 978 90 272 1091 3 (HB)
ISBN 978 90 272 5792 5 (E-BOOK)

John Benjamins Publishing Company · https://benjamins.com

Table of contents

Part II. Case studies

List of figures

List of tables

List of abbreviations

1	first person
2	second person
3	third person
ACC	accusative
AP	adjective phrase
AD/R	addressee/reader
ADV	adverbial
AUX	auxiliary verb
CF	communicative function
CircAdv	circumstance adverbial
Cxn	construction
DAT	dative
D1	discourse segment 1
D2	discourse segment 2
DEM	demonstrative
DF	discourse function
DM	Discourse Marker (highly pragmatic Connector)
DSM	Discourse Structuring Marker
1DSM	minimally pragmatic Discourse Structuring Marker
E	expression
F	form
IITSC	Invited Inferencing Theory of Semantic Change
M	meaning
MORPH	morphology
N	noun
NP	noun phrase
OBJ	object
PHON	phonology
PIE	Proto Indo-European
PL	plural
PRAG	pragmatics
Prep	preposition
SEM	semantics
SG	singular
SP/W	speaker/writer
SUBJ	subject
SYN	syntax
V	verb
VP	verb phrase

Preface and acknowledgments

The idea of writing this book developed as I delved into the histories of more and more discourse markers in the narrow sense of expressions that signal how the speaker or writer intends the relationship between two discourse segments to be understood. For example, is it a relationship of elaboration (*also, after all* in some of its uses), of contrast (*but, all the same*), of digression (*by the way, incidentally*), or of return to a prior topic (*back to my point*)? In the process I realized how little work has been done in recent cognitive linguistics and especially in construction grammar on pragmatics. So I set out to respond to a call for reflection on how pragmatics and context may best be accounted for in constructional terms. This book aims to contribute to the growing field of Diachronic Construction Grammar, with focus on historical pragmatics. It is written for advanced students and for researchers interested in construction grammar, historical linguistics, and functional approaches to linguistic pragmatics.

Many people inspired me to write this book by inviting me to conferences or challenging me to think further about the theoretical issues the data raise, among them Oliver Ehmer and Malte Rosemeyer, Benjamin Fagard and Michel Charolles, Bruce Fraser, Maj-Britt Mosegaard Hansen, Yu Lin, Miguel Oliveira, Noriko O. Onodera and Salvador Pons Bordería. Laurel J. Brinton's work on the rise of various kinds of pragmatic markers, many of them epistemic, has been a guiding light for many years in my pursuit of historical pragmatics and now of the histories of discourse structuring markers. Bernd Heine and his colleagues Gunther Kaltenböck, Tania Kuteva and Haiping Long have suggested radically different ways of thinking about discourse markers and helped shape my thinking. Yang Kun inspired me to probe the underpinnings of my approach. Fuyin (Thomas) Li invited me to present ten lectures at the China International Forum on Cognitive Linguistics (CICFL) 2020 (presented virtually in 2021). This and the lively chats afterwards greatly sharpened my ideas. These lectures are appearing in the Ten Lectures series at Brill, entitled *Ten lectures on a Diachronic Constructionalist Approach to Discourse Structuring Markers*.

Special thanks to Graeme Trousdale who, in virtual meetings during the COVID-19 pandemic, tirelessly discussed almost every theoretical aspect of this work and helped guide my thinking in new directions. I am also grateful to Jan-Ola Östman for welcoming this book into the CAL family, and to two anonymous reviewers for detailed and valuable comments. Many thanks too to the editorial staff at John Benjamins, especially Patricia Leplae, for her exemplary patience and guidance.

ECT
December 2021

CHAPTER 1

Introduction and overview

1.1 Introduction

This book is a contribution to the growing field of historical cognitive linguistics, in particular historical constructionalist pragmatics.

Cognitive linguistics is a perspective on the study of language that has as its subject-matter language-users' "knowledge of the full range of linguistic conventions" (Langacker 1987: 494; see also Goldberg 2003: 223). There are several broadly compatible approaches to cognitive linguistics (Geeraerts and Cuyckens 2007: 3). The one adopted here is construction grammar, of which there are also several varieties (Hoffmann and Trousdale 2013). The constructionalist approach on which the present volume is based is that of Goldberg (1995, 2006), and to some extent that of Croft (2001).

It is a basic principle of cognitive linguistics and its various subtypes, e.g. construction grammar, that there is no fundamental modular distinction between semantics and pragmatics. This position is especially clear in work by Fillmore and his colleagues during the 1980s and 1990s. Among Fillmore's many linguistic interests was his concern for pragmatics. For example, the *let alone* construction, as in *Fred won't eat shrimp, let alone squid*, is constrained in the following way: there must be some relevance between *let alone X* and what precedes. More specifically there is a scalar relation between the proposition (in this case *eat shrimp*) and *let alone X* (in this case *eat squid*): shrimp and squid are both a type of seafood. Squid is presented as of even higher on a scale (for example, of distastefulness) to Fred than shrimp (Fillmore et al. 1988). Expressions of the type *What's this fly doing in my soup?* are usually not intended as questions (though they could be in some contexts) but as requests for explanation of some incongruity and inappropriateness (Kay and Fillmore 1999). Such constraints on the pragmatics of expressions are conventionalized aspects of a speaker's knowledge of language because speakers share them and replicate them in use.

In other strands of cognitive linguistics, however, less attention has been paid to pragmatics or markers of pragmatic meanings such as discourse markers (DMs). There is surprisingly little on the role of pragmatics in ground-breaking works on cognitive grammar such as Langacker's *Foundations of Cognitive Grammar* (1987, 1991) and Goldberg's *Constructions: A Construction Grammar Approach to*

Argument Structure (1995) and *Constructions at Work: The Nature of Generalization in Language* (2006), which lay the foundations of Goldberg's model of construction grammar. DMs have in fact been regarded as marginal in many theoretical approaches because they are not integrated into the syntactic structure of the clause. For example, Huddleston, Payne and Peterson (2002), treat as "supplements" interjections (e.g. *Oh*) and what Brinton (2008) calls comment clauses (e.g. *I mean*). However, DMs are essential, not supplementary, to communication, whether spoken or written. Furthermore, to quote Joseph (1997: 197) "despite the fact that the core has always attracted more intense attention among linguists than the periphery has, nonetheless there is much of value in considering the margins and the peripheral aspects of language".

"Knowledge of the full range of linguistic conventions" has for the most part been interpreted in cognitive linguistics as accounting for the use and understanding of standard sentence types like ditransitive *Kim gave Sue a book* and resultative *He wiped the table clean* (Goldberg 1995), and of event structures like *John swam into the cave* in which manner of motion (swimming) is lexically merged with motion (compare *John entered the cave swimming*, in which manner and motion are lexically separate, a more typical configuration in Spanish) (Talmy 1985, 2000). Langacker theorizes linguistic ways of expressing alternative construals of and perspectives on the same event or situation. For example, while *enter* is a process, *entrance* "construes the component states of the corresponding process as a set of interconnected entities and imposes on them the collective profile of a thing" (Langacker 1987: 247).[1] Knowledge of language is, however, not limited to knowledge of construals of events and what options are available for expressing them. It also includes knowledge of pragmatics, as is coming to be recognized. Recent important contributions to the study of the role of pragmatics in cognitive linguistics include a handbook on *Cognitive Pragmatics* edited by Schmid (2012) and Fischer's (2017) handbook article on "Cognitive linguistics and pragmatics".

An early exploration of the role of pragmatics in construction grammar is a paper on pragmatic markers in Czech and Solv (a Finland-Swedish dialect) by Fried and Östman (2005). In this paper, pragmatics is understood in the broader sense of cultural as well as more strictly internal linguistic pragmatics. Two special issues of *Constructions and Frames* focus on aspects of the role of pragmatics in construction grammar: Fried and Nikiforidou (2015) explores interactions of constructions with register and genre, and Finkbeiner (2019a) explores pragmatics and semantics from a primarily language internal perspective. Finkbeiner (2019a) includes Ariel (2019) on the pragmatics of *or* from a constructionalist perspective (see also Ariel and

1. See Verhagen (2007) for an overview of work in cognitive linguistics on construal and perspectivization.

Mauri 2019). To my knowledge, Ariel's papers are among the first theoretical constructionalist synchronic papers on an English pragmatic Connector. Finkbeiner (2019b) seeks to find principled ways of infusing more pragmatics into constructional thinking (see also Schmid 2016; Cappelle 2017).

In what follows, I outline the goals of the book (Section 1.2) and introduce the empirical domain to be analyzed: Discourse Structuring Markers in Section 1.3. Section 1.4 provides an overview of the chapters. Data sources and method are discussed in 1.5. Section 1.6 summarizes.

1.2 Goals

I have three intertwined goals in this book. My primary goal is to take up the challenge to find ways of infusing more pragmatics into constructionalist thinking using pragmatic Connectors as my data.[2] The empirical domain of investigation is the development of constructions that can be used to do discourse structuring work. My goal in studying these is to develop a systematic and nuanced approach to the history of Discourse Structuring Markers from the perspective of Diachronic Construction Grammar (see e.g. Traugott and Trousdale 2013; Barðdal et al. 2015). The third goal is to refine aspects of Diachronic Construction Grammar in the process. Studying Discourse Structuring Markers requires a view of what discourse is. For present purposes I characterize it as consisting of sequences of utterances that speakers/writers intend as a coherent whole and that addressees/readers attempt to understand as a coherent whole.

1.3 The empirical domain: Discourse Structuring Markers

Constructions such as *after all, but, by the way*, *in addition, instead*, and *now* have been referred to by a number of different terms, as will be discussed briefly below. I will be using the term "discourse structuring marker" (DSM). I refer to "discourse structur-ing markers" to highlight the fact that they are used not merely to reflect intended relationships but to signal and even shape such relationships. The term is a partial adaptation of Fraser's (2009a: 893) term 'discourse structure markers'. Whereas Fraser cites *first, in summary, I add* as examples of discourse structure markers, I am using the term "discourse structuring markers" to cover a far larger set of markers, in fact all Connectors that he calls "DMs".

---

2. Some non-pragmatic markers that are discussed in Chapter 9 highlight how pragmatic the ones that are the main topic of this book are.

DSMs are Connectors that allow the speaker/writer (SP/W) to signal what relationship they wish the addressee/reader (AD/R) to deduce from the linking of discourse segments in a non-subordinate way. *After all* can be used to justify what has just been said in anticipation of the AD/R questioning why it was said, and *but* can be used as a contrastive marker inviting AD/R to cancel potential expectations arising out of segment 1. Examples from the spoken component of the *Corpus of Contemporary American English* (COCA) are:

(1) a. You know, for a lot of folks, sports, and in particular football, can be like a religion. **After all**, there's a Sunday ritual and a fervent belief in the powers of the home team. (2000 *CNN – SunMorn* [COCA])

b. Countries separate, they break relations, they – leaders don't speak to each other. **But** there's always a coming back together. (1990 *ABC_Nightline* [COCA])

In (1a) the addressee (the TV anchor) and listeners generally are invited to cancel any question they might have about the statement that sports can be like a religion, and in (1b) they are invited to cancel a conclusion from hearing segment 1 such as 'Countries will continue to break relations'. Without these cues to interpreting the relationships between what precedes and what follows, the hearer or reader would have to work harder to make the connection. While DSMs have been used throughout the historical record of English and before, particular markers like *after all* and *but* were created in the course of the last millennium.

I posit a continuum of DSMs. DSMs at one pole of the continuum are primarily contentful and monofunctional expressions (e.g. *in addition*). At the other pole they are highly multifunctional expressions (e.g. *after all*), and in between are markers like *incidentally*, as in Figure 1.1:

Figure 1.1 The semantics-pragmatics continuum of Discourse Structuring Markers

Largely contentful/semantic		**Largely non-contentful/pragmatic**
monofunctional		multifunctional
further, instead, parenthetically	*incidentally*	*also, by the way*

The semantics-pragmatics continuum in Figure 1.1 is synchronic. It shows that similar connective functions can be served by markers that are pragmatic to varying degrees. Among markers of "digression" compare relatively contentful *parenthetically* with significantly less contentful *by the way*. As will be shown in Chapter 8.3, *incidentally* is used in a more pragmatic way than *parenthetically* but less so than *by the way*. In the next section I discuss how the continuum also has a historical dimension as DSMs arise out of partially contentful monofunctional expressions, many of them adverbials in English and in other languages as well.

Fraser (1996 and elsewhere) has used the term DM in a restrictive sense to refer to pragmatic markers that connect discourse segment 1 and discourse segment 2, but as is explained immediately below, I further restrict it to multifunctional expressions. The term DM was originally used by Schiffrin (1987) for a wider range of markers including not only Connectors like *and, but* and *now* but also markers like *y'know*. Like Fraser, I use the term pragmatic markers (PMs) for this larger set and consider DMs to be a subset of the larger class of PMs.

Because the empirical domain of investigation is the development of constructions that can be used to do discourse structuring work, I use the term "Discourse Structuring Markers" (DSMs) to refer to expressions on this continuum. By contrast, Fraser (1996 and elsewhere) calls both relatively monofunctional and multifunctional markers "discourse markers" (DMs). This obscures not only the kinds of knowledge that language users have about how to use these markers but also the historical trajectory of DSMs, which, I will show, always start out on the contentful, monofunctional pole and only sometimes are used on the multifunctional pole. DMs on this view are a multifunctional, pragmatic subset of DSMs. I call DSMs that are largely contentful and monofunctional 1DSMs. Those that are multifunctional are DMs.

DSMs are a subset of a larger set of Pragmatic Markers. Calling them all Discourse Markers, Schiffrin (1987) distinguished three kinds: social markers like *please*, what she called "epistemic markers" like *y'know*, and textual markers like *and, but*.[3] The hierarchy of Pragmatic Markers is represented in Figure 1.2, using Schiffrin's three types of Discourse Markers. Pragmatic Markers that are not DSMs are here called Discourse Particles, following Aijmer (2002). The hierarchy presents prototypical distinctions only. In keeping with the concept of the semantics-pragmatics continuum, the distinctions are dynamic and flexible, with marginal as well as central members.

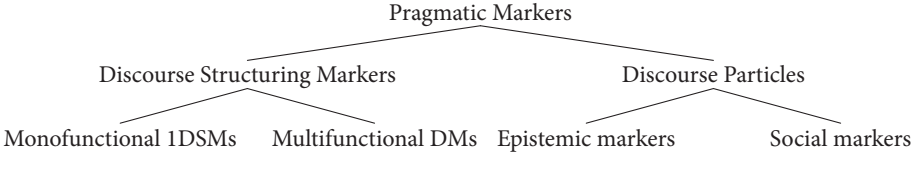

Figure 1.2 Hierarchy of Pragmatic Markers (drawing on Schiffrin 1987)

My concern in this book is with members of the left hand branch of the hierarchy.

3. Östman (1981: 17) suggests that *y'know* is used to seek cooperation by the Addressee, and/or to check the extent of mutual background knowledge. On this analysis it is not epistemic.

1.4 Overview of the book

The book is organized in three parts: Foundations, Case Studies, and Three Open Issues for Historical Construction Grammar.

Part I, Foundations, introduces the three main sets of concepts that underpin the discussion: (a) cognitive linguistics, and within it, construction grammar of the type associated with Goldberg (e.g. 1995, 2006) and to some extent Croft (2001) (Chapter 2); (b) a constructionist view of language change (Chapter 3); and (c) background on Discourse Structuring Markers and how they arise (Chapter 4). In addition, some alternative approaches to the rise of DMs are discussed briefly in Chapter 5.

Chapter 2 lays out key tenets in current work on cognitive linguistics, especially in the model known as construction grammar as developed by Goldberg. The most important of these for the present book are that (i) knowledge of language consists of form-meaning pairs (constructions) of various levels of abstraction that are unified in speech, writing, or sign, (ii) knowledge of language is dynamic and acquired over a life-time, and (iii) analysis should be usage-based. I understand a usage-based approach to be one that is grounded in instances of language-users producing and understanding language (see Kemmer and Barlow 1999; Bybee 2010; Noël 2017; Diessel 2017, 2019; Tomasello 2003, among many others). Croft's (2001) model of a construction with three form and three meaning components is introduced. It is used later in the case studies to model changes. In keeping with recent thinking about the semantics-pragmatics divide articulated in Jaszcolt (2019) and Finkbeiner (2019b), I suggest that semantics can be characterized as [+truth-conditional, +conventional] and pragmatics as [−truth-conditional, ±conventional]. These distinctions are not absolute, but prototypical.

Chapter 3 presents a usage-based constructional approach to language change. It builds on and updates Traugott and Trousdale's (2013) distinction between constructionalization and constructional changes. In the diachronic perspective adopted, change is understood as the result of change in usage (Bybee 2010), not of parameter settings in universal grammar (as proposed in earlier generative linguistics, see Kiparsky 1968). Several much discussed issues are introduced in this chapter. Among them, in a constructionalist perspective on change what are the roles of innovation, of analogy and neoanalysis (also known as "reanalysis") and contexts? Regarding the role of innovation, it is argued that individual innovations may lead to change but do not constitute it. Change is understood as arising in context when aspects of form and meaning that differ from those of an earlier generation or different community are conventionalized, that is, come to be shared by some community of interlocutors (see e.g. Croft 2000; Schmid 2017).

Discourse contexts that are interpreted as assemblies of constructions (Petré 2016, 2019) play an important role in constructional characterizations of context. The kinds of contexts that are particularly relevant prior to constructionalization are replicated uses in discourse associated with the expression undergoing modification. For example, discourse structuring uses of spatial *by the way* 'en route', arise in replicated contexts of reference to talk, such as *It was **told** him **by the way** that his wyf was deed ('dead')* (1482). This contextual use is consistent with Lenker's (2010) observation that *however* and other English adverbial Connectors arose in locutionary contexts. A further context is topicalized use of the source adverbial (*and by the way tell me how Wales* …) (1608), which places the dynamic spatial adverbial at the beginning of the clause, a position typically occupied by Connectors. A third context is use in non-restrictive relative clauses, for example, *which by the way is worth our obseruation* ('which is worth our discussion in passing') (1616). This places *by the way* in a parenthetical "aside", again in the context of locution. The contexts here are linguistic "co-texts" (Catford 1965: 31). Some contexts are not recoverable in historical work, for example, prosodic contexts. We cannot recapture the kind of continuum between main clause and discourse marker discussed in Dehé and Wichmann 2010's study of the uses of epistemic *I think (that)* and *I believe (that)*), nor can we recapture gestural ones (see e.g. Hata 2016 on the communicative role of gestures associated with discourse markers). Several types of "external" contexts, such as contact, culture change, changes in rhetorical and genre traditions and sociolinguistic factors like age, gender and class are essential factors in change. They can, however, be incorporated in models of construction grammar (see Fried and Östman 2005: 1772) and would ideally be so. But unfortunately we often do not have enough information about such factors, particularly in periods prior to about 1500, to be able to incorporate them consistently with precision and explanatory power in historical work, so the contextual uses have to be gleaned from the texts themselves. Such "external" contexts are beyond the scope of this study.

With regard to analogy, a distinction is proposed between analogical thinking, a process that we engage in on a daily basis, and analogization, a linguistic change based on an analogical match of X to Y. Every instance of analogization is understood as an instance of neoanalysis (also known as "reanalysis"), because an X analogized to Y will be different from its source. However, not all instances of neoanalysis are instances of analogization. Work on Diachronic Construction Grammar privileges analogical thinking, analogization and similarity because construction grammar is concerned with similarities such as are manifest in abstract patterns and schemas.

Chapter 4 articulates criteria for recognizing pragmatic markers in general, and the subset of Discourse Structuring Markers in particular. Most Discourse

Structuring Markers in English are syntactically Conjuncts. They link discourse segment 1 (D1) with discourse 2 (D2) in a [D1__D2] template. In constructional terms DSMs are form-meaning ([[F] ↔ [M]]) pairings of the type: [[Conjunct] ↔ [Discourse Structuring Marker]]. DSMs have a range of uses on the continuum in Figure 1.1. 1DSMs are monofunctional and largely contentful while DMs are multifunctional and non-contentful.

Historically many DSMs in English originate in Circumstance adverbials of location and temporality. A Discourse Structuring Marker Trajectory Hypothesis is put forward in which the form change is from Circumstance adverbial to Conjunct and the meaning changes are from [+contentful] via partially contentful to largely pragmatic features. The latter shift is optional since not all DSMs come to be used as multifunctional DMs. The main factors in this Hypothesis are modeled in Figure 1.3 (to be slightly modified in Chapters 4 and 7):

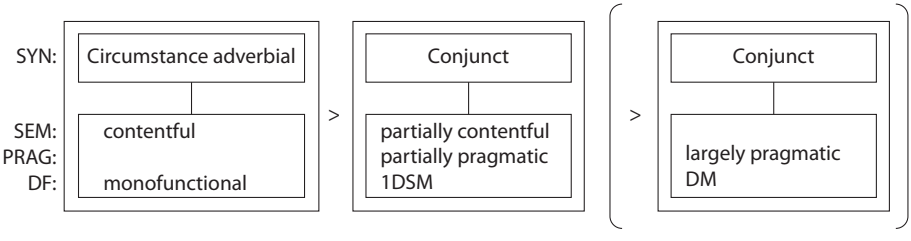

Figure 1.3 The Discourse Structuring Marker Trajectory Hypothesis

The change from Circumstance adverbial to Conjunct entails a fairly significant shift in distributional constraints (syntactic form) and in meaning (the new use has a discourse connective function). This is a constructionalization. Further shifts toward DM function are interpreted as changes in degree of pragmaticality and the development of a new discourse function (DM). These are changes in the meaning component of a construction and are constructional changes. The chapter ends with a summary of the development of DM *after all* as a concrete example of the kinds of changes and evidence for them discussed earlier in Chapter 4.

There has recently been extensive discussion of which historical model best accounts for the kinds of changes evidenced by the development of DMs. In Chapter 5, the final chapter of Part I, some different approaches to the changes leading to the development of DMs are presented and evaluated. One is a grammaticalization approach suggested in Traugott (1997[1995]). As has often been noted, DMs have many of the characteristics of grammatical markers: for the most part they have lexical origins, they are fixed units, and they are procedural. Typical examples of the outcome of grammaticalization processes are tense, aspect, modality and case, all of which are procedural in the sense that they "constrain inferential

procedures" (Blakemore 2002: 79) and guide interpretations in context (Hansen 2008: 20, 2012: 596). However, because they are mainly multifunctional and have scope over the associated clause rather than over V, N or Adj, DMs present a problem for work on grammaticalization, which has been conceptualized as a process of reduction in terms of contentful and segmental substance, scope, and similar factors (see e.g. Heine et al. 1991; Lehmann 2015[1995]). Some researchers have suggested a "pragmaticalization" rather than a grammaticalization approach (e.g. Erman and Kotsinas 1993; Aijmer 1997). Pragmaticalization privileges the pragmatic and scopal properties of DMs, but is not sufficiently distinct from grammaticalization to have received wide acceptance.

A third approach is to propose a dual level discourse grammar consisting of two different subgrammars (Kaltenböck et al. 2011; Heine et al. 2017; Heine et al. In press). One level is sentence grammar, which "is organized in terms of propositional concepts and clauses and their combination" (Heine 2019: 418). The other is thetical grammar, which provides material for spontaneously occurring discourse events and more stable, patterned "external" phenomena such as non-restrictive (appositive) relative clauses, imperatives, comment clauses, and discourse markers. The proposal is that some DMs may undergo some grammaticalization, for example, morphological boundary reduction and "chunking", a gradual process of change, but the shift to DM status is instantaneous: speakers coopt an expression into thetical grammar, where they may be further grammaticalized. For the most part this approach ignores transitional stages in which (i) a Circumstance adverbial is topicalized and used to frame the upcoming content, (ii) the Circumstance adverbial is used as a Conjunct. I show in the case studies that there is no evidence that DSMs, whether 1DSMs or DMs, are instantaneously coopted. They arise gradually. It is proposed instead that constructionalization offers the kind of architecture that accounts well for the development of DSMs under discussion as is demonstrated in the case studies in Part II.

Part II, Case studies, provides empirical evidence for the development of markers that serve a topic-orienting function (see Fraser 2009a) in at least pre-clausal position: elaborative markers like *also* (Chapter 6), contrastive markers like *but* (Chapter 7), digressive topic shift markers like *by the way* (Chapter 8), and markers of return to a prior topic like *back to the point* (Chapter 9). The DSM Trajectory Hypothesis in Figure 1.3 is shown to be supported in all its stages by several DMs such as *also, after all, but, by the way*. However, many expressions with similar functions but less or minimal pragmatic multifunctionality that are listed in Fraser's extensive work on classifying DMs (e.g. Fraser 1996, 2015) are argued to not (yet) be used as DMs. Rather, they have largely monofunctional 1DSM properties. These are markers like *back to my point, in addition, in contrast, instead.*

Each of Chapters 6–9 provides several "semasiological" accounts of the ways in which uses of a particular expression *E* developed by outlining its history. Cumulatively, these accounts provide a partial "onomasiological" perspective on the options available to a language user for a particular discourse structuring concept. In a semasiological approach an expression is traced though its history, keeping form relatively constant, whereas in an onomasiological approach, the focus is on a concept and changes in the ways in which it is expressed over time. The distinction between semasiology and onomasiology has traditionally been theorized within the domain of the lexicon (see Geeraerts 1997, Grondelaers et al. 2007). It is also useful in conceptualizing the distinction between a constructional perspective on the rise of procedural aspects of language on the one hand (an essentially semasiological approach), and schemas and the "constructional space" within which particular semasiological changes occur on the other (an essentially onomasiological approach).

The last chapter in Part II, Chapter 10 is devoted to the development of possible combinations of 1DSMs and DMs and complements synchronic work by Koops and Lohmann (2015) and Lohmann and Koops (2016) on combinations of the eleven "DM"s that are the topic of Schiffrin (1987). Focus is on the development of the unit DMs *now then*, and incipiently *Oh, by the way* in one of its uses.

Part III, Three open issues, draws on analyses in earlier chapters to discuss a proposal about the roles of subjectification and intersubjectification in a historical constructionalist model, and two current questions about the constructional approach to grammar. One is whether position of DSMs relative to the clause is a construction, and the other is how to think about nodes and networks from a historical perspective.

The first chapter of Part III, Chapter 11, addresses the question how we might think about the meaning changes known as subjectification and intersubjectification from a constructionalist perspective, and especially how these changes interact with the rise of DSMs. The hypothesis is that since DSMs are procedurals, they are used by speakers and writers, consciously or not, to guide addressees and readers to interpret the relationships between D1 and D2. This means that when an expression that is a circumstance adverbial is used as a DSM, there will at the same time be both subjectification and intersubjectification of that source expression. For example, when used as a DSM, *by the way* is more (inter)subjective than when used as a dynamic spatial adverbial. Further subjectification or intersubjectification may occur later, for example, *by the way* came to be used more intersubjectively when it was used as a hedge attenuating an upcoming face threatening D2.

Chapter 12 addresses the question whether position with respect to the clause can be analyzed as a construction. This question arises out of extensive recent interest

in the roles of "left periphery" and "right periphery" (e.g. Beeching and Detges 2014), also called "front field" and "end field" (Auer 2005) and especially in the relationship of different pragmatic meanings and functions at turn beginnings and endings (e.g. Heritage 1994, 2002; Deppermann and Günther 2015; Haselow 2019).

After all is an example of a DM that was consistently used with distinct functions in different positions in its earlier history. In 19thC and early 20thC data from COHA use of *after all* in pre-clausal position marks the upcoming segment as a justification or explanation for saying what precedes, as in (2a), in medial position especially in BE-predications (2b), it marks the segment as not only an explanation of the proposition in which it is used but also as implicating that the information is or should be epistemically accessible and obvious to AD/R (roughly equivalent to 'of course'), while in post-clausal position (2c) it marks the proposition in the associated segment as referring to something unexpected or that counters what was earlier thought (a "concessive" use that is often used to refer back to an expectation raised at some textual distance). In contemporary English post-clausal *after all* is associated with a different prosody from *after all* in other positions. As mentioned above, we do not have access to prosody at earlier times. Although it is likely that there were different prosodies, we do not know what they were. I consider the uses in different positions to be polysemous and probably prosodically distinct.

(2) a. Mabel's uniform policy was that of outward submission to the mandates of her chief. "**After all**, it makes little difference!" she fell into the habit of saying… (1870 Harlan, *At last* [COHA])

b. The United States christens its splendid cruisers after great cities, its battleships after States. We might take a leaf from our cousins' book, especially as there is English precedent for it, and it would interest the people in the navy, which is, **after all**, one great object to be striven for.
(1895 *New York Times editorial* [COHA])

c. what I meant to imply was, that Kentuck's situation is not so bad, **after all**.
(1834 Thomas, *Clinton Bradshaw* [COHA])

After all is a prime example of a marker that was subjectified by default when it was used as a Conjunct, and was later further subjectified as a marker of the three perspectives on D2: justification, obviousness, or counterexpectation. In the second use it was also intersubjectified as the addressee is specifically called on to agree with the proposition. While the resulting distinctions between uses continue to apply in contemporary data, they have been somewhat relaxed, and are now rather weak tendencies only, especially concessive use in post-final position. Haselow (2019) has suggested that the serial order of turn beginnings and endings in a conversation depends on the communicative tasks that interlocutors are engaged in, and

the question has arisen whether the position of DMs in linear combinations such as are discussed in Chapter 10 is meaningful and therefore constitutes a construction. The same question can be asked about position in the clause. It is tempting to answer the question in the affirmative, because post-clausal position tends to be associated in English with concessive interpretations (Lenker 2010). However, concessive appears to be relevant only to contrastive Connectors and correlations between use of individual DMs in a particular position are highly item-dependent, so there is not sufficient evidence that position is a construction.

The third chapter in Part III, Chapter 13, attempts to answer Smirnova and Sommerer's (2020: 3) question in a preliminary way: "How can node creation and node loss be implemented in the network model?". In this Chapter ways are suggested in which changes in networks can be formalized using Croft's (2001: 18) model of constructional form-meaning features, building on earlier analyses of *after all* and *by the way*.

The book ends with a brief summary of main points and suggests some further topics of research (Chapter 14).

1.5 Data and methodology

The data used for this book are drawn from a number of electronic corpora. For the full set of sources, see Data Resources and Corpora at the beginning of the References. The main corpus sources used are:

1. For Old English the *Dictionary of Old English Corpus* (DOEC), a corpus of over 3 million words of Old English and 1 million of Latin. According to the website at https://www.doe.utoronto.ca/pages/index.html, it is "a database consisting of at least one copy of every surviving Old English text". It is not balanced or parsed. The part of the *Dictionary of Old English* itself (DOE) that is available so far (for the letters A-H) was also consulted.
2. For Middle English, the *Harvard Geoffrey Chaucer Website* (HGCW).
3. For Early Modern English *Early English Books Online* (EEBO), a corpus of books in print during the period 1470–1700. This corpus has the great advantages of being large (755 million words) and of providing some figures for frequent lexical collocations. However, it is not balanced or parsed. Each text bears the date of publication, not composition, e.g. Chaucer's *Canterbury Tales* is dated 1477, but they were written from 1387–1400. Texts in EEBO are largely historical, homiletic, philosophical and religious, and therefore not contexts in which large numbers of Pragmatic Markers might be expected. However, since DSMs are central to discourse organization, the subtypes of expressions

analyzed here are in fact quite widely attested, as is metatextual commentary on observations, digressions and insertions. Many texts are translations from Latin and French, so ultimately comparison with Latin, Medieval Latin and Medieval French would be needed to fully understand to what extent the examples in question represent native English usage.

4. For the Modern period, the *Corpus of Historical American English* 1810–2009 (COHA), a balanced corpus originally of 400 million words divided into decades and representing fiction, magazines, newspapers and other non-fiction 1810–2009. In 2021 COHA was expanded to 475 million words representing the 1820s-2010s. Data from TV and Movies have also been added to data for 1930s on. I state explicitly which release was used when counts are involved.
5. For Present Day English, the *Corpus of Contemporary American English* 1990–2019 (COCA). When I started my research for this book, I used a version covering the years 1990–2019 with 600 million words and providing data from a balanced corpus of spoken data (mainly TV, radio and other public speech), magazines, newspapers, and fiction. Unless otherwise specified, any counts provided are based on this version. In March 2020 the corpus was expanded to 1 billion words, including extensive examples of blogs, TV, movies and web entries. Examples selected from these registers are identified, as in the corpus, as BLOG, MOVIE, TV, WEB.

In addition some other corpora and resources were used such as *The Corpus of Late Modern English Texts*, version 3 (CLMET_3.0), 34 million words 1710–1920. The *Middle English Dictionary* (MED) and *Oxford English Dictionary* (OED) are useful resources for etymologies and approximate dating of new uses. They are designed to explicate meanings of words, not to provide representative samples of English at different times and therefore were used only as references. The OED in particular is problematic if used as a corpus (Hoffmann 2004) because it does not provide consistent access to clauses prior to or following the target one, and many entries are still being updated.

The corpora and data bases were chosen because they are large and relatively easy to access. For the most part early corpora like DOEC and EEBO represent written genres such as treatises on religion, philosophy and institutional documents. The advent of printing with Caxton's printing press in 1476 made documents relatively easy to produce and replicate, and materials representing conversational speech came to be more widely available. However, because it is a corpus of printed books, EEBO largely consists of writing-based genres, rather than broadsheets and other more colloquial printed texts. It does nevertheless include sermons and other genres that Culpeper and Kytö (2010) consider to be

relatively speech-like.[4] They conceptualize genres close to speech as "speech-like" (personal correspondence), "speech-based" (trial proceedings which purpose to be reports of what was said), and "speech-purposed" (plays which represent fictional interaction, sermons and other works that even if written are meant to be spoken) (Culpeper and Kytö 2010: 17). Based on this classification, sermons in EEBO are "speech-purposed". COHA includes a significant amount of data from fiction, much of it also "speech-purposed" writing (for example, Dickens' novels were intended to be read aloud). Although the SPOKEN component of COCA is transcribed from speech, for the most part it does not represent ordinary conversation, but rather public talk such as is found in TV and News shows. The transcriptions can be considered to be "speech-purposed" as well, and some can be considered to be "speech-based". In sum, although the corpora are written they exemplify a wide range of genres, many of them speech-purposed. This is an important factor in the development of Discourse Structuring Markers, as their function is to guide the argumentation process.

The corpora were compiled assuming fairly traditional views of periodization in English. Periodization is a debated topic because it is in part based on non-linguistic factors such as the Norman Conquest for the transition from Old to Middle English and the establishment of printing for the Early Modern period (see Lass 2000) and because particular changes often do not coincide with traditional dates (Gries and Hilpert 2012). However, the traditional periodization as in Table 1.1 serves as a useful guide for thinking about stages of development of DSMs in English and is retained here.

Table 1.1 Traditional periodization for the history of English

Old English (OE)	c650–1100
Middle English (ME)	c1100–1500
Early Modern English (EModE)	c1500–1700
Modern English (ModE)	c1700-present
Present Day English (PDE)	c1970-present

Data were drawn manually and the methodology is qualitative. This means that the history of Discourse Structuring Markers is outlined in broad strokes only. The kind of fine-grained analysis that quantitative analysis makes possible is exemplified in Culpeper and Kytö's (2010) study of changes in use of one marker, *and*. Comparing phrasal, coordination uses of *and* like *go and get a Wedding Ring*

4. Culpeper and Kytö (2010) base their work on *A Corpus of English Dialogues 1560–1760* (CED) and the *Corpus of Early English Correspondence 1414–1680* (CEECS). These are relatively small corpora (1.5 million words and half a million words respectively), so data of the kind discussed in this book is relatively sparse in them.

and clausal, conjunct uses of *and* in interactional data, Culpeper and Kytö present evidence for differences in use among text-types and genres within them, e.g. narratives in witness depositions within trials, or within didactic works. This work serves as an important reminder not only that "speech and writing are in no way uniform varieties of language" (Culpeper and Kytö 2010: 182), but also that data such as are presented in the present book give only a broad-scale view on highly variable usage patterns. Ideally a quantitative study might be used to complement the qualitative one provided here (see Evers-Vermeul et al. 2011 on the merits and limitations of both types of analysis), but as several of the expressions studied in this book appear only in small numbers in the corpora, quantitative analyses would not be particularly revealing in all cases.

Candidates for analysis were identified based on extrapolating back from lists of alleged examples of DMs in PDE such as are provided in Fraser (2009a) and dictionary entries for members of these lists in OED and MED, where variant spellings are provided.

Since a crucial characteristic of DMs in the restrictive sense used here and the larger class of DSMs is that they are clause-Connectors, data were restricted to finite causes with subjects and to imperatives introduced by a potential DSM. This excludes examples of *after all* in (3) where the explicit subject *patrons* precedes in D1 and is null in D2:

(3) that the patrons of image-worship are at a mighty loss what to say to it, and **after all** are forced to cry out against it as supposititious ['not genuine'].
(1675 Cave, *Primitive Christianity* [EEBO])

Most examples of DSM use were clearly cases of pre-clausal, clause-medial or clause-final use, as in (2). However, there are instances of DSMs that are not obviously one of these, for example, when a DSM precedes a complementizer, as in (4). Is *after all* post-clausal here or clause-medial because *that*-complements are arguments? Taking the latter position, I coded a DSM that precedes a *that*-complement as in (4) as clause-medial:

(4) I would observe, **by the way**, that it costs me nothing for curtains.
(1854 Thoreau, *Walden* [COHA])

A fundamental problem in studying the history of DSMs with lexical sources is how to distinguish reliably between contentful/referential and procedural uses (see Hansen 2008 for detailed discussion of the differences, and Hansen 1998: 204–205 on the difficulty even with contemporary data of obtaining reliable native speaker judgements about uses and paraphrases). Defour's (2007) study of the history of uses of *well* and *now* provides a model of how to address the problem in highly nuanced detail. It follows that another problem is how to distinguish clause-initial (referential) from pre-clausal (more procedural) uses. This kind of

problem presumably in part mirrors problems that Hearers and Readers face and supports the hypothesis that change is gradual.

Note that some possible criteria do not lead to adequate distinctions. Among these are:

a. Clause-initial position, because topicalized Circumstance adverbials can occur there. The terms "clause-initial" is best reserved for the position in which argument structure material may be used. "Pre-clausal" is preferable for DSMs (see Gregory and Michaelis 2001). However, since it is often not possible to distinguish clause-initial from pre-clausal without prosody or extensive context, searches were conducted for expressions that are attested at the beginning of a discourse segment with a finite verb.
b. Presence or absence of commas, because in EModE punctuation was not as codified as now (Parkes 1992). Furthermore, views have shifted about faithfulness to originals in editions. Editors often change or add punctuation, and EEBO is a normalized corpus of printed books that are likely to have been heavily edited or not faithfully typeset from the original manuscript. Therefore expressions both with and without commas were collected, except when the number of hits is very high, e.g. *also*. In such cases, the expression was searched with a comma (*also*,).

Plausible paraphrases in context can be invaluable. They can, however, also be problematic as they are contemporary and may have associated nuances that were irrelevant at an earlier period, so pragmatically neutral paraphrases are needed. For example, a paraphrase for *by the way* used as a location adverbial is 'where?'. A paraphrase for the DSM Conjunct use of this expression is 'in passing/along the textual route'. For DM use a useful paraphrase is 'I note that X and invite you to regard X as not particularly important'. 'If I may say so' was used as a paraphrase for hedging use.

Some of the researchers cited, for example Fraser (1996, 2009a), have tended to use constructed data in their analysis of DMs. This provides valuable insight into the patterns that a speaker can assemble, and what a speaker knows about an expression, but may lead to claims that are at odds with data. In my view, constructed data are valid and valuable data, because they tell us something about a language-user's knowledge and capacity for language use, especially when constructed data provide more information than corpus data (see further Chapter 9.5). However, before an expression or sequence of expressions is excluded on the basis of constructed data, support from a corpus is useful. If there is a discrepancy it is important to acknowledge that. For example, in a squib on *well* cited in Aijmer (2002: 31), Murray (1979: 731) uses some constructed examples, including:

(5) a. Oh, by the way…
b. *Well, by the way…

(6) a. ?Oh, so we'll meet at three.
b. Well, we'll meet at three.

Murray concludes: "Clearly, only *well* may introduce an anticipated topic, and only *oh* a new one". The March 2020 version of COCA provides 20 examples of *Well, by the way*.[5] These include (7) where Ms. Allred says she has been addressed by the wrong name, a topic certainly unanticipated in light of the question:

(7) I would like to know, Ms. Albright, as a lawyer, why if it had nothing to do with it, would you have taken this case if he were just an average Joe? *ALLRED*: ***Well, by the way***, I'm not Ms. Albright. (1997 *CNN_Talkback* [COCA])

The examples are few, but in situations such as this, a judgment based on constructed data should be questioned rather than assumed to be reliable.

1.6 Summary

In sum, the main theoretical questions for this book are:

1. How can pragmatic factors best be incorporated in Diachronic Construction Grammar?
2. How can contextual factors best be incorporated in Diachronic Construction Grammar?

The empirical domain investigated is the rise of Discourse Structuring Markers (DSMs) that signal the relationship that speakers and writers intend (or claim to intend) between discourse segment 1 and discourse segment 2. The particular types of DSMs investigated all play a role in signaling topic-orientation. They are:

1. Elaboratives
2. Contrastives
3. Digressives
4. Markers of Return to a prior topic

The methodology is qualitative and the data are drawn from large electronic data bases.

---

5. There are 59 hits, but some are interrupted and others involve *well* used as a clause-final adverb (*you've done well, by the way)* and are not examples of what Murray had in mind.

PART I

Foundations

CHAPTER 2

Cognitive linguistics and construction grammar

2.1 Introduction

In this chapter I outline the main assumptions characteristic of cognitive linguistics and construction grammar that guide my work. In Section 2.2 some foundational ideas on cognitive linguistics laid out in Langacker (1987) are briefly reviewed. Section 2.3 outlines some key principles of construction grammar, mainly as developed in Goldberg (1995, 2006 and elsewhere). Croft's (2001) model of a construction is presented in Section 2.4. Section 2.5 proposes how the semantics-pragmatics interface can be conceptualized in a constructionalist framework. 2.6 summarizes the key points.

2.2 Cognitive linguistics

Cognitive linguistics in its various manifestations goes back to the 1970s. Langacker (1987, 1991) articulated many of the foundational assumptions of cognitive grammar.[6] Among assumptions that are particularly important for the work discussed in this book are:

a. Knowledge of language consists of form-meaning pairs of various levels of abstraction that are unified in speech. In other words, language is a symbolic sign system.
b. "Linguistic relationships are not invariably all-or-nothing affairs" (Langacker 1987: 14). Linguistic categories are gradient, without sharp boundaries.
c. Meaning is not fixed. There is "meaning potential". This means that there is an "essentially unlimited number of ways in which an expression can prompt dynamic cognitive processes" (Fauconnier 2008: 661). However, context grounds and restricts interpretation (Langacker 1987: 497).
d. There is no, or very little, innate linguistic knowledge. "Blueprints for language" are learned (Langacker 1987: 13).
e. Constructions are linked in networks (Langacker 1987, 1988).

---

6. See Goldberg (2013) for bibliographical references to earlier articulations of these foundational concepts.

Another manifestation of cognitive linguistics is Frame Semantics (Fillmore 1976; Fillmore and Baker 2001) and more recently FrameNet (see Petruck 2011 for a brief overview). Key to this approach is the argument that knowledge of language is based in experience, and experience involves situation frames. For example, a concept like *buy* involves knowledge of a frame that includes an agent who is a seller, a theme (the object to be bought), and intended transfer from seller to buyer.

Arguing that a cognitively based theory takes "not the objective 'real world,' but human perception and understanding of the world to be the basis for the structure of human language" (Sweetser 1990: 2), Sweetser developed a highly influential theory of polysemy based in metaphorical frames. She starts the book with a historical account of why verbs like *grasp* and *see* have both physical and intellectual/ cognitive readings, showing that metaphorical frames have a long history in Indo-European. She goes on to investigate uses in contemporary English of modals (e.g. *may* is polysemous between 'permission', and 'possibility'), causals and adversatives, coordinating conjunctions and conditionals. In most domains Sweetser found three kinds of meaning: real world/sociophysical, epistemic and speech act meanings. The "sociophysical" meaning is a representation of what the speaker conceptualizes as a "real world" phenomenon. "Epistemic" meaning pertains to representation of speaker's knowledge and conclusions drawn from that knowledge, and the "speech act" meaning to the locutionary act being performed. Example (1) illustrates with *so* (Sweetser 1990: 79, italics original):

(1) a. He heard me calling, **so** he came.
(The hearing caused the coming, in the real world.)
b. (You say he's deaf, but) he came, **so** he heard me calling.
(The knowledge of his arrival causes the *conclusion* that he heard me calling.)
c. Here we are in Paris, **so** what would you like to do on our first evening here?
(Our presence in Paris enables my *act of asking* what you would like to do.)

However, not all connectives she discusses are used in all three domains. While not wanting to "state categorically that there is no such thing as a content-domain use of *but*" (Sweeter 1990: 103), drawing on Lakoff (1971) Sweetser concludes that uses of *but* like (2) that might be thought to pertain to the real world content-domain are actually in the domain of the speaker's conclusion because clash or contrast depend on a speaker/writer's (SP/W's) beliefs:

(2) John eats pancakes regularly, *but* he never keeps any flour or pancake mix around.

Of (2) she says "the naturalness of a pancake-eater's stocking flour would lead us to *conclude* from John's habits that he stocks flour. However, this conclusion clashes with the (otherwise well-supported) fact that he doesn't stock flour" (Sweetser 1990: 103, italics original). Like *so*, *but* can be used with a speech act function, as in (3):

(3) I love you *but* PLEASE take those wet boots off the carpet!

But in (3) "might express a conflict between the supposed expectations set by a speech act like 'I love you,' and the actual reproofs which follow" (Sweetser 1990: 106).

The main title of Sweetser's book, *From Etymology to Pragmatics*, highlights the importance in her thinking of pragmatics and she writes: "the choice of a 'correct' interpretation depends not on form, but on a pragmatically motivated choice between viewing the conjoined clauses as representing content units, logical entities, or speech acts" (Sweetser 1990: 78). Given that she discusses *and, but, because*, and *so*, which are four discourse markers in the restrictive sense among the 11 discourse markers in the looser sense discussed in Schiffrin (1987), it is surprising that little connection seems to have been made subsequently between discourse markers and cognitive analyses of these kinds of expressions.[7] Instead, most attention has been paid to topics in the second half of the title of Sweetser's book: *Metaphorical and Cultural Aspects of Semantic Structure*.

As stated in Chapter 1.2, the purpose of the present book is to explore what the role of pragmatics and constructions should be in a constructional approach to language change and thereby to fill part of the pragmatic gap in cognitive linguistics.

2.3 Goldberg's model of construction grammar

An outgrowth of cognitive linguistics that has inspired the present study is construction grammar as conceptualized in Goldberg (1995, 2006 and elsewhere). While the tenets of cognitive linguistics mentioned above in Section 2.2 are essential for Goldberg's work, she refined some and developed some additional ones (Goldberg 2013). Among the tenets particularly relevant for the present study are:

a. Knowledge of grammar is dynamic and can change over a life-time (see e.g. Croft 2001, MacWhinney and O'Grady 2015).

7. However, there are brief mentions in Onodera's (2004) analysis of Japanese discourse markers.

b. Grammar consists of constructions: [[Form] ↔ [Meaning]] pairings, which are stored in a constructional lexicon known as a "constructicon" (Goldberg 2019: 35–37; to highlight its morphology, she spells it 'construct-i-con' there).
c. There is no strict dichotomy between lexical and grammatical constructions. Lexical constructions (e.g. *constructionalization*, ditransitives like *Mary gave me a scarf*) are on the pole of content meaning. Grammatical constructions (e.g. present tense, possessive case) are on the procedural pole (see Figure 1.1 in Chapter 1.3, and Section 2.5 below).
d. When speakers produce an utterance, constructions stored in the constructicon are assembled (Goldberg 2003: 221; Lyngfelt 2018), provided they do not conflict. For example, *Why didn't you go?* is conceptualized as an assembly of, among other constructions:
   i. an event construction ('You went'),
   ii. a purposive adverbial construction ('for a reason'),
   iii. the negation construction ('n't'),
   iv. the interrogative construction ('subject-verb inversion').
e. "Semantics, information structure, and pragmatics are interrelated; all play a role in linguistic function" (Goldberg 2013: 16). This tenet underlies much of the present book.
f. Knowledge of language is usage-based and grounded in "instances of a speaker's producing and understanding language" (Kemmer and Barlow 1999: viii; also Bybee 2010; Noël 2017).[8]

The term "usage-based" is interpreted in different ways in cognitive linguistics. For Langacker (1988), "usage-based" refers to an approach that is "maximalist", "non-reductionist" and "bottom up", in contradistinction to the tenets of generative grammar that he characterizes as concerned with economy, generativity, and reductionism. For Goldberg (2013: 16) it means that "[k]nowledge of language includes both items and generalizations, at varying levels of specificity". However, she also cites Kemmer and Barlow (1999). It is their definition in (f) immediately above that I adopt.

In earlier work on construction grammar, e.g. Goldberg (1995, 2003, 2006), constructions are said to be of any size (morpheme, word, phrase, clause, larger discourse). However, in sign-based construction grammar, the minimal construction is a phrase. More recently, Diessel (2019: 11) defines a construction as a pattern involving "at least two meaningful elements" because lexemes (monomorphemic expressions) are "processed in very different ways" from sequences of meaningful

8. It should be noted that not all versions of construction grammar are usage-based, among them sign-based construction grammar (e.g. Boas and Sag 2012; Michaelis 2013).

elements. However, for historical work this is problematic, as what is now a word may once have been a phrase (e.g. *yesterday* < Old English *giestran dæg* 'yesterday day') or clause (*goodbye* < *God be with you*). In Chapter 8.3.2 we will see how the overall development of the manner adverb *incidentally* into a digressive and hedging discourse marker is similar to that of the phrase *by the way*, which despite its origins and spelling is no longer phrasal but a unit equivalent to a word. I therefore consider constructions to be of any size, including words.

Constructions with phonological form, e.g. *also, by the way*, are called "substantive". Constructions can be conceptualized at several levels of abstraction: substantive micro-constructions that are abstractions away from utterances ("constructs") and stored in the constructicon, and more schematic constructions which may be partially filled (i.e. contain substantive micro-constructions along with abstract variables as in *What's X doing Y?)* or completely abstract, e.g. ditransitive SUBJ V OBJ1 OBJ2 (an instance of which is *Sally gave Bill a book*), which are also stored in the constructicon. Constructions are organized in taxonomic "inheritance" hierarchies, which show "vertical" mother-daughter relationships. While assessments can be made in psychological experiments of the neuronal reality of abstract schemas and subschemas in contemporary language use (e.g. Perek 2012) or can be deduced from the way languages are learned (Goldberg 2019), we have to acknowledge that as historical linguists we can only hypothesize that speakers of earlier centuries made the kinds of abstractions that we deduce from the data (see further Chapter 3.2.2).

Recently it has been proposed that to account for variation such as can be found in morphophonology, as in /s, z, əz/ plurals (cf. *cats, dogs, horses*) or in syntax as in ditransitive-prepositional alternations (cf. *Sally gave Bill a book ~ Sally gave a book to bill*) "horizontal" relationships are needed as well as vertical ones (Cappelle 2006; Van de Velde 2014; Perek 2015). The architecture of such relationships will be discussed in more detail in Chapter 13 on networks. For the present they can be visualized as in Figure 2.1, where "Cxn" is short for 'construction'. Figure 2.1 represents an abstract vertical network (represented by vertical lines). In addition, one pair of abstract micro-constructions is in a horizontal relationship (represented by the horizontal arrow).[9]

9. There are three constructional levels in Figure 2.1. In earlier work I referred to micro-, meso-, and macro-constructions. While three levels are useful, there is no principled reason to posit only three. How many levels are posited depends on the degree of abstraction that the researcher finds useful and realistic.

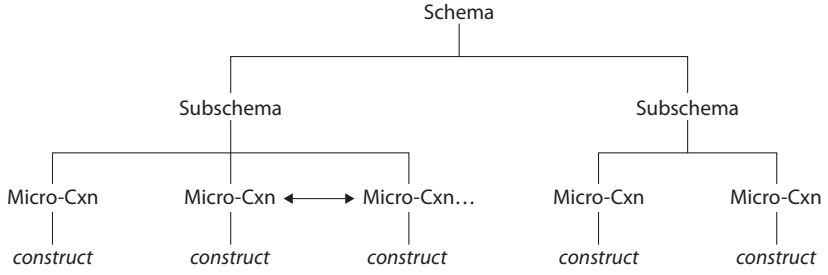

Figure 2.1 Minimal architectural representation of a constructional hierarchy

2.4 Croft's (2001) model of a construction

A much-cited constructional model of the link between form and meaning appears in Croft (2001) and is represented in very slightly modified form in Figure 2.2. The properties are considered to have the potential to overlap, so they are not be interpreted as discrete. Note the symbolic link between the form and the meaning components, and the fact that meaning subsumes pragmatic and discourse-functional properties as well as semantic properties. "Discourse-functional properties" include information structuring and particular types of discourse marking functions.

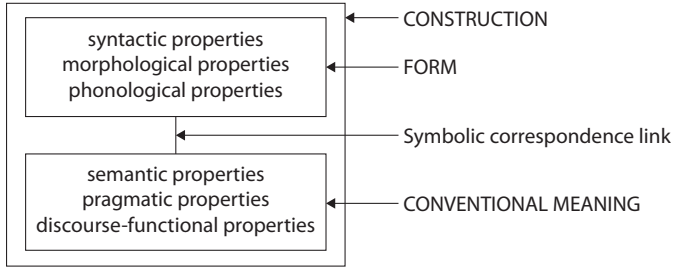

Figure 2.2 The symbolic structure of a construction (based on Croft 2001: 18)

Croft (2005: 280) expands on this model, saying:

> Constructions are symbolic units, a pairing of form and meaning, where both form and meaning are construed broadly: the former including morphology, syntax and even phonology and prosody, and the latter including semantics, information structure/discourse function, and also social parameters of use.

In this book the model in Figure 2.2 will serve as the template for several historical analyses.

Constructions are "symbolic units" (Croft 2001: 18, 2005: 80) and, as has often been pointed out, akin to Saussure's (1983[1916]: 65–70) concept of the "sign", except that constructional signs encompass clauses and complex sentences as well as words. One of Saussure's principles of the sign is arbitrariness. Constructions are arbitrary in that the [[F] ↔ [M]] link is a matter of social and historical convention. Another principle is that a sign is an "entity" or "whole". I understand this to mean that, as a pairing, a particular construction can be used as a pattern. It is often suggested that such patterns are "holistic". Properties may be "linked to the entire grammatical pattern [e.g. passive] rather than to particular components" (Diessel 2015: 299). It is not useful, however, to think of a construction as a "holistic" unit that is an "unanalyzable whole" the properties of which cannot be accessed separately, as suggested in Smirnova and Sommerer (2020: 12).

The unit property of a construction is represented in Croft's model by the outside box. As Langacker (1987: 59) says, "[d]espite their internal complexity, units are effectively simple for purposes of manipulation and combination with other structures". They serve as convenient "packages" of form and meaning that are easily retrievable and usable. They form nodes in constructional networks. But within the constructional envelope represented in Figure 2.2, there are two smaller boxes, one for form and one for meaning, and these two boxes specify three properties each of which is subject to change. In other words, a construction is a unit with complex internal structure. If a construction were truly unanalyzable, change could not happen. As discussed in Chapter 7.3, *all the same*, which meant 'in exactly the same way', came to mean 'in contrast' in negative contexts. In this case meaning, especially pragmatic properties, was reinterpreted, but form did not change. Form may also change independently of meaning, as in the case of phonological reductions such as the chunking or univerbation of *all the same* that preceded use as a contrastive DSM, and the phonological assimilation of *want to* to *wanna*. It follows logically that components of constructions are accessible, even though constructions are "wholes". In historical work, the contingencies of accessibility that depend on a particular speaker's current goals, focus of attention, priming, etc., cannot be recaptured. However, accessibility can be hypothesized to depend on evidence for use in replicated contexts and outcomes of contextual uses.

2.5 The semantics-pragmatics interface

In this section, I turn to the question how the semantics-pragmatics interface can be conceptualized in a constructionalist framework, with special attention to the concept of "procedural" meaning.

In Croft's model, semantics and pragmatics are on separate lines and might appear to be conceptualized as discrete subdomains. However, the distinctions are meant to be gradient, not rigid, so they do not counter the principle of cognitive linguistic mentioned in Chapter 1.1 that that there is no fundamental modular distinction between semantics and pragmatics. Semantic properties have to do primarily with potentials for referential and contentful meaning, pragmatic properties have to do primarily with potentials for inferences, procedural meanings, and so forth. As noted in Chapter 1.4, procedurals guide interpretations in context (Hansen 2008: 20, 2012: 596). Procedurals point "hearers to particular – more or less – schematic frames of interpretation for the utterance hosting such expressions" (Hansen 2012: 595). Consider (4):

(4) Well, I really don't understand why other people don't leave diplomats do their work. **After all**, we have been in contact. (1992 *CNN_King* [COCA])

Here the speaker uses *after all* to cue that what follows (*we have been in contact*) is to be understood as a justification or explanation of what he has just said. The schematic frame has the discourse function 'upcoming discourse segment 2 justifies prior discourse segment 1'.

In Figure 2.2 "discourse-functional properties" include:

a. the illocutionary force of an utterance in a particular context, e.g. statement, question, command, justification,
b. features of information structure such as topic, focus, given, new, background-foreground,
c. evaluations such as surprise.

Why then has so little been said in cognitive linguistics in general and construction grammar in particular about pragmatics? And, it follows, why has so little been said about the role of discourse markers? Finkbeiner suggests that the reason little attention has been given to theorizing pragmatics in cognitive linguistics is that there has been a mind-set about semantics that has precluded or avoided contextual modulation. It has been assumed that semantics is context-invariant, while pragmatics is highly context-dependent:

> semantics deals with truth-functional, context-invariant aspects of meaning, while pragmatics deals with meaning aspects that are derived from actual utterances via inferential reasoning, i.e. with aspects that vary with context (Grice 1989).
>
> (Finkbeiner 2019b: 174)

An example of this kind of thinking is: "pragmatics is largely part of invisible backstage cognition, with exceptionally sparse formal or symbolic marking" (Fauconnier 2008: 674). On this kind of view, semantics is [+truth-conditional] and [+conventional] (i.e. shared across some community of speakers), and pragmatics is [−truth-conditional] and [−conventional]. Most of the contexts that Finkbeiner associates with Gricean thinking are real-world, conversational inferences that arise on the fly, e.g.:

(5) A. Do you have time for coffee?
B. My train goes at eight o'clock.

Speaker A has to compute in the particular context whether 8 o'clock is too soon for there to be time for coffee or whether it leaves enough time for coffee, and therefore whether Speaker B's answer can be interpreted as *No* or *Yes*. The traditional view of the semantics-pragmatics interface, as outlined in Finkbeiner (1919b: 185), assumes a semantics-pragmatics divide that can be pictured as in Table 2.1:

Table 2.1 The traditional semantics-pragmatics divide

Semantics	**Pragmatics**
+ truth-conditional	−truth-conditional
+ conventional	−conventional

Another example of the traditional thinking represented in Table 2.1 is logicians' view that *and* and *but* are both semantically [+truth-conditional and +conventional] and therefore equivalent. In formal logic both *and* and *but* are represented by '∧'. But in language use they do not mean the same thing pragmatically. If they did, (6) would be meaningless.

(6) But me no buts. (1709 Centlivre, *The busie body* [Wiktionary])

Contrast *and me no ands*, which requires considerable mental search for contexts to be understood as meaningful. Although Grice (1989[1967]) firmly maintained the [±truth-conditional] divide (Jaszczolt 2019), he proposed dealing with the problem of *but* by suggesting that it is associated with a generalized/conventionalized implicature of contrast. Since then a partial (and still strongly debated!) picture has emerged "that semantics and pragmatics both participate in the determination of the truth-conditional content" (Finkbeiner 2019b: 177).[10] Earlier Pons (2008)

10. Finkbeiner (2019b: 176) cites Recanati (2010) and Depraetere and Salkie (2017) as having similar views of the relationship between semantics and pragmatics.

argued that both conceptual and procedural meanings can co-exist within a single marker and Mauri and van der Auwera (2012) among others moved to put more pragmatics in semantics and more semantics in pragmatics (see also Jaszczcolt 2019 and elsewhere).

On a view of an enriched pragmatics and of overlap between semantics and pragmatics, some expressions can be characterized as both non-truth-conditional and truth-conditional, both conventionalized and non-conventionalized, in other words. As will be discussed in later Chapters, some Discourse Structuring Markers like *further*, have partially contentful meaning. This is especially true in early stages of development, for example at the time of transition from lexical adverbial to Connector. When it was first used not as a spatial adverbial but as a Connector, *by the way* was used to meanu 'in passing', which is partially contentful, but when it came to be used as a digressive and dismissive DM, it was no longer contentful. Therefore, the possibility that some Connectors can be [±truth-conditional] at least at some point in their history, needs to be acknowledged. However, here I highlight prototypical characteristics of semantics and pragmatics as I understand them at the extreme poles of the pragmaticality continuum (see Figure 1.1 in Chapter 1.3). I have specified that pragmatics is prototypically [−truth-conditional] in Table 2.2 below. Just as the characteristics of constructions in Croft's model in Figure 2.2 are not meant to be discrete, so the '+' and '−' features in Table 2.2 are not meant to be discrete. Rather, they are features that allow for gradience, as a dynamic model requires. The grid in Table 2.2 replaces the grid in Table 2.1. It preserves the long-recognized characteristic of pragmatics that processes of inferencing may enrich utterances with meaning in a specific situation ([−conventional]) and adds the now widely recognized characteristic that pragmatic meanings can also be conventionally associated with certain expressions ([+conventional]).

Table 2.2 Prototypical characteristics of semantics and pragmatics adopted here

Semantics	**Pragmatics**
+truth-conditional	−truth-conditional
+conventional	±conventional

2.6 Summary of key points

This chapter has surveyed some key points in cognitive linguistics (e.g. Langacker 1987), especially in the model known as construction grammar as developed by Goldberg (e.g. 1995, 2006). These will be foundational for the rest of the book. The most important of these points are that:

1. Knowledge of language consists of form-meaning pairs (constructions) of various levels of abstraction that are unified in speech or writing,
2. Semantics and pragmatics overlap,
3. Knowledge of language is dynamic and acquired over a life-time,
4. Analysis should be usage-based, that is, the approach should be one that is grounded in instances of a speaker or writer producing and understanding language.

   Additional points addressed in this chapter are:
5. Constructions are conventional in the sense of being shared by a community.
6. Some expressions can be characterized with an enriched pragmatics that is both non-truth-conditional and conventionalized.
7. Croft's (2001: 18) model of a construction predicts that different components of a construction can be accessed independently from others even though the construction is functionally a unit.

In the next chapter I consider ways in which constructions can be viewed from a historical perspective.

CHAPTER 3

A Diachronic Construction Grammar view of language change

3.1 Introduction

Language is always in flux and changes over time. How best to account for this fact is the domain of historical linguistics. From the 1970s on morphosyntactic change was theorized from a functionalist perspective primarily in terms of grammaticalization, whether the term was used or not (see, for example, Givón 2018[1979]; Heine et al. 1991; Traugott and Heine 1991; Bybee et al. 1994; Hopper and Traugott 2003[1993]). Constructional perspectives on morphosyntactic change began to be theorized in the 2000s and were pioneered in Bergs and Diewald (2008), Bergs and Diewald (2009a) and Barðdal et al. (2015). Inevitably, much work was and still is done comparing constructional approaches with approaches from grammaticalization (e.g. Coussé et al. 2018a). In this Chapter I outline aspects of Diachronic Construction Grammar as represented in Traugott and Trousdale (2013) and present some updating of our thinking. A brief comparison with grammaticalization is also provided.

All historical linguists pose the question "What changes?", but different approaches to change focus on different answers (Section 3.2). In Section 3.3 I briefly discuss widely mentioned mechanisms underlying change that will be referred to in later chapters, first the three most often discussed in the literature: neoanalysis (aka reanalysis), analogy, borrowing, and one that is specific to meaning change: pragmatic inferencing. Subjectification and intersubjectification are introduced separately in Section 3.4 as they are major factors in the rise of DSMs. Section 3.5 discusses the characterizations of constructionalization and constructional changes as developed in Traugott and Trousdale (2013) and updates those characterizations. Section 3.6 briefly compares the objectives and data of grammaticalization and constructionalization. Context plays a highly significant role in change and is discussed in Section 3.7. Section 3.8 summarizes.

3.2 What changes and how?

In this section I take the now fairly widely held functionalist position that change is the result of shifts in usage patterns, not of parameter shifts in universal grammar (3.2.1). In 3.2.2 I argue that change is not innovation, because change requires conventionalization, in other words, replication among a network of language users. Section 3.2.3 touches on the issue of gradualness vs. abruptness in change.

3.2.1 "Usage changes" vs. "grammar changes"

A fundamental point of debate in historical linguistics is what the basic assumption about change should be in historical linguistics. In the traditional generative view, it is that "grammars change" (Kiparsky 1968). From this perspective, the prime task of historical linguistics has been to account for ways in which fairly specific innate parameters are set, given primary linguistic input (e.g. Lightfoot 1991). However, more recently Biberauer and Roberts (2017: 143) have suggested that parameters may not be "pre-specified in the innate endowment". Rather, they emerge from the interaction of three factors identified in Chomsky (2005): underspecified innate endowment, primary linguistic data, and cognitive optimization principles.

By contrast, most constructionalists and functionalists in general argue that "usage changes", not grammars (see e.g. Croft 2000, Hawkins 2004, Bybee 2010), that is, change arises from processing and interpretation with only minimal, if any, assistance of innate grammar. As stated in Chapter 2.3, the position taken in this book is that knowledge of language is usage-based and grounded in "instances of a speaker's producing and understanding language" (Kemmer and Barlow 1999: viii). Historical data is necessarily based in evidence from textual records. Working with corpora, I take "usage-based" to mean "based in evidence from actual language use found in corpora".

I adopt the hypothesis that grammar accounts for linguistic knowledge and that this knowledge involves shared conventions and is therefore fairly stable. It is also dynamic and subject to change, emerging but not "intrinsically indeterminate" (see Traugott and Trousdale 2013: 47–48). This means that while I adopt a usage-based perspective, I do not identify it with emergent grammar in Hopper's sense as some researchers do, e.g. Diessel (2015). The hypothesis of "emergent grammar" (e.g. Hopper 1987, 2008), which is often referred to as an "emergentist approach", holds that structure and regularity arise out of discourse use, that grammar is "intrinsically unstable and indeterminate", and that the notion of a mentally represented set of rules needs to be set aside (Hopper 1987: 155). Most construction grammarians reject rules, but not mental representations. Hopper

(2011) makes a useful distinction between emergent and emerging phenomena. "Emergent" refers to the view that "a grammatical structure is always temporary and ephemeral" (Hopper 2011: 26), whereas "emerging" refers to the view that grammar that is dynamic but arises out of relatively stable phenomena. As Auer and Pfänder (2011: 15) point out, it is useful in a usage-based grammar to think of both "categorized linguistic knowledge" and use leading to innovation. Without some fairly stable structure, speakers would not recognize linguistic experimentation and improvisation (see also Harder 2012 for arguments in favor of emerging rather than emergentist models of grammar).

A usage-based view of change in Kemmer and Barlow's sense requires us to think of speakers and writers (SP/Ws) and addressees/readers (AD/Rs) doing things with language (Noël 2017). This means we should try to avoid reifying language using such concepts as constructions "prefer" some context. It is language-users who prefer. We should also avoid the notion of "competition". A *lot* and *much* do not "compete" before adjectives: SP/Ws prefer to use one over the other. It is worthwhile to recall Joseph & Janda's (2003: 79) injunction:

> given the transmissional discontinuity of language … across individual minds, it behooves us to resist the temptation to view particular linguistic constructions … as if they were organisms with lifespans larger than those of humans by several orders of magnitude (much less as entities independent of people).

However, it is almost impossible to avoid reification. "Language changes" is a useful shorthand, as is the ability to say that any aspect of a construction can change, although strictly this should be expressed as "Language-users may over time have a different interpretation of any aspect of a construction from their peers and predecessors".

If we recall Croft's (2001: 18) model in Figure 2.2 of Chapter 2.4, we may think of SP/Ws and AD/Rs bringing about changes on the form side of a construction: syntax, morphology, phonology or on the meaning side: semantics, pragmatics, discourse-function. They may even bring about changes in the symbolic linkage between the form and meaning components. In most cases of morphosyntactic change, such changes are preceded by changes in linguistic context. A basic constraint is that there should be a plausible path of interpretation from one step to the next (but see Section 3.3.2 below). Another constraint is that it is assumed that language-internal pragmatic processes are universal and do not themselves change. What changes is the extent to which the processes are activated at different times, in different contexts, in different communities, and to which they shape change.

3.2.2 Innovation vs. change

Each individual needs to interpret input and learn a language, so the likelihood that the knowledge of language that is acquired is exactly the same as that of others, especially older care-givers, is almost nil. Acquisition is a life-long affair (e.g. Bybee 2010; Sankoff 2019), especially in contact situations. Therefore:

i. Innovations constantly occur,
ii. Innovation occurs in the individual.

What then is the relationship of innovation to change? In my view, and that of most functionalists, innovations are not changes. Change is the development of shared, conventionalized usage (see, among others, Weinreich et al. 2017[1968]; Milroy and Milroy 1985; Schmid 2020). Change occurs only when certain innovations are adopted by groups of language-users. As Milroy and Milroy (1985: 347) say:

1. A speaker innovation may fail to diffuse beyond the speaker.
2. A speaker innovation may diffuse into a community with which he/she has contact, and go no further.
3. A speaker innovation may diffuse into a community with which he/she has contact, and then subsequently diffuse from that community into other communities … When the results of this process are observed, we tend to label the results as 'linguistic change'.

It is difficult to define a group or community, especially in historical work, as so the view that change only when innovations are adopted by groups of language-users has been debated. Every individual plays several roles in society and participates in several social networks. With the advent of electronic corpora and micro-statistical analysis, it is becoming possible to trace developments in individuals (e.g. Schmid & Mantlik 2015; Petré & Van de Velde 2018; Petré and Anthonissen 2020) and therefore it has become possible to track innovation, at least in contemporary data.

3.2.3 Gradualness vs. abruptness

Morphosyntactic change is typically gradual. Change is never A > B, but always A > A/B (> B), i.e. there is always variation for a while. A may never cease to be used, but the link between old and new may be weakened or lost. For example, *BE going to V* 'motion expression' has coexisted with *BE going to* 'future' from the 17thC on. The older and newer uses are "layered" (Hopper 1991). But Petré and Anthonissen (2020: 200) find distributional evidence for "dissociation" and divergence of motion and future uses from the earliest times. Some speakers nowadays may not link the

older and newer uses because future *BE going to V* is usually pronounced *BE gonna V*. Steps in change are currently considered to be tiny micro-adjustments, "sneaky" shifts (De Smet 2012: 608), that may be hardly noticeable in a corpus, or across individuals. Gradual change occurs in linguistic contexts, and these contexts themselves change. Focus on gradualness is in contrast to earlier work (e.g. Lightfoot 1979) that privileged "catastrophic" change and "saltations" or big jumps in the context of research that assumed large parameters in generative syntax. The shift in thinking is at least partially influenced by corpus work and quantitative studies (see also the current focus on micro-parameters, even features in Minimalism).

However, it should be noted that in the lexical domain, once a lexical word formation or idiomatic pattern has arisen, innovations based on it are usually instantaneous. For example, there is a highly productive and instantaneous word formation process N > V, e.g. *to calendar, to bicycle.* As Clark and Clark (1979) show, the N semantically constrains the way the V is understood. For example, the N may be interpreted as a locus on or into which something is put, e.g. *to calendar something* means 'to put something on/in a calendar', *to bottle something* means 'to put something in a bottle', not 'to put something on a bottle' of 'to use a bottle'. Alternatively, the N may be interpreted as an action using an instrument, e.g. *to bicycle/rake* means 'to use a bicycle/rake in order to X', not to put something in or on a bicycle/rake. Such expressions are understood in terms of both the linguistic context and knowledge of the world. When first used they are instantaneous innovations. They may not be taken up, e.g. *to porch* as in *The boy porched the newspaper* (Clark and Clark 1979: 787), has not become conventionalized, even though it easily understood as involving the action of delivering the newspaper onto a porch, probably by throwing it. But if innovations are taken up the process of conventionalization is gradual. Eventually speakers internalize them as members of the inventory of constructions, the constructicon. Clark and Clark (1979: 799) note that some different word-formations may become so entrenched that denominals based on N > V may be "blocked" or avoided, e.g. the existence of *hospitalize* preempts innovation of *to hospital* meaning 'to admit someone to hospital' (like all cases of entrenchment, such pre-emption is itself subject to change).

3.3 Mechanisms underlying change

Strictly speaking, since in this book "change" is defined as conventionalization or innovations, mechanisms discussed in this section are mechanisms of innovation that may underlie change. However, they are often referred to as "mechanisms of change".

Three mechanisms, neoanalysis, analogy and borrowing are often mentioned as the basic mechanisms of change (cf. Harris and Campbell 1995 on syntactic change) and are outlined in Section 3.3.1 together with frequency, which Bybee (2003) argues is also a major mechanism in grammaticalization, the process by which grammatical expressions come into being. Two mechanisms that have been associated with meaning change are inferencing (Section 3.3.2) and (inter)subjectification. As I will mention in Section 3.4. and discuss in detail in Chapter 11, although to date I have considered subjectification and intersubjectification to be mechanisms, thinking about them in diachronic constructionalist terms and in connection with the rise of DSMs suggests that they are not mechanisms, but actually default concomitants of constructionalization as procedurals.

3.3.1 Neoanalysis, analogy, borrowing, and frequency

Neoanalysis has been a foundational concept in historical linguistics. Originally the term used was "reanalysis", but since neither children nor second language learners can "re-"analyze a structure they have not internalized, it is a problematic term. "Neoanalysis" (Andersen 2001: 231, ft. 3) is preferable as a more neutral term. Work on neoanalysis asks: "How is expression X different from expression Y?". The question was originally asked mainly of form, e.g. bracketing, category assignment, etc. (Langacker 1977). Typical examples are:

(1) a. [[be going] [to V]] > [[be going to] [V]]
b. [[in] [deed]] > [indeed]

However, since the publication of Eckardt's (2006) important contribution to the study of meaning change as neoanalysis (called reanalysis in the book), semantic neoanalysis has also been studied in depth. Eckardt's examples include the development of *BE going to V*, negation markers in French, focus particles and determiners in German.

By contrast to work on neoanalysis, work on analogy asks: "How similar are expressions X and Y?", and on what grounds, e.g. form, meaning, both? A basic example is the matching of the form of various plural markers in earlier English to *-s* (plural of one of several noun classes in Old English), e.g.:

(2) *hund* 'dog' *hundas* 'dog-PL'
boc 'book' *bec* 'book-PL' (see *foot-feet*)
> *books*
dohtor 'daughter' *dohtor* 'daughter-PL' (see *sheep-sheep*)
> *daughters*

The importance of analogy in morphosyntactic change was for a long time downplayed as "too general a notion" (Lightfoot 1979: 361). But there has been a dramatic shift in thinking since the beginning of the 21st century. Bybee argued that frequency and repetition are major factors in grammaticalization, including "connections or associations of both phonological and semantic nature … among items" (Bybee 2003: 610). Drawing on Anttila (2003) and neurological research such as Pulvermüller (2002), Fischer (2007, 2011) argues that analogy is "a deep-seated cognitive principle that is relevant not only to language processing and language change but also to learning processes outside language" (Fischer 2011: 37; see also De Smet 2009), and therefore should play a significant role in historical morphosyntactic analysis. Fischer proposed that rather than demonstrating "catastrophic" reanalysis as Lightfoot (1979) suggested, the development of auxiliaries like *may, might, must* can be understood in part as the outcome of word order changes and especially analogization of auxiliary verbs (which ceased to have verbal inflections) to adverbial emphasizers like *definitely, actually*, amplifiers like *completely, quite*, and downtoners like *hardly, almost* which came in Early Modern English to be used before the main lexical verb (Fischer 2007: 197).

A distinction has been drawn between exemplar analogy in which one expression is matched to another (the exemplar) and constraints-based analogy in which optimization is a key principle (Kiparsky 2012). Exemplar analogy is strongly associated with constructional approaches to change, since it involves patterns and pattern match in the sense of a "best fit" match which allows for some properties in a "matched" pair to be different.[11] Match is typically not to an individual micro-construction but to a (sub)schema. This (sub)schema may be partially filled or fully abstract. For example, the match of [N plural] to $[s]_{\text{NPlCxn}}$ is partially filled as *s* is specified.

In my view, it is important to distinguish:

a. Analogical thinking, which language users do all the time (see "Humans are simply analogical animals", Anttila 2003: 438). It may enable change and is therefore a motivation for change.
b. Analogization, the modification of an *E* that is the result of such thinking (see Traugott and Trousdale 2013: 37–38). This is a mechanism underlying change and results in new uses of constructional components.

11. The principle of best fit "relative to the context" was developed in Winograd (1976: 280–282) to account in computational work for match of a conceptual entity with another that has approximately similar properties "relative to the current context" (p. 280) and has been extended to linguistic processing (see e.g. Hudson 2007: 48).

Further, each analogical change involving a specific microconstruction is an instance of neoanalysis, e.g. *boc* + vowel alternating plural > *boc* + general plural involves a new analysis regarding which plural marker is appropriate for this item. Compare *foot-feet* which was neither analogized nor neoanalyzed despite the analogical thinking that leads children to refer to *foots* (later *feets*) and presumably enabled the analogization change from plural *bec* to *books*.

As regards the third mechanism, borrowing, many of the homilies, histories and philosophical works that are strongly represented in the earlier textual record are translations from Latin or French: Latin in the Old English period, Latin and French in the Middle English period; Greek was an additional influence in EModE. Therefore the influence of Latin, French and Greek rhetorical traditions has been strong. With regard to pragmatic marker development, there is lexical borrowing of expressions such as *return, point* (from French). These are just two examples of the massive influx of loanwords from French in the Middle English period (Durkin 2014). They are loanwords that are substantive parts of monofunctional DSMs (cf. *to return to my point* in Chapter 9). It is probable that Latin models with pre-clausal connectives (see e.g. Kroon 1995; Rosén 2009) reinforced use in English of an already robust pre-clausal slot for epistemic adverbs like *witodlice* 'truly' (see Swan 1994) and were enabling factors in extension to non-epistemic, specifically conjunct adverbial expressions. It is also very probable that there was influence from French (see Ingham's 2015 study of some DSMs in Anglo-Norman). This, however, remains to be established.

Increased frequency is closely associated with change, but whether it is a mechanism and causal factor or a result is a matter of debate. Bybee (2003) argues that frequency contributes to grammaticalization. She distinguishes token frequency and type frequency. Token frequency refers to the number of times a particular expression *E* occurs, while type frequency refers to the number of instantiations of a category *E* may coocur with. For example, *the* is the most token frequent *E* in Present Day English texts, and its type frequency is high (most Ns in English). Bybee argues that the token frequency observed in studies of grammaticalization comes about as a result of the increase in types of context in which the *E* can occur (Himmelmann 2004 calls this "host-class expansion"), e.g. *BE going to* future increased in token and type frequency as the types of context in which it could be used expanded. Specifically, the lexical motion meaning does not occur in stative verb contexts, while the grammatical one does: *She is going to like the play* is not normally understood as an instance of the motion sense of the phrase. In the case of DSMs, token frequency expands by comparison with topicalized uses of the lexical source expressions, but not necessarily by comparison with uses of that source expression in default adverbial position. This phenomenon will be discussed further in Section 3.6 on contexts for change.

3.3.2 Pragmatic inferencing

It is now widely accepted that pragmatic inferencing in the flow of speech and writing contributes extensively to meaning change (see e.g. Hansen and Visconti 2009). Traugott and Dasher (2002) elaborated on my initial interest in the ways in which pragmatic meanings may come to be coded as semantic meanings (semanticized) and on a proposal called the Invited Inferencing Theory of Semantic Change (IITSC) (Traugott 1999). In Traugott and Dasher (2002: 35) it was proposed that:

> historically there is a path from coded meanings to utterance-token meanings to utterance-type, pragmatically polysemous meanings to new semantically polysemous (coded) meanings.[12]

In the particular formulation cited, context is not mentioned but rather assumed. However, close attention to context in corpus work suggests that utterance meaning arises in replicated contexts, for example, repeated use of literal *by the way* 'en route' represented as the context in which talk occurs by hypothesis strongly contributed to interpretations of the route as background to foregrounded talk (see Chapter 8.2). When such interpretations are replicated, the associated pragmatic meanings may lead AD/Rs in their role as SP/Ws to expand contextual uses. These expanded interpretations and uses may come to be conventionalized in the sense that they come not only to be commonly associated with an expression but also come to be used by a network of language users. In the chapters that follow, focus will be on the replicated contexts for change that as analysts we can find in corpora.

The goal of the IITSC is: "to account for the conventionalizing of pragmatic meanings and their reanalysis as semantic meanings" (Traugott and Dasher 2002: 35). This is too narrow a definition to be maximally useful for a usage based study of the development of pragmatic markers and is better stated as follows:

> Linguistic data provides evidence that over time language users may come to associate particular meanings with an expression *E* and later generations may interpret these as conventionalized meanings. By hypothesis, the mechanism that enables such changes is that in certain contexts speakers and writers implicate meanings and invite addressees and readers to interpret them, but addressees and readers may infer different meanings.

---

12. Hansen and Waltereit (2006) object that utterance-type meanings (GCIs) are not normally a factor in change. However, IITSC does in fact not refer to GCIs but to GIINs (utterance-type, pragmatically polysemous meanings), although a descriptive comparison to them is made. In the current book, conventionalization of replicated contexts is regarded as essential to many changes, and is well attested in the histories of some DSMs like *by the way* (see Chapter 8).

This statement is purposely broader than the original in order to cover the development of both conventionalized pragmatic meanings that are [–truth-conditional, +conventional] and semantic meanings that are [+truth-conditional, +conventional] (see Chapter 2.5).

While the SP/W has been privileged in the ITSC, "invited" is meant to highlight the fact that not only SP/Ws but also AD/Rs are involved in negotiation and change of meaning. Some researchers have focused on AD/Rs rather than SP/Ws (see Ehmer and Rosemeyer 2018 for a brief summary). For example, Heine et al. (1991: 70ff) refer to "context-induced reinterpretation", a term that highlights the role of AD/R. Detges and Waltereit (2002) emphasize the importance of listeners' strategies in meaning change in grammaticalization. Schwenter and Waltereit (2010) argue that some change is accountable for only in terms of AD/R's inability to accommodate default presuppositions in context. This results in the development of a new presupposition. One of their examples is use of *too* not as an additive but as an adversative refuting the propositional content of the previous speaker's utterance, as in:

(3) A: No, you can't!
B: I can, **too**!
(1916 Tarkington, *Penrod and Sam* [Schwenter and Waltereit 2010: 93])

The argument is that additive *too* as in *John had dinner in New York too* requires a presupposed first proposition such as *Mary had dinner in New York*, but if no prior proposition is available, or is too difficult to accommodate in context, the hearer may choose a simpler but novel interpretation (in this case refutation). A similar argument is put forward in Eckardt (2009). She investigates some other cases of scalar expressions such as *even* in which, she argues, presuppositions may be too hard to accommodate when new contexts do not support the older meaning. In such situations of "presupposition overload" or "failure" AD/Rs may interpret new meanings that do not arise by normal processes of implicature. Eckardt explicitly associates the changes with innovation but does not address in detail the puzzle that the same kinds of meaning shifts becomes conventionalized and occur cross-linguistically. None of the Discourse Structuring Markers discussed in this book involves parsing problems of the kind discussed above. However, the potential problem of presupposition accommodation serves as an important reminder that AD/Rs play a role in change and it is not always easy or perhaps possible to identify a direct coherent link between older and newer meanings, despite the constraint on analysis mentioned in Section 3.2.1 that "there should be a plausible path of interpretation from one step to the next".

Since change, not innovation, is under discussion, therefore conventionalization among a group of language-users, there is no question that AD/Rs as well as SP/Ws contribute to change, but in my view, it is only as speakers and writers that

addressees can pass on their innovation. My focus is on the outcome of SP/W and AD/R negotiation of meaning, as it is reflected in new uses of *E*s. Figure 3.1. summarizes the hypothesized process of innovation replication that leads to "manifest conventionalization" or change. Note that the first step involves a single speaker/writer and a single interpreter, and the second step involves more than one because a single innovative reinterpretation will not lead to change:

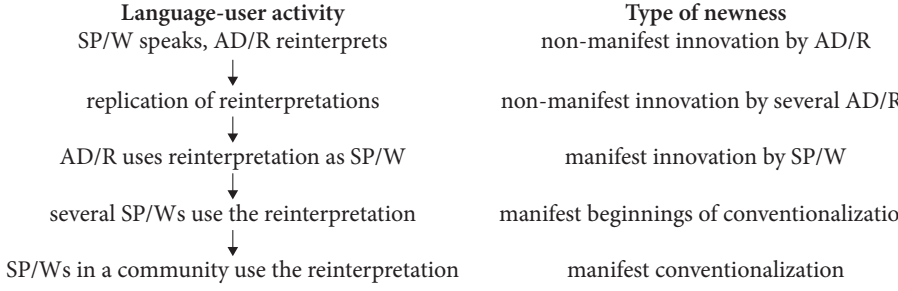

Figure 3.1 Steps from innovation to change

In Traugott and Dasher (2002) examples of invited inferencing were drawn from a range of meaning changes, some of which are contentful and truth-conditional, such as the rise of verbs that name speech acts of various kinds, for example *to promise*, which was borrowed in the early 15thC from French, but ultimately derives from Latin *pro* 'forward' – *mis-* 'sent' [past participle of *mittere* 'to send'). This had come to be used with the meaning 'to vow' in late Latin, perhaps in both non-linguistic contexts metonymic to vowing, such as sending a pledge, and in the linguistic context of passive syntax ('This message is sent forward as a pledge to X'). Another example is *to see* 'to experience visually' > 'to understand/experience cognitively').

However, most of the examples in Traugott and Dasher are procedural and illustrate the rise of [+conventional, -truth-conditional] meanings, e.g. *in fact*. The noun *fact* was borrowed from Latin in the 16thC. It derives from the past participle of *facere* 'to do', and means 'done thing, deed'. In EEBO it is used as a Circumstance Adverbial, sometimes paired with a noun like *right*, *law* or *deed* (itself derived from the past participle of *to do*) (4a) or contrasted with *fiction* (4b):

(4) a. all those that … doe walke with the wicked in betraying **in fact** and deed, that which their hart embraceth for veritie.
'all those that … walk with the wicked in betraying in their actions what their heart embraces as truth'.
(1583 Foxe, *Actes and Monuments* [EEBO])

b. our deliuerance from that intended mercilesse and matchlesse massacre both **in fact** and fiction, the fifth of Nouember, in the yeare 1605.
(1613 Boys, *Exposition of the last psalm* [EEBO])

The hypothesis is that AD/Rs interpreted the implicature of 'in truth' from *in fact* (also *in deed*) as the *E*'s chief meaning, especially in contexts of contrast with *fiction* such as (4b). Replicated reinterpretations enabled reinterpretation as an epistemic marker. This is a profile shift, from background (implicated) to foregrounded (expressed) meaning.

3.4 Subjectification and intersubjectification

Later changes undergone by *promise* and *in fact* suggest more than conventionalizing of inferences took place. Once *promise* had been interpreted as a speech act verb, it came to be used performatively with illocutionary force in first person, present tense contexts (*I (hereby) promise to X*). This is a change specific to some speech act verbs. Not all speech act verbs are used performatively, e.g. *threaten*. Both *promise* and *threaten* can be used as speech act verbs, typically with inanimate subjects, to express the speaker's epistemic belief (*It promises/threatens to rain*), but with different evaluations, one positive, the other negative (for cross-linguistic evidence of such uses see Heine and Miyashita 2008), but only *promise* can be used performatively. During the 18thC *in fact* came to be used as an epistemic adverb with contrasting, adversative implicatures, mainly in clause-medial position, and in the 19thC as an epistemic PM in pre-clausal position. These epistemic uses are instances of what I have called "subjectification". The development of discourse structuring Connectors exemplifies both subjectification and intersubjectification, as will be shown throughout this book and will be discussed at some length in Chapter 11. As mentioned in Section 3.3, I have in the past conceptualized subjectification and intersubjectification as mechanisms of change. However, they will be argued not to be mechanisms of change but rather communicative function processes.

Crucial to an understanding of the terms "subjectification" and "intersubjectification" as intended here is that they refer to change.[13] Subjectification is not the same as subjectivity, although it is grounded in subjectivity and results in subjectivity. Nor is intersubjectification the same as intersubjectivity, although negotiation of meaning always involves some degree of intersubjectivity (see especially Verhagen 2005, who sees "tailoring" of points of view with respect to other interlocutors as foundational to language). Attention to subjectivity in semantics has a long history going back at least to Bréal [1964[1900]). Theoretical approaches to subjectivity for the most part draw on Benveniste's (1971[1958]) observation that all language is

13. An alternative view of subjectification has been developed by Langacker (1990, 1999, 2006). For him it is largely a synchronic concept having to do with construal and covert subjects, for example, in *It is going to rain*, the locus of consciousness (the speaker) is "off-stage" and implicit.

both subjective and intersubjective. The speaker speaks (subjective) to an addressee who interprets (intersubjective). "Subjectification", on the other hand, as defined in Traugott and Dasher (2002) and here, is the development of meanings that are more based in the SP/W's perspective than earlier meanings. For example, *further* used as a DSM to mean 'in addition' is more subjective than *further* used as an adverbial of measure of distance (see Chapter 6.3.1), as it cues SP/W's perspectives on the relationship of D2 to D1. "Intersubjectification" is the development of meanings that are more centered on AD/R than earlier ones, for example, *by the way* used as a hedge mitigating the face-threatening content of D2 is more intersubjective than *by the way* used to mark D2 as just having come to mind (see Chapter 8.2). Both subjectification and intersubjectification are gradient and may implicate less or more intensity of attention to SP/W or to AD/R. Properties of (inter)subjectification will be discussed in greater detail in Chapter 11, along with the hypothesis mentioned above that they are not properly thought of as mechanisms, but as processes accompanying constructionalization of *E*s as procedurals.

One issue of debate is whether SP/W or AD/R is the agent of (inter)subjectification as a change. Hansen (2012: 602–603) suggests that both speakers and hearers play a role in subjectification. In the process of speaker-hearer dyads negotiating meaning, the hearer is interested in the speaker's point of view and is "predisposed to reinterpret linguistic items as expressing precisely that point of view". Pointing to the importance of hearers in change, Detges and Waltereit (2002) and Detges (2016) conceptualize hearers as "ratifying" new interpretations. As mentioned above, and modeled in Figure 3.1, my hypothesis is that it is only as a SP/W that an AD/R can ratify or manifest a reinterpretation. It follows that SP/Ws are the agents of subjectification. But by hypothesis, the path of development is indirect as it is likely that AD/Rs have made the inference that repeatedly encountered constructs are in certain contexts subjective. This does not exclude the possibility that an individual SP/W can innovate a subjective use directly. This is in fact likely in the case of use of speech act and cognitive verbs for performative use (e.g. *promise, recognize*). However, since as analysts we are looking for evidence of replicated use and gradual steps in change, small steps in subjectification need to be identified in terms of replicated constructs that are interpretable as expressing the subjective evaluation of AD/Rs acting as SP/Ws.

3.5 Constructionalization and constructional changes

Since form and meaning are combined in a construction Traugott and Trousdale (2013) proposed to use the term "constructionalization" (a term apparently first used in Rostila 2004 and taken up in Noël 2007), to refer to:

> the creation of form$_{new}$-meaning*new* (combinations of) signs. It forms new type nodes,[14] which have new syntax or morphology and new coded meanings, in the linguistic population of speakers. It is accompanied by changes in degree of schematicity, productivity, and compositionality (Traugott and Trousdale 2013: 22)

Since changes do not always result in an obvious new form-meaning pairing, we proposed a distinction between constructionalization and constructional changes and used the example of reduction of *BE going to V* to *BE gonna V*. Such changes were hypothesized to affect "one internal dimension of a construction" and to "not involve the creation of a new node" (Traugott and Trousdale 2013: 26).

In the years since the book came out these characterizations have been extensively discussed and tested by ourselves and others. The distinction between constructionalization as the coming into being of a form$_{new}$-meaning$_{new}$ pairing and constructional changes as changes that affect one component of a construction has been criticized by Börjars et al. (2015), Hilpert (2018) and Flach (2020) among others, primarily because what may appear to be change on a single dimension is often coupled with some other change. For example, the rise of the reduced form *BE gonna V* is associated with casual speech. If casual speech is regarded as a discourse function, then the emergence of *BE gonna V* should count as a new pairing of form and meaning. Some of the criticism is based in the theoretical assumption that ideally historical work should privilege innovation and the individual rather than the group or that tracking spread of a construction is a more valid object of research than its coming into being, partly because it is more amenable to operationalization in corpus studies. Nevertheless, the points are well taken and the original definitions need reworking.

A particularly interesting criticism is raised by Flach (2020). She argues that there are two readings of "constructionalization": one is a process reading suggested by "accompanied by changes in degree of schematicity, productivity, and compositionality", the other is a point/result reading associated with "creation" of form$_{new}$-meaning$_{new}$ signs and nodes (Flach 2020: 48). She suggests that constructionalization as characterized in the quotation from Traugott and Trousdale above "refers simultaneously to constructional changes surrounding the new node and the new node itself". Her concern is that what she calls "con-constructionalization constructional changes" (Flach 2020: 47) muddies the distinction between constructionalization and constructional changes, between process and point change. She proposes that "constructionalization" be reserved for the point reading, which entails that constructionalization is instantaneous, and "constructional changes" for

14. "Nodes" in construction grammar are constructions with links to other constructions (see Chapter 13).

the process reading, which may be gradual. However, this proposal is problematic because "point" implies instantaneous change and this can be identified only for individual SP/Ws (manifested individual texts). This links constructionalization to innovation, not to the rise of a conventional pattern of the type that can be identified as a change, which necessarily emerges gradually as it involves conventionalization across interlocutors (see Sections 3.2.2 on change vs. innovation and 3.2.3 on gradualness vs. abruptness above). In Flach's view, the distinction she proposes avoids "arbitrary distinctions of pre-, con-, or post-constructionalization change" (Flach 2020: 49). A distinction between pre- and post-constructionalization and constructionaliztaion is, however, not arbitrary. They are different in ways to be discussed below.

While the distinction between constructionalization and constructional changes is admittedly somewhat difficult to maintain when sophisticated quantitative corpus analysis is engaged in, for a relatively coarse-grained analysis such as is used in this book, it is helpful in identifying approximately when and how a new construction has come to be part of some subset of language-users' knowledge of the language. But the concepts as articulated in Traugott and Trousdale (2013) need to be modified. In what follows I provide new characterizations of constructionalization and the kinds of constructional changes that enable it and that follow after it, drawing on Trousdale and Traugott (2021). These characterizations respond to some of the criticisms and also address the kind of pragmatic data that is the topic of this book. Before presenting the revised characterizations an explanation of our objectives may be useful. This will be followed by a brief review of the concepts "compositionality", "analyzability", "schematicity", "productivity" and "coded meaning" that appeared in the definition cited above.

Our objective is to focus on how a new form-meaning pairing comes into being, using a constructionalist usage-based model. The focus is on internal language change (see Chapter 1.2), as the grounding on which interactional, sociolinguistic or psycholinguistic accounts can be built. One example of the kind of change Trousdale and I have in mind is the development of the *BE going to V* future. In Middle English *BE going to V* was used to express motion with a purpose. But a relative future came to be entrenched in the early 17thC within the abstract Auxiliary. Schema. We analyze this as a constructionalization (a new [[Form] ↔ [Meaning]] pairing that is added to the constructicon). In the century after the first appearances of relative future *BE going to V*, there is evidence that SP/Ws began to use it in contexts that cumulatively led to the rise of the deictic future (see Petré 2019). This is a constructional change because no new construction has been added to the constructicon; an extant one has been modified. Another example is that in earlier Old English there was no article as we understand that term now, but a definite article originating in the demonstrative *þæt* came to be firmly entrenched by late

Old English around 1100. Corpus data attest a number of examples that suggest SP/Ws started to use the demonstrative *þæt* in ways that led cumulatively during Old English to the development of a new definiteness marker *the* within a Determiner construction (Sommerer 2018). This was a constructionalization. Later conventionalized pragmatic uses such as distancing from a plural referent as in *the Republicans* vs. *Republicans* (Acton 2019) are constructional changes.

As discussed earlier in this chapter, innovations by individuals are manifested in the constructs they use (the "textual evidence"). These are tokens that exemplify both recapitulation of *E*s that have been used before and modulations of *E*s that have already been heard. Over time, new types of *E*s may appear in the textual record. These new types are conventionalized "micro-constructions", form-meaning pairings that generalize over accumulated constructs. They are conventionalized in the sense that some individuals in a community share knowledge of the new micro-construction and produce constructs that are unambiguously different in meaning and distribution from earlier expressions. In practice this means that in a corpus at least two different authors need to be found using the new pairing, but preferably more.

"Compositionality" has to do with the degree to which the meaning of a construct can be understood on the basis of its parts. Over time, and as a result of processes like chunking, idiomatization and metaphorization, compositionality may be reduced. The internal parts of the construct may come to be less identifiable and therefore less analyzable (e.g. Bybee 2010; Vincent 2015). Both Bybee and Vincent cite the example of suppletion: compositionally, *went* means *go* + PAST, but it is not analyzable. Examples of non-compositional but analyzable *E*s are uses of the French present perfect to express the past, as in *Pierre a visité Manchester l'année derniére* 'Pierre visited Manchester last year', where a periphrastic present perfect tense form *a visité* is used to express past tense. The form *a visité* (literally 'has visited') is analyzable, but it is not compositional because the meaning of the present perfect is not past (Vincent 2015). Finally, some constructions may be non-compositional and non-analyzable. Börjars et al (2015: 375) cite English *though* and its German cognate *doch*, which "can be shown to be the outcome of convergent diachronic trajectories on both the form and content side < PIE *to- + k^{w}e 'and in that case'". Changes in compositionality usually result in decrease in structural and semantic accessibility. However, in a few cases increases also occur. These are known as "folk etymologies": a semantically opaque structure is given a structure that appears to be less opaque, e.g. neoanalysis of Old English *bryd guma* 'bride man' as *bridegroom*.

The term "schematicity" has been used in two senses. In the grammaticalization literature it often refers to degree of generalization over specific items (Coussé et al. 2018b: 7). In the constructionalist literature, however, it is usually used to refer to the availability of open slots (Bybee 2010; Coussé et al. 2018b: 9). These slots can

be instantiated by a variety of words and phrases. *E*s discussed in this book can be used in a Connective.Schema which has the schematic form [discourse segment 1 ___ discourse segment 2] ("D1__D2" for short) (see further Chapter 4). The slot is filled by the Connector.Cxn, which licenses several *E*s the discourse function of which is to signal the relationship between D1 and D2 (elaboration, contrast, etc.). A schema can be studied in terms of the paradigmatic and syntagmatic dimensions.[15] Paradigmatically, the questions is how many and what kind of options or micro-constructions are available for any given (sub)schema. For example, in PDE expressions of the future include *will V, shall V, BE to V, BE going to V.* Nesselhauf (2010) traces the frequency with which these options have been used since 1600 and argues that in addition certain uses of the progressive and of a separate auxiliary *ll V*, which neutralizes differences between *shall* and *will*, should be included among the options. I will discuss the changing paradigm of Elaboratives in Chapter 6. Syntagmatically, the question is to what extent a slot licenses new combinations, or is "type productive" (Barðdal 2008). If a construction becomes more schematic and covers a wider range of situations, it follows that "a wider range of lexical items are compatible with the schema" (Perek 2020: 147). This is referred to as the "extensibility" of a construction in Barðdal (2008). Changes in schematicity and productivity tend to involve increase, but constructions may also be lost, reducing both schematicity and productivity, as will be exemplified in Part II.

"Coded" meaning is often understood as truth-conditional, "literal" semantic meaning (e.g. Sperber and Wilson 1995[1986]; Frawley 2013[1992]: 11). This is a book about Discourse Structuring Markers, the meaning of many of which is for the most part conventionalized, non-truth-conditional and pragmatic. Therefore "coded" meaning is too restrictive in the original definition of constructionalization for present purposes and for other domains of study, such as performative use of speech act verbs. "Conventionalized" meaning is less restrictive and captures the idea of replicated, shared meaning change.

Turning now to new characterizations of constructionalization and constructional changes, Graeme Trousdale and I propose to characterize constructionalization as follows:

> Constructionalization is the establishment of a new symbolic link between form and meaning which has been replicated across a network of language users, and which involves an addition to the constructicon.

15. The distinction between the "paradigmatic" and "syntagmatic" axes, between choice and combination, referred to here is central to work by Saussure (1983[1916], Jakobson (1956) and Lehmann (e.g. 2015[1995]). A related pair of concepts usually used with reference to meaning change in the lexicon, is the distinction between an onomasiological and a semasiological approach (see Chapter 1.4).

This new characterization is shorter than the definition offered previously in Traugott and Trousdale (2013) and cited above. It is meant to highlight conventionalization of a pairing of form and meaning and storage of that conventionalized pairing in the constructicon. "Establishment" differs from Flach's "point" because the outcome is interpreted as the result (however temporary) of changes involving gradual spread across members of a community, not innovations. The characterization also differs from that proposed by Smirnova (2015), which is that constructionalization is a process of gradual accumulation of contextual restrictions prior to structural reorganization of material such as form-meaning change. In our view, these processes lead to constructionalization (see below on pre-constructionalization), but do not constitute it.

Constructionalizations are the result of communally shared generalizations over constructs. Such generalizations give rise initially to micro-constructions. In Chapter 4 we will see how two separate microconstructions, temporal *after* and pronominal *all* 'everything', are generalized over when they cooccur in certain contexts and are reinterpreted as a unit with Connector function. In Chapter 8 we will see how three microconstructions *by, the* and *way* meaning 'on the road', are generalized over and similarly reinterpreted as a unit with Connector function. Over time generalizations may be made over microconstructions with similar function, leading to the development of a schema, as in the case of Digressives: *by the by* came to be used in addition to *by the way* to signal digression from the main argument; by hypothesis SP/Ws generalized over the two and a digressive subschema of Connectors was formed. This subschema expanded when *incidentally* also came to be used as a Digressive (see Chapter 8).

Constructionalization of most *E*s and schemas is the gradual outcome of multiple small changes.[16] These pre-constructionalization changes, which are tendencies and not determinative, often involve the following effects on a construction A that is undergoing change. In English the relevant A in the development of a DSM is usually a Circumstance Adverbial.

1. chunking and some loss of compositionality within construction A,
2. some loss of analyzability within construction A,
3. adjustments that involve replication of semantic content or syntactic contexts that are only relatively loosely connected with the emerging new construction.

16. However, some constructionalizations based on word-formation schemas arise instantaneously, without pre-constructionalization changes leading to use of the particular new micro-construction, for example those built on word formation patterns such as X-ization, X-ity, where X is an Adjective (see Traugott and Trousdale 2013: Chapter 4; Hüning and Booij 2014).

They include changes in routinized assemblies of constructions (see context type d) in Section 3.6 below) and belong to a "co-text that … is generally quite large" (Budts and Petré 2020: 331) or to fixing of strings,

4. increased frequency of assemblies of co-textual shifts.

All of these are exemplified in the history of Connectors in Part II.

Post-constructionalization, in a process of propagation across linguistic contexts known as "actualization" (see e.g. Andersen 2001; De Smet 2012), the modifications are on the whole more local than those prior to constructionalization (Budts and Petré 2020). Again, post-constructional changes are tendencies. The new construction B often undergoes:

a. collocational expansion (spread to new immediate environments) (Hilpert 2013), such as loosening of restrictions on V and on the subject in the case of *BE going to* 'future', or, in the case of Connectors, being used with scope over the following clause; this is increase in type frequency,
b. increase or decrease in frequency of use, in other words, change in token frequency,
c. generalization over microconstructions and alignment to other microconstructions may lead to incorporation into a more abstract, schematic type-node. Examples are the alignment of *after all* and *by the by* with extant Discourse Structuring Markers,
d. in the case of decrease of use, failure to learn and use a construction or failure to add new members to a schema, may lead to obsolescence,
e. morphophonological reduction of the type discussed extensively in the grammaticalization literature, e.g. reduction of Old English *eal swa* 'exactly like this' > *also* (in this case, loss of both compositionality and analyzability).

The development of Discourse Structuring Markers exemplifies all of these kinds of post-constructionalization changes.

On this view, constructional changes can be characterized as:

> Modulations of contextual uses prior to and following constructionalization; they do not involve additions to the constructicon.

Constructional changes include language-internal changes in the features that make up a construction in the sense of Croft's (2001) model (e.g. in the case of *BE going to V* a shift from relative to deictic tense, or phonologically to *gonna*) and also contextual shifts prior to and post constructionalization.

3.6 A brief comparison of work on constructionalization and on grammaticalization

Grammaticalization has been referred to several times in the preceding pages. Here I briefly compare the main objectives of work on constructionalization and grammaticalization, and the relevant domains of study. Both terms refer to the research approach and to the processes studied. The comparison between constructionalization and grammaticalization is explored in great depth in Coussé et al. (2018a). Furthermore, a grammaticalization approach to DSMs is discussed in Chapter 5, so only a few key points pertaining to differences and similarities in approach will be mentioned here. The most important point is that the two approaches are complementary as they ask different questions and in some cases investigate entirely different data.

As mentioned at the beginning of this chapter, functionalist approaches to morphosyntactic change were theorized primarily in terms of "grammaticalization" from the 1970s on. While many historical linguists in the 19thC were well aware of the tendency of grammatical items to be derived from lexical ones, and had identified "bleaching" or loss of content meaning as a factor in the change, it was Meillet who gave the phenomenon the name "grammaticalization" in 1912. He characterized it as: "l'attribution du caractère grammatical à un mot jadis autonome" 'the attribution of grammatical character to an erstwhile autonomous word' (Meillet 1958[1912]: 131). The processes involved in this kind of change have been considered to include various kinds of reduction, including what is widely known as "bleaching" (loss of "semantic integrity"), loss of syntactic freedom especially within a phrase ("univerbation" in the grammaticalization literature, "chunking" in the construction grammar literature) and syntactic scope reduction. The hypothesis that unidirectionality is irreversible (Haspelmath 1999) was central to research through the end of the 20thC. However, irreversibility was challenged, for example in Campbell (2001), where several examples of "degrammaticalization", the reverse of unidirectional change, were proposed. Haspelmath (2004: 21) grants that some counterexamples exist but argues that "it is undeniable that the unidirectionality of grammaticalization is by far the most important constraint on morphosyntactic change". Norde (2009) is a detailed study of the examples proposed in Campbell (2001) and elsewhere, testing them against parameters of grammaticalization put forward in Lehmann (2015[1995]) (these will be discussed in Chapter 5.2 with respect to the development of DSMs). In her 2009 book Norde showed that while many of the examples proposed are not genuine reversals (also Haspelmath 2004's point), nevertheless there are enough cases to confirm that unidirectionality is not irreversible.

One of the major research questions addressed in grammaticalization studies has been how lexical items give rise to grammatical *E*s such as tense, aspect, modality, case and complementizers (see e.g. Heine et al. 1991; Traugott and Heine 1991; Hopper and Traugott 2003[1993]; Lehmann 2015[1995]; Kuteva et al. 2019). Another has been how "free", discourse-based word order comes to be linearized, fixed and syntactic. The conclusion has largely been that fixing of word order is not a type of grammaticalization, but rather the result of several factors such as on-line discourse organization and contact as well as processes of grammaticalization (see e.g. Hawkins 1983; Heine and Reh 1984).

Most early practitioners of Diachronic Construction Grammar and especially constructionalization in the first decade of the 21stC, had already worked on grammaticalization, so many papers on Diachronic Construction Grammar were devoted to rethinking work on grammaticalization (see e.g. Bergs and Diewald 2008). However, this is not true of the first paper taking a diachronic constructionalist perspective (Israel 1996). Israel studied the development of the *way* construction, and its expansion from expressions like *wend one's way to X* through *force one's way through X* and eventually *giggle one's way through X*. The interest here is basically in what types of verb classes came to be used in the *way*-construction, one of many three-argument constructions discussed in Goldberg (1995). This was an early study in what we now think of as collexemic expansion.

Diachronic Construction Grammar, especially the branch known as constructionalization, asks the question: How do constructions, including schemas, come into being, and what mechanisms are involved? This means that both form and meaning are attended to equally in the research. This is different from most grammaticalization research, where the practice has been to focus mainly on either form or meaning (while always recognizing the importance of the other). Constructions include tense, aspect, modality and case constructions, so there is substantial overlap in the domains of study. However, Diachronic Construction Grammar also studies domains where grammaticalization is irrelevant, such as the development of ditransitives and "dative alternation" (e.g. *Beth gave Kim a puppy, Beth gave a puppy to Kim*) (Zehentner 2019; see also Wolk et al. 2013 on the alternation).

In this respect constructionalization covers far more than grammaticalization. But in some respects it covers less. Because to date there is no adequate constructionalist theory of phonology, it has not been possible to analyze phonological change well from a Diachronic Construction Grammar perspective, except where meaning is involved, which is usually a matter of prosodic pattern. So there has been little substantive discussion of late-stage grammaticalization phenomena, where there may be extensive phonological reduction (such as *BE gonna V* and further reductions).

So among major differences are subject matter and coverage. More importantly, there are considerably different mindsets, in that many researchers on grammaticalization focus on reduction, even if the irreversibility of unidirectionality is questioned. Constructionalists, on the other hand, embrace pattern and context expansion, incorporation into larger networks and schemas (see comments on increase in schematicity in Section 3.5 above). When researchers work on grammaticalization, they have until recently been concerned primarily with difference and the mechanism of neoanalysis. By contrast, historical constructionalists work extensively with similarity, analogy and pattern match (but, as I pointed out in Section 3.5, particular instances of analogization are in fact neoanalysis). Another major area of difference is that much work on grammaticalization is founded in typological work (see Kuteva et al. 2019), almost none is in Diachronic Construction Grammar. This is in part because Bernd Heine and Christian Lehmann, both early theorists of grammaticalization, are typologists. It is also in part because Croft (2001: 32–34) questioned the cross-linguistic validity of many categories "based on the well-known and pervasive grammatical diversity of languages" (p. 33).

Areas where grammaticalization and constructionalization overlap are that both address the development of grammatical units, frequency, productivity, bleaching and other kinds of reduction, e.g. obsolescence. Himmelmann (2004) argued that the result of reduction, for example, to an affix, is expansion of the contexts in which it can occur (a version of Bybee's "type expansion"), and therefore grammaticalization is in essence the study of context expansion. This is consistent with constructionalist work on collexemic analysis, e.g. Hilpert (2008, 2013).

Simplifying even more, the complementary foci of work on grammaticalization and constructionalization can be summarized as in Table 3.1 (A "gram" is a minimal grammatical unit. "Cxn" is short for 'construction'):

Table 3.1 The complementarity of work on grammaticalization and constructionalization

Foci of Grammaticalization	**Foci of Constructionalization**
How do grams arise?	How do constructions arise?
Form or meaning change	Form and meaning change
Reduction of components of an *E*	Changes to components of a Cxn
Unidirectional reduction	Increase in schematicity
Expansion of contexts	Expansion of contexts
Changes in frequency	Changes in frequency

3.7 Contexts for change

As characterized above, constructional changes are largely contextual changes. Context has been a matter of prime importance in recent years in various lines of work. As Harder (2012: 525) noted, "Cognitive Linguistics arose as a recontextualizing movement (cf. Geeraerts 2003), after the structuralist century of removing language from its anchoring in the rest of the world". Context has been central in historical work as well. In a much-quoted statement about meaning change in grammaticalization, Bybee et al. (1994: 297) said: "everything that happens to the meaning of a gram happens because of the contexts in which it is used". Bergs and Diewald (2009b: 1) seek to investigate the role of contexts as "loci for change". As analysts we cannot exhaustively identify what contexts, including mutual cognitive environments, a particular locutionary participant pays attention to in a particular situation, even when studying contemporary data (Fetzer 2012). This is particularly true where historical texts are concerned. It is also difficult to analyze what is called the "context of the situation" (e.g. Halliday and Hasan 1976), for example who is speaking, where, to whom and in what manner. This kind of context is the specialty of historical sociolinguists (see e.g. Hernández-Campoy and Conde-Sylvestere 2012). Contexts that can be gleaned from corpora are primarily linguistic, specifically the co-texts surrounding an expression. For practical reasons, this is what is meant by context here, although I recognize that ultimately any interlocutor's understanding of co-text cannot be totally separated from context of situation, which may be political and ideological as well as cultural (see e.g. Fried and Östman 2005; Terkourafi 2009; Misković-Luković and Dediać 2012). Like situational contexts, linguistic contexts bias toward interpretations. *Promise* is illocutionary primarily in the context of first person subject *I* and present tense. These contexts bias toward the interpretation that the construct *I promise to X* functions as an act and is non-truth-conditional, whereas a past tense or third person context as in *I/she promised to X* biases interpretation toward a truth-conditional and non-illocutionary interpretation.

Working with Word Grammar, a cognitive model that has much in common with construction grammar, Hudson (e.g. 2010: 89) points out that speakers attend only to active, relevant contexts. From a historical perspective this raises the question which ones are or come to be relevant over time and whether any predictions about likely contexts can be made. The contexts that are identified as important for change and that are identified in this book are linguistic, replicated and conventional. Meanings may arise on the fly. For example, *I am going to get a haircut* could be contingently understood as a boy's promise in the context of a parent fussing about her son's unruly hair, but it is not conventionally one. Finkbeiner (2019b: 184) calls such meanings that contingently arise in context "context-dependent meanings".

They are akin to Grice's (1989[1967]: 37) and Levinson's (2000: 16) "particularized conversational implicatures" (PCIs). They do not enable change unless they are replicated, conventionalized and come to be available context-independently.

According to Bergs and Diewald (2009b: 2–3) "[t]he notion of context is located in the overlapping area between pragmatics and discourse". Three broad interacting types of linguistic context have been considered especially important in constructional change (e.g. Traugott 2018b; Zehentner 2019) and will be addressed in the case studies:[17]

a. The linear flow of speech and writing and replicated co-occurrences on the axis of combination and syntagmatic relations. In constructional terms, the question is what particular sets of co-occurrences (collocations) of expressions appear to be relevant in a particular change.[18]
b. The alternatives available on the axis of similarity and choice ("paradigmatic" relations). In constructional terms, the question is how schemas are instantiated and how this changes over time.
c. Systemic changes such as word order changes. Such changes are not particular to any specific construction. They may enable change prior to constructionalization, as in the case of dative alternation, which Zehentner (2019) links to the development of periphrastic case in late Middle English. They may support ongoing changes after constructionalization, as in the case of benefactive alternation (Zehentner and Traugott 2020). Or they may have no significant affect before or after constructionalization, as in the case of the Great Vowel shift that led to phonological changes.

Particularly important is an additional type of context that has only recently been discussed:

d. Assemblies of discursive uses of constructions (Petré 2019). As mentioned in Chapter 2.3, from a constructionalist perspective such as the one articulated in Goldberg (2003), an utterance manifests unification of several micro-constructions. *What did Liza buy the child?* unifies not only the substantive micro-constructions *Liza, buy, the, child, what, did*, but also the schematic ditransitive, interrogative, subject-auxiliary inversion, VP and NP

17. The distinction between the "syntagmatic" and "paradigmatic" axes referred to below are central to work by Saussure (1983[1916], Jakobson (1956) and Lehmann (e.g. 2015[1995]). A related pair of concepts usually used with reference to meaning change in the lexicon, is the distinction between a semasiological and an onomasiological approach (see Chapter 1.3).

18. Statistical analysis of collocations can be conducted in terms of "collostructional analysis", the study of patterns of syntagmatic combinations and the degree to which items are "attracted" to (cooccur with) a construction, and how this changes over time (see Hilpert 2008, 2013).

constructions (Goldberg 2003: 221). Thinking this way helps the researcher notice replication of certain larger contextual and constructional uses prior to constructionalization, in addition to the more local contextual collocations typically found after constructionalization (see Section 3.5 above).

The model of change espoused here (A > B/A (> B)) in which A and B coexist reflects the hypothesis that there will be polysemy at least briefly.[19] The hypothesis of polysemy suggests that interpreters may find some utterances ambiguous. Which meaning is to be interpreted in a context that does not provide adequate cues? In early work on generative syntax, grammaticalization and especially reanalysis, it was proposed that ambiguity was a prerequisite to change (e.g. Timberlake 1977 and Haspelmath 1998 on reanalysis) and frequent mention was made of "bridging contexts" (e.g. Evans and Wilkins 2000; Heine 2002), or "critical contexts", (Diewald 2002), both of which are defined as ambiguous. More recently the importance of ambiguity prior to change has been questioned. Focus on the role of clusters of subtle contextual modulations in change has established that ambiguity is not a necessary factor in change, although it may be found in textual data. It has in fact been shown that ambiguity often arises after rather than before change (e.g. Detges and Waltereit 2002: 170, Detges 2016).

Finally, a reminder about speech and writing, which were mentioned in (a) above and Chapter 1.5. This is a very broad distinction that glosses over register (see e.g. Biber et al. 1999). "Speech" can range from casual conversation to public speech such as presentations to institutions like Parliament or Congress. "Writing" generalizes over practices ranging from letter writing to newspaper reporting, academic writing and fiction. While speech is without doubt a main factor in change, the role of writing in (semi-)literate societies cannot be ignored. In particular, "speech-purposed writing", or writing intended to be spoken, as for example translations of the Bible and homilies (Culpeper and Kytö 2010) clearly had significant influence on not only the written textual record but also on speech. There has been an increased use of colloquial, spoken features of language in fiction and of economy and condensation for purposes of communicating information in science writing and to some extent in newspaper writing (Biber and Gray 2011). Such factors are considered "external" to broad kinds of developments outlined in the present book.

19. An alternative hypothesis of monosemy has been developed primarily in work on Relevance Theory, e.g. Sperber and Wilson (1995[1986]) and Carston (2002). Monosemy is problematic for historical work because it does not readily allow for revealing distinct new meanings (but see Nicolle 2011 on a Relevance Theoretic approach to grammaticalization).

3.8 Summary of key points

To summarize, I have argued that what changes is usage. Therefore interlocutors play a major role in change. Innovation is not change. Change is brought about when expressions are used in replicated contexts and are eventually conventionalized with aspects of form and meaning that differ from those of an earlier generation or different community. Those contexts of particular interest here are linguistic, especially assemblies of constructional uses prior to constructionalization, and pragmatic implicatures that come to be conventionalized through replication. Most change is gradual, progressing sneakily through constructional space. Analogical thinking is a prerequisite for many changes, which are brought about by analogization. All cases of analogization are neoanalyses. Meaning changes often involve the conventionalization of invited inferences. Some meanings may come to be subjectified and in some instances these subjectified meanings may come to be intersubjectified.

New characterizations of constructionalization and constructional changes were proposed and the main objectives and domains of study in work on constructionalization and grammaticalization were compared.

In the next chapter I characterize Discourse Structuring Markers, what their typical sources are, and exemplify with the history of one marker, *after all*.

CHAPTER 4

Discourse Structuring Markers and some generalizations about how they arise

4.1 Introduction

In this chapter I provide an introduction to what is usually meant by "Pragmatic Markers" in general, including "Discourse Markers" (4.2). In Section 4.3 I elaborate on the characterization in Chapter 1.2 of the subset of Connectives that are "Discourse Structuring Markers" (DSMs), the topic of this book. By way of reminder, I understand DSMs to be procedurals that signal the intended relationship between discourse segment 1 (D1) and discourse segment 2 (D2).[20] Multifunctional DSMs are DMs. Some generalizations about how DSMs, including DMs, arise in English are presented in Section 4.4. The history of *after all* is outlined in Section 4.5 to illustrate the theoretical points in prior sections. Section 4.6 summarizes.

4.2 Pragmatic Markers

As mentioned in Chapter 1.3, many terms have been used for pragmatic markers. In her ground-breaking book, *Discourse Markers*, Schiffrin (1987) analyzed use of the pragmatically interpreted expressions *and, because, but, I mean, like, now, Oh, or, so, then, y'know, well* and called them "Discourse Markers". These expressions include the mainly epistemic marker *y'know* and the mainly social marker well, as well as the Discourse Structuring Markers and, *but, now, then*. The 11 functionally rather disparate markers that Schiffrin discussed have also been referred to as "pragmatic particles" (e.g. Schourup 2016[1985]; Foolen 1996; Fischer 2006), "pragmatic markers" (e.g. Fraser 1996) or "discourse particles (e.g. Aijmer 2002). They are all associated with conventional pragmatic meaning (Haselow 2019), including *Oh* (see Heritage 1994 *et passim*), *Uh* and *Um* (see Tottie 2014).

---

20. This is an idealized abstraction. In some cases of actual production, either D1 or D2 may be absent (see Blakemore 2002 for discussion). Inferential *so* is well-known to be usable at the beginning of a discourse, provided that relevance to the discourse situation is inferable. This can be considered to be a special property of *so*. In my data, when D2 is absent (mainly in the SPOKEN component of COCA) there has usually been an interruption.

I follow Fraser (1996) in identifying an "umbrella" category of "Pragmatic Markers" (PMs) with several subcategories. In PDE there are several dozen PMs (under various names). For lists Fraser (1988, 1996, 2009a) and Aijmer (2002).

PMs signal what the speaker/writer (SP/W) intends the addressee/ reader (AD/R) to understand regarding such factors as:

1. SP/W's (represented) social relationship with the interlocutor (*please*),
2. SP/W's degree of epistemic certitude regarding the upcoming discourse segment (*I guess, of course*),
3. The relationship between discourse segments that SP/W intends to convey (*and, also, but, by the way*).

This third type of Pragmatic Marker is what I call a "Discourse Structuring Marker". The particular markers cited in 3. are multifunctional and are what I call Discourse Markers. They are Connectors and are syntactically Conjuncts (Quirk et al. 1985).

Characteristics of Pragmatic Markers are outlined in 4.2.1, and of Discourse Markers in particular in 4.2.2.

4.2.1 Characteristics of Pragmatic Markers

Many characterizations of the umbrella category called PMs here have been proposed (often with different names), among them Schiffrin (1987), Foolen (1996), Fraser (1996), Aijmer (2002) and Brinton (2017). It is generally agreed that, among other things, PMs are:

a. non-truth-conditional,
b. contextualizing cues and processing instructions about how to interpret the host clause (procedural),
c. multifunctional,
d. (inter)subjective,
e. not syntactically integrated with the host clause,
f. often mobile, i.e. usable in several positions in the clause; typically these positions are pre-clausal, post-clausal, and medial, but not all PMs occur in all positions (see Chapter 12),
g. short,
h. set off by a prosodic envelope (or comma in writing).

Haselow (2019: 3) regards PMs (which he calls DMs, following Schiffrin 1987) as "cues for processing". They "contribute to the development of a coherent mental model of ongoing interaction and discourse by integrating single units of discourse into a unified whole".

4.2.2 Characteristics of Discourse Markers

Fraser (1996 and elsewhere) restricts the term "Discourse Marker" to the subset of PMs that "can signal the intended relationship between adjacent discourse segments" (Fraser 2009a: 893). These are metatextual markers in the sense that they signal text relationships construed by SP/W. Fraser typically recognizes DMs only in pre-clausal position. Therefore, although DMs as Fraser uses the term have most of the characteristics listed above for PMs, they are not mobile. In this book, use in any position is recognized, provided the expression *E* in question marks a connective relationship between discourse segments.

For example, in (1), *by the way* is used in the typical pre-clausal position. Here it is conventionally understood as a marker of digression, a move away from a straight line of argument, in this case introducing an interruption to make an explanation.

(1) let's go back to the TNA machine. The TNA machine- *FRANKEN*: **By the way**, let me explain what that is 'cause it keeps coming up.
(1990 *CNN_NewsSun* [COCA])

Pre-clausally it is used with other functions as well, like hedging and topic-shifting, as will be discussed in Chapter 8.2, but it is to be contrasted with contentful *by the way* meaning 'along the road/path', an expression only rarely found clause-initially and topicalized in Present Day English. By contrast, when used post-clausally *by the way* has less pragmatic effect, often implicating that the preceding clause should be understood as having been presented as an after-thought.

(2) Did you suspect that he might have fooled around with the models? That doesn't make him a murderer, **by the way**. (1990 *CNN_King* [COCA])

Even though it is in post-clausal position in (2), *by the way* still functions as a Connector, this time orienting the AD/R back and signaling the relationship of the host D2 (*That doesn't make him a murderer*) to the D1 before that (*Do you suspect that…?*). Note that by using the post-clausal *by the way* in (2) SP pretends to downplay the importance of the preceding D2, and mitigates an accusation of poor reasoning on the Addressees's part. This can be seen by contrasting the serious assertiveness of *That doesn't make him a murderer*, without the DM.

A highly influential proposal is that there is a cross-linguistic systemic tendency toward asymmetry in the functions of pre-clausal and post-causal uses, see especially Beeching and Detges (2014a), Degand and Fagard (2011) on *alors* in French, and Haselow (2015) on *anyway*.[21] Using the terms "left periphery" (LP) and "right

21. It should be noted that the asymmetry exemplified here is a tendency only and not deterministic (see Traugott 2014 on *no doubt*).

periphery" (RP), it is proposed that, especially in conversation, there is a tendency for pre-clausal *E*s to host:

> elements that create discourse coherence, which includes functions such as the indication of topic shifts or disalignment from ideas expressed or implied in the preceding discourse. (Haselow 2015: 158)

In conversation the pre-clausal *E*s tend to have scope over longer discourses. In post-clausal position, however, *E*s tend to have scope mainly over the immediately preceding discourse, and to:

> host markers that renegotiate the effect and illocutionary force of an utterance, i.e. it is the preferred place for a modification of the modality of an utterance, but it may be exploited for retrospective textual linkage. (Haselow 2015: 158)

Renegotiating the illocutionary force of an utterance and retrospective textual linkage are exemplified by use of *by the way* in (2) above, where the epistemic certitude of D2, *That doesn't make him a murderer* is modulated by *by the way*. Examples of modulation of the modality of a D2 include the 'despite what was expected' use of post-clausal *after all* discussed in Section 4.5 below, and concessive-conditional use of post-clausal *anyway* (Haselow 2015).

And and *but* are Discourse Markers in some of their uses in both Schiffrin's and Fraser's senses, specifically when they are used to combine clauses. As discussed in Chapter 2.5, from the point of view of logic and formal semantics, both express coordination and are truth-conditionally, therefore semantically, equivalent (Kempson 1975). But in language use they do not mean exactly the same thing. Blakemore (2002: 99–103) explores this in some detail in connection with clause combining uses such as are illustrated by (3) and (4).

(3) Larry, Sue and Simon want coffee **and** Bob, Jane and Tom want wine.
(Blakemore 2002: 101)

Blakemore suggests that "*but* would only be acceptable here if there was something surprising either about the fact that Bob, Jane and Tom did not want the same drink as the others or about the fact that they wanted wine" (p. 101). On the other hand, *but* in (4) contradicts the assumption that Chicago is the windiest city in the US:

(4) New York was the windiest city in the United States today, **but** Chicago had light winds. (Blakemore 2002: 102)

Discourse Structuring Markers are Connectors that are "procedural". They guide interpretations in contexts (e.g. Blakemore 1987, 2002; Rouchota 1998; Hansen 2008: 20, 2012: 596) and point AD/Rs "to particular – more or less schematic – frames

of interpretation for the utterance hosting such expressions" (Hansen: 2012: 595). They are "highly compressed ways of expressing relations between ideas" (Lewis 2011: 440). In the following chapters, I highlight the multifunctionality of DMs and the importance of contexts in the development of DM meanings. In the view of DMs developed here, they are a multifunctional, non-contentful pragmatic subset of DSMs.

4.3 Discourse Structuring Markers

As has been mentioned from the beginning, the monofunctionality of some DSMs and the multifunctionality of others suggests that they should not all be treated as a single category if we are to account for language users' knowledge of DSMs. The categorization problem is somewhat diminished if categorization is based, as in the case of Fraser (e.g. 1996, 2006, 2009a), on the basis of prototypical use in pre-clausal position, but even in that position, whether a DSM is potentially multifunctional or not can be debated. For example, Fraser (2009a: 895) lists *anyway* as an attention-marker. However, Haselow (2015) convincingly shows that in pre-clausal position in conversation it has several functions. It is used primarily to introduce a new topic and to dismiss the earlier one as not highly relevant "for the speaker's communicative goals" (Haselow 2015: 168).

Building on Lenk (1998: 50) and Cuenca (2015), I distinguish Connectors that are:

a. lexical and fully contentful (*back to my point*),
b. partly contentful, partly pragmatic and monofunctional and partly contentful (*further*),
c. primarily pragmatic, multifunctional and primarily pragmatic (*after all, by the way*).

I call the first set "lexical Connectors", using Cuenca's term, and do not consider them to be DSMs. I discuss them in Chapter 9 mainly to highlight the distinctions among the three kinds of Connector.Cxns above. As discussed in the present book, DSMs are on a cline of pragmaticality from largely contentful, truth-conditional to largely pragmatic, non-truth-conditional, as in Figure 4.1. A preliminary rendering of Figure 4.1. was presented in Figure 1.1 in Chapter 1.2. It is slightly modified here to include the abbreviations used in this book: 1DSM (monofunctional DSM), DM (multifunctional DSM):

Figure 4.1 The semantics-pragmatics continuum of Discourse Structuring Markers

Minimally pragmatic 1DSM		**Maximally pragmatic DM**
largely contentful		largely pragmatic
monofunctional		multifunctional
further, instead, parenthetically	*incidentally*	*also, by the way*

Constructional properties of DSMs are discussed in Section 4.3.1 and types of Discourse Structuring Markers in 4.3.2. 4.3.3 provides a short interim summary.

4.3.1 Constructional properties of Discourse Structuring Markers

Syntactically, most Discourse Structuring Markers in English are Conjuncts. They:

1. serve linking textual functions,
2. have scope over clauses and more extended discourses,
3. can be used pre-clausally and are not fully integrated into clause structure.

I characterize expressions with these three properties as Connectors. Following the constructional view that expressions are [[F] ↔ [M]] pairings, I identity the syntactic properties in (ii) and (iii) above with Conjunct status (form) and the textual discourse-functional property of connectivity with DSM status (meaning). As defined in Quirk et al. (1985: 631–645), Conjuncts "have the function of conjoining independent units" and they provide assessment of how the SP/W "views the connection between linguistic units". Large numbers of examples of Conjuncts are cited in Quirk et al. including *after all, all the same, however, in addition, likewise*, which may occur in several positions in the clause, and are therefore different from conjunctions, which typically only precede the clause. In this book I do not make this distinction between Conjuncts that are mobile and coordinators that may only be used pre-clausally and those that are mobile. I regard all Connectors as expressions that may be used pre-causally to connect D2 with D1 and to provide an assessment of how SP/W views the connection between D1 and D2. Therefore *and* and *but* are included within an abstract Connector.Cxn. Constraints on position are typical of all Connectors and may change over time. They will be discussed in Chapter 12.

In construction grammar schemas are conceptualized as generalizations over particular micro-constructions. They have default properties that are passed down to those particular micro-constructions; exceptions have overrides. I represent the abstract schematic DSM pairing as [[Conjunct] ↔ [DSM]]. Default properties of DSMs are the syntactic property Conjunct and the connective discourse function. As will be discussed in Chapter 11, because it is the SP/W who conceptualizes the

type of connectivity between D1 and D2, there will by default also be at least a small degree of subjectivity in the pragmatic component. Also, because the SP/W cues the AD/R to the kind of connectivity intended, there is also by default be a small degree of intersubjectivity.

The kind of subjectivity involved in a DSM is in many cases a kind of stance-to-text (see Traugott 2020b), since DSMs usually signal an evaluation of the textual relationship between D1 and D2. This is particularly clear with DSMs like *all the same*, which is used to concede the content of D1 and by *the way*, which is used to imply that the content of D2 is unimportant. The suggestion that DSMs convey stance-to-text is an extension to Connective adverbials of the category of "stance" adverbials proposed in Biber and Finegan (1988, 1989), Biber et al. (1999), Conrad and Biber (2000). Stance adverbials signal the speaker's evaluation of the content of the clause (*regrettably* in *regrettably these do not match*), or of how it is said (*frankly, I'm bored*). Connective DSMs have a similar evaluative function, but the evaluation in this case concerns whether D1 is to be taken as an elaboration (*and, also*), contrast (*but, all the same*), etc. I call this "connectivity stance".

Figure 4.2 draws on Croft's (2001: 18) model of a construction (Figure 2.2. in Chapter 2.4), modified to show the default properties of DSMs in English discussed so far.

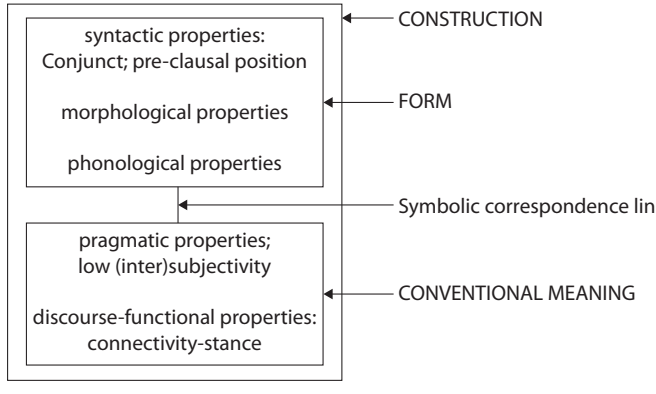

Figure 4.2 A preliminary model of the default symbolic structure of a DSM (based on Croft 2001: 18)

The model in Figure 4.2 will be further elaborated in Chapter 11. Use in particular non-initial positions and the particular kind of stance-to-text can be specified in the models of individual DSMs and the changes they undergo, as is illustrated in Figure 4.6 below in Section 4.5.2.

Given a model like that in Figure 4.2, we can think of specific DSMs as having:

a. individual phonological form paired with Conjunct syntax and with discourse-functional properties of a particular type of connectivity (for types, see Section 4.3.2 immediately below),
b. a relatively specific position on the pragmaticality continuum.

For example, *instead* has the default syntactic form Conjunct and the particular phonological form /InstEd/. On the meaning side it has the particular property of a relative degree of contentfulness. The default discourse-functional property is connectivity and the particular discourse-functional stance is Contrast.

DSMs are Connectors that signal a relatively symmetric connective relationship between D1 and D2 and are pragmatic to some degree. Connectors contrast with subordinators that signal a relatively asymmetric relationship between D1 and D2, and structural dependency of one on the other, e.g. [*although* D1, D2]. For DSMs an abstract constructional non-subordinate Connective.Schema is proposed with the form-meaning pairing [[D1 Connector.Cxn D2] ↔ [signals the relatively symmetric connective relationship between D1 and D2]]. Likewise, a Subordinate. Schema generalizes over subordinate, asymmetric connectivity between D1 and D2. Both the Connective.Schema and the Subordinate.Schema are daughters of a highly abstract super-schema that I call the Connectivity.Schema, as shown in Figure 4.3. To avoid the complexities of indicating clause order in the case of the Subordinate.Schema, which is irrelevant to discussion of DSMs, I specify only the meaning side of the latter:[22]

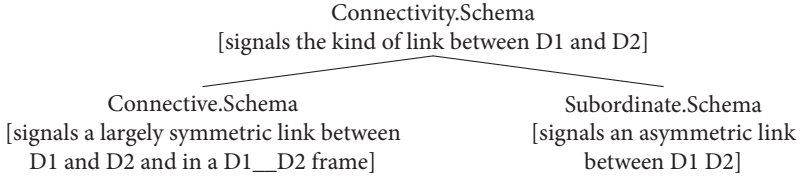

Figure 4.3 The connectivity.schema and its chief daughters

In this book I am concerned only with the Connective.Schema. Note that I use the term Connector to refer to the class of micro-constructions that connect D1 and D2. This category includes lexical Connectors as well as DSMs. "Connective" is used either as an adjective or, as here, in the expression "Connective.Schema" to refer to a hypothesized higher level schema.

22. To address ordering issues in the case of dative alternation, Perek (2012) proposed a "constructeme" that generalizes over order of nominal arguments. How best to formalize order of subordinate clauses remains to be considered.

4.3.2 Types of Discourse Structuring Markers

DSMs (including DMs) are Connectors licensed by the abstract Connective.Schema posited in the previous section. Fraser (1996 and elsewhere) identifies four types of what I call DSMs in English (he calls them DMs). While it is likely that such markers are available in most, perhaps all, languages (Heine et al. In press), the extent of the inventory of some types may differ from language to language, as may the extent to which the type is instantiated by DMs. For example, English has rather few reformulation markers such as *I mean, namely* within the category of Elaborative markers, but a significant number has been identified for Catalan and Spanish in Cuenca and Bach (2007). As will be discussed in Chapter 9, the function "Return to prior topic" is largely instantiated in English by *E*s that are lexical Connectors and not pragmatic.

Fraser's four types are:

1. **Elaborative markers**: used to implicate that D2 expands on D1: *and, also, further*. This set includes "reformulation markers" such as *in other words, namely, I mean*,
2. **Contrastive markers**: used to implicate that the content of D2 contrasts with that in D1: *but, instead, on the contrary*,
3. **Inferential markers**: used to implicate that D2 can be inferred from D1: *after all, for, now, so*,
4. **Topic change markers**: *as far as, by the way*, the function of which is primarily to signal the nature of the topic-orientation that SP/W intends. "Topic" is here understood as the discourse topic or what the discourse is about. DSM topic markers are above all markers of how coherence is construed and negotiated.

SP/Ws use such markers to express their (adopted) stance toward the relationship between what is said in D1 and D2. That the stance can be adopted and fake is particularly clear in the case of digressive topic change markers used to implicate that D2 is relatively unimportant when it may actually be the main point of the contribution (see Chapter 8).

It should be noted that In Fraser (2009a: 893) the fourth type, topic change markers, are called "topic-orientation markers" and accounted for in a separate subset of PMs, called "discourse management markers" along with a class of attention markers like *hey, oh, ok.* Fraser's reason for treating them as a separate category is that he sees them as "metatextual comment[s] on the structure of the discourse" (Fraser 2009a: 893).[23] The idea is that the first three types (Elaboratives,

23. In Fraser (2009b: 316) he defines DMs as semantic and says change in discourse topic is not a semantic notion, therefore topic-changing markers are not DMs.

Contrastives and Inferentials) mainly have scope over clauses, but discourse management markers mainly have scope over larger discourse. Fraser's distinction between DMs and discourse management markers will not be used here[24] because, as Schiffrin (1987) pointed out, many markers can be used at both "local" and "global" levels of organization. While she is concerned with organization of dyadic conversation, local and global scope are relevant in monologic written texts as well. From a cognitive point of view any procedural kind of cueing of the relationship between D1 and D2 contributes to cueing of coherence and how it is to be interpreted.

It should also be noted that among "discourse management markers" Fraser (2009a) identifies a class of topic-orientation markers that includes:

1. continuation with the current topic (e.g. *as I was saying*, *but*),
2. digression from the current topic (e.g. *by the way, I almost forgot*),
3. return to the previous topic (e.g. *back to my point*),
4. introduction of a new topic (e.g. *to change the topic, to return to the prior topic*).

As mentioned above, because the examples in the return to previous topic category are lexical and contentful I do not treat them as DSMs but as lexical Connectors. However, I discuss them briefly in Chapter 9 to spotlight the differences between them and DSMs and within DSMs between 1DSMs and DM.

4.3.3 Interim summary

To summarize this far, in this book I cut across some of Fraser's more recent categorizations and distinguish lexical Connectors that are fully contentful from DSMs that are partially or not contentful. The kinds of DSMs discussed are as follows:

1. Elaborators: Chapter 6
2. Contrastives: Chapter 7
3. Digressives: Chapter 8
4. Inferentials

Inferentials are illustrated below in Section 4.5 by the history of *after all.* Because introduction of a new topic has been exemplified in Haselow's (2015) extensive historical analyses of *anyway*, this kind of topic shift will not be discussed here.

24. The distinction is, however, used in Traugott (2020c).

4.4 Generalizations about the rise of Discourse Structuring Markers

While in English most Discourse Structuring Markers (and many PMs) arise historically via Conjunct adverbials from contentful Circumstance adverbials, a few arise via different phrasal configurations, e.g. *all the same*, which originates in a quantifier and N modifier.

Essential to the change from Circumstance Adverbial to [[Conjunct] ↔ [DSM]] is a shift in prototypicality associated with position. As will be shown in the case studies, there is evidence that this change to Conjunct status can occur only if the Circumstance adverbial (or other source expression) has been used repeatedly in clause-initial position, where it is a topicalized and serves as a frame for what follows.

In this section, I outline typical developments from Circumstance adverbial to [[Conjunct adverbial] ↔ [Discourse Structuring Marker]] (4.4.1) and from monofunctional 1DSM to multifunctional DM function (4.4.2).

4.4.1 From Circumstance adverbial to [[Conjunct adverbial] ↔ [Discourse Structuring Marker]]

Adverbials have long been considered to be adjuncts and not core elements in a clause (see recently Kim and Davies 2020, citing Kim and Sag 2005). From a syntactic point of view it is therefore possible to interpret the development of DSMs in English as a history of gradual shifting of some syntactically non-core expressions to ones that are even more marginal syntactically because they are syntactically not fully integrated. However, just as use of adverbials adds to the clause crucial material that frames and perspectivizes the content of the clause in terms of space, time, manner, etc., so use of DSMs adds crucial perspectivizing of the textual organization of what is being said or written. They are essential to the architecture of discourse.

The history of DSMs in English requires consideration of classes of adverbials. Adverbials have been analyzed in a number of different ways (e.g. Greenbaum 1969; Quirk et al. 1985; Ramat and Ricca 1994; Biber et al. 1999; Huddleston and Pullum 2002; Hasselgård 2010; Lenker 2010). It has become customary to distinguish:

a. "adverbs": a formal category of expressions that in PDE typically end in *-ly*, and are syntactically dispensable (adjuncts),
b. "adverbials": adverbs having the functional property of modifying verbs (*sang loudly*), adverbs (*very loudly*), adjectives (*extremely loud*), and clauses (*frankly, he did a poor job*).

Drawing on Hasselgård (2010), who in turn draws on Greenbaum (1969) and Quirk et al. (1985), among others, I classify adverbials according to various types of functions. Types particularly relevant to the present study are Circumstance adverbials and Conjunct (linking) adverbials.

Circumstance adverbials are adjuncts with semantics of space (*at home, along the road*), time (*at that time, after all the speeches*), manner (*loudly*), instrument (*with a hammer*), etc. I abbreviate them as CircAdvs. Prototypically CircAdvs:

i. are syntactically focusable and "addressable" (they can be *it*-clefted, negated and questioned, see Mittwoch et al. 2002),[25]
ii. are semantically relatively contentful and monofunctional,
iii. have scope over a phrase,
iv. are integrated into clause structure.

Prototypically, CircAdvs have been used in clause-final position from Middle English on, but they can also be used in other positions. Topicalized use in initial position is atypical in Present Day English but they frame the content that follows, especially in rhetorically parallel argumentation:

(5) **In Greece**, farmers expect disaster should there be no rain next year. **In France**, the drought has been the worst on record.
(*The environment digest* [BNC]; Fagard and Sarda 2014: 199, bolding original)

Fagard and Sarda (2014: 199) point out that "framing adverbials delimit discourse segments that are cohesive with the interpretative criteria they set up". By hypothesis the use of a CircAdv in initial position is key to its reinterpretation as a Conjunct.

As stated above in Section 4.3, Conjuncts are linking adverbials. They:

i. serve connective textual functions (*after all, by the way, in addition, instead*),
ii. have scope over clauses and more extended discourses,
iii. tend to be used pre-clausally and are not fully integrated into clause structure.

They do not have the first characteristic of CircAdvs mentioned above as they cannot be clefted and they are not addressable.

The beginning of the clause is a position in which discourse framing occurs (Schiffrin 1987; Auer 1996; Aijmer 2002). As mentioned in Chapter 1.5, absent prosody, it is often not possible to distinguish pre-clausal (DSM) and initial (topicalized CircAdv) position syntactically (Gregory and Michaelis 2001). Therefore use at the beginning of the clause is a locus of change from CircAdv to DSM status.

25. For discussion of focusability and addressability see Boye and Harder's (2007, 2012) hypothesis about grammaticalization in Chapter 5.3.

As will be discussed later, some adverbials with connective meaning can be used at the end of the clause as well, e.g. *after all*, see Section 4.5.1 below, and *all the same*, Chapter 7.3. Again, it is not always possible, absent prosodic information, to distinguish clause-final adverbial and post-clausal DSM use. ("Comma punctuation" is not a useful criterion since commas are not used systematically, and present day punctuation practices have been conventionalized only fairly recently, see Parkes 1992).

Table 4.1 models prototypical and atypical alignments of CircAdvs and Conjuncts with respect to position relative to the clause in Early Modern English, the period in which many of the changes discussed occurred. In this Table, distinctions between pre-clausal/initial and post-clausal/final position are ignored as these cannot be assessed well in written historical data, where we have no prosodic information.

Table 4.1 Prototypical positions for CircAdvs and Conjuncts in Early Modern English

	Typical	**Atypical**
CircAdv	Final	Initial
Conjunct	Initial	Final

The hypothesis is that in the communicative act interlocutors repeatedly reinterpreted certain topicalized CircAdvs as univerbated rather than phrasal chunks. Given replication in topicalized initial position where it was used with framing function, a contentful, monofunctional [+truth-conditional] CircAdv could be reinterpreted as a semi-contentful, largely monofunctional and partially [–truth-conditional] Conjunct. The result was a profile shift, for example, a topicalized *after all* 'after everything' (a temporal adverbial phrase with an alternant *after all this*) was reinterpreted as a univerbated form associated with a default connective meaning 'in the end' (an inferential relationship, used only in the pronominal form *after all*). This was a constructionalization because there was a shift in default position and morphophonology (form) and in function (meaning). A new construction was added to the constructicon. The profile shift and subsequent reconfiguration of one of the uses of CircAdvs exemplifies what Van de Velde (2014: 173) characterizes as "form-function changes [that] involve strengthening of already available resources with extension to new domains". He says this happens "when a new subsystem comes under pressure". However, there is no obvious pressure on a subsystem in the development of Conjunct uses of CircAdvs.

Because DSMs are used by default by SP/W to signal to AD/R how the relationship between D1 and D2 should be understood, all DSMs as they develop

undergo some degree of (inter)subjectification, the change whereby later meanings come to be more grounded in the SP/W's and AD/R's perspective than earlier uses. That is, they come to be used more (inter)subjectively. Association with at least a small degree of (inter)subjectivity is one of the main differences between DSMs and CircAdvs.

4.4.2 From monofunctional to multifunctional Discourse Structuring Marker function

In some cases, as more pragmatic nuances come to be associated with a particular Conjunct, it may come to be used in multifunctional ways not associated with the original [[Conjunct] ↔ [DSM]] pairing. This is a reshaping of the pragmatic features of the Conjunct resulting in a constructional change. All DMs are conventionally pragmatic. They are all multifunctional, but not all are mobile (e.g. *and* and, until recently, in most varieties of English, *but*), therefore availability with different pragmatics in different positions is not an essential component of reorganization of a Conjunct with largely monofunctional meaning as a multifunctional DM.

As will be seen in each of the cases studied in Part II, some DSMs have not come to be used as DMs in the sense discussed here. In other words, some DSMs are largely monofunctional and partially contentful, e.g. *in addition, instead* (1DSMs). The evidence from the corpora suggests that just as CircAdvs are not used as Conjuncts without being used in initial position first, so Conjuncts are not used with DM function without prior monofunctional DSM use. This means that there is always the potential for a Conjunct to be used with multifunction DM function, but this potential is often not exploited.

Figure 4.4 (Figure 1.3 in Chapter 1.4 repeated here for convenience, with a slight modification) summarizes the main kinds of changes just discussed as the Discourse Structuring Marker Trajectory Hypothesis (to be modified slightly in Chapter 7.5 in connection with Connectors without adverbial sources). It assumes that [[Conjunct] ↔ [Discourse Structuring Marker]] is a constructionalized [[F] ↔ [M]] pairing and that DMs are an optional discourse functional subset of DSMs. While all Conjuncts are largely monofunctional DSMs (1DSMs) with partially contentful/partially pragmatic meaning, only some come at a later stage to be used as largely pragmatic multifunctional DMs. Unlike in Figure 4.4, Conjunct is not repeated in the third representation of a construction as there has been no change.

For any particular shift, there will be assignment to a particular discourse-function such as the ones mentioned in 4.3.3 above: Elaborative, Contrastive, Inferential or Digressive.

DSMs tend to have meanings related to those of their CircAdv sources, and DMs tend to have meanings related to those of their Conjunct sources. This is

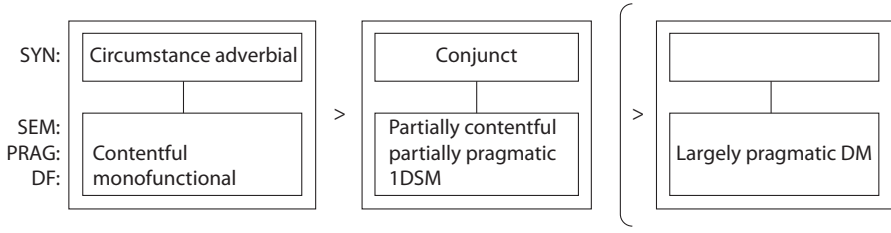

Figure 4.4 The Discourse Structuring Marker Trajectory Hypothesis

a kind of polysemy, but as Percillier (2020: 220) cautions, traditional "polysemy", understood as multiple meanings associated with a form, is not adequate for constructional analysis. Rather, we need to think of "a single construction with related constructional meanings". That said, it is useful to recall that Hansen (2012: 598) observed that "[t]ypically, synchronic polysemy [read: "multiple meanings associated with a construction"] will be the result of diachronic sense extensions". The layering can persist for quite some time. Indeed, Indo-European uses of CircAdv temporal and Conjunct relatively monofunctional DSM (later multifunctional DM) uses of NU 'now' have persisted since before Old English times (Auer and Maschler 2016). However, in many cases layering is observable only for a time. There is, for example, little remaining relationship, either formal or meaningful, between Old English *eall swa* 'all equally' and *also*.

4.4.3 Contexts for the rise of Discourse Structuring Markers

Modifying Bybee et al.'s well known statement (1994: 297) cited in Chapter 3.6 about grammaticalization: "Everything that happens to the meaning of a gram happens because of the context in which it is used", we can say "Everything that happens to the meaning of a construction happens because of the context in which the construction is used". Contexts that are especially important for the development of DSMs are the surrounding text ("co-text") and the assembly over time of multiple different contexts (Petré 2019).

An example of co-textual contexts for the development of Conjuncts out of lexical CircAdvs is use of an *E* in the immediate context of descriptions of locutionary events, as noted in Lenker (2010):

> many of the new adverbial connectors which have developed in the history of English started their life as circumstance adverbials on the phrasal level, modifying a verbal phrase containing a *verbum dicendi* ('verb of speaking' ECT).
>
> (Lenker 2010: 38; italics original)

Lenker (2010) cites such contexts for, among other adverbials, *now* 'at this point in the text' in Old English (6a) and *also* in Early Modern English (6b):

(6) a. We willað **furðor** ymbe ðas emnihte swiðor sprecan on
We will further about those equal-nights more-correctly speak in
gedafenlicere stowe; & we **secgað nu** sceortlice þæt…
more-suitable place; and we say now briefly that…
'We will speak further and more correctly about the equinox in a more suitable place; and we say now briefly that…
(c1000 Ælfric, *De Temporibus Anni* [HC: CoTempo; Lenker 2010: 38])

b. **Also** it is **said** that (Theriaca Athanasia) doe both resolue, breake and digest humours.
'Also it is said that Theriaca Athanasia [a drug thought to be anti-poison] resolves, breaks down and digests [diseased] body fluids.'
(1602 Clowes, *Treatise for the artificial cure of struma* [HC CeScie2a; Lenker 2010: 39])

In later chapters we will see the importance in some cases, especially *by the way*, of discursive assembly with a verb like *say, talk*. (7) is an example:

(7) [beasts] of the which some are of the greatnesse and height of a great pigge, the others lesse, and this much will I **say by the way**, their flesh is tender and pleasant for to eate.
'some of which are the size and height of a large pig, the others smaller, and this much I will say in passing, their flesh is tender and pleasant to eat.'
(1568 Thevet, *The new found vvorlde* [EEBO])

Contexts also include the syntactic positions in which a DSM can occur. The positions most often discussed are what I am calling "pre-clausal" and "post-clausal". However, clause-medial position is also used in some cases. As will be seen immediately below in connection with *after all*, DMs can be used with very different meanings in different positions (and see further Chapter 12).

4.5 A preliminary case study: The development of *after all*

To exemplify the rather abstract issues raised above and in Chapter 3, here I provide a sketch of the history of *after all*. A more detailed account of this history in constructional terms can be found in Traugott (2018a). In this section I first outline uses in PDE (4.5.1) and then the history (4.5.2).

4.5.1 *After all* in contemporary American English

After all is listed in Fraser (1996: 340) as an inferential DSM signaling that D2 is a conclusion that follows from D1. In many cases it marks SP/W's justification for saying D1, while at the same time implicating that AD/R is often expected to share the assumptions behind D2, as in:

(8) So any ideas or pointers would be helpful. **After all** computers are supposed to make life easier, not harder!
(2012 *Multiple live broadcast at the same time*? [COCA BLOG])

This is how it is normally understood in pre-clausal position, where it appears most frequently. But *after all* is pragmatically polysemous, depending in part on position, as was illustrated by Example (4) in Chapter 1.5.

In medial position, especially after *BE*, it can be understood as meaning 'as is well known, of course'. In (9) it has a stronger effect than pre-clausally, and functions intersubjectively to signal 'I expect you to remember/use your encyclopedic knowledge'. In (9) the speaker (Gingrich) justifies why he said *I think too many people are dismissing Bachmann*, and at the same time reminds AD/R that Bachmann's being a woman is 'of course' important in a political campaign. In other words, he invokes knowledge that both Bachmann and McCain (not Cain!) were Republican candidates for the Presidency of the US:

(9) I think too many people are dismissing Bachmann. She is **after all** a woman and since Cain has been slandered so much by the democratic party it only makes sense she would be nominated.
(2012 *Gingrich tells ABC News* [COCA WEB])

In post-clausal position *after all* is often used with a concessive meaning ('despite expectation'), as in (10), where *after all* evokes 'despite the claim that wolves are easy to live with'.

(10) Once they [wolves] even tried for a horse in my hometown but that was a vey [sic] peculiar case. So they are not too easy to live with **after all**.
(2012 *Would real wolves act like the wolves of 'The Grey'?* [COCA BLOG])

It is this concessive use that is most often cited in dictionaries. For example, the online Google dictionary defines *after all* as 'in spite of any indications or expectations to the contrary'. The on-line Merriam Webster dictionary gives two definitions: the first is 'in spite of considerations or expectations to the contrary', with a post-clausal example, and the second is 'in view of all circumstances', with a clause-medial example. No pre-clausal example is provided, nor is a justification

interpretation.[26] Constraints on meanings of *after all* in particular positions do not predict meaning in PDE and the distinctions between (8), (9) and (10) are not always readily discernable. They are the residue of historical developments.

4.5.2 A sketch of the history of *after all*

DSM *after all* originates in a Circumstance temporal construction meaning 'after everything, at the end'. The first likely example of *after all* (pronominal) in EEBO is (11). Here it can be interpreted as 'after all this' and is a clause-medial adverbial with an adjective *all* used pronominally (*anone* 'immediately' is probably a topicalized temporal adverbial that is an argument of the clause):

(11) they were dampned to the deth as fals traitours and anone **after all** the britons of the lande by commune assent crouned vortiger.
'they were condemned to death as false traitors and soon after everything the Britons of the land by common assent crowned Vortiger.'
(1480 Caxton, *In the yere of thyncarnacion of our lord* [EEBO])

As a reviewer pointed out, (11) is ambiguous as cited. It could be parsed as 'soon after, Vortiger was crowned by all the Britons'. However, this interpretation is unlikely if the fuller context is considered. The passage follows a sequence of events: *and anone let take the hondred knyghtes of pehites and bynde hir hondes behynde hem and lede hem to london* 'and right away had the hundred Pictish knights seized and their hands bound behind them and led them to London'. Note that this first set of episodes is introduced by *and anone*. As I interpret it, the narrative here is structured as two main episodes, each introduced by *and anone*, and *after all* refers back to the first set of episodes.

In the context of reasoning, *after all* later came to be understood to mean 'in the end, eventually, on reflection', sometimes 'in conclusion', a conclusive use that Lewis (2000, 2007) regards as a crucial step toward its later use as a DSM. Lewis (2007: 96) finds that in her data contexts are "often polar or scalar, insofar as the idea in the scope of *after all* is typically something smaller than, or less than, expected, or is the contrary of what was expected". An 'in the end' reading is exemplified by (12) because a set of alternatives precedes the final choice.

26. "After all." *Merriam-Webster.com Dictionary*, Merriam-Webster, https://www.merriam-webster.com/dictionary/after%20all. Accessed 14 Aug. 2021.

(12) man hangs in a ballance like a young virgin vvhich hath many suters:
man hangs in a balance like a young virgin who has many suitors:
some she fauoures for parentage, some for personage, some for friends,
some she favors for parentage, some for personality, some for friends,
some for vvealth, some for vvit, some for vertue: and **after all**, chuseth
some for wealth, some for wit, some for virtue: and in the end, chooses
the vvorst of all:
the worst of all: (1588 Smith, *The Christians sacrifice* [EEBO]; Traugott In press: 72)

An explicit mention of coming to a conclusion after deliberation appears in a 1620 translation from Italian of a famous text, Bocaccio's *Decameron* (note the locutionary context):

(13) he resorted to his other brethren, and told them what he had seene in the time past, betweene their sister and lorenzo: many deliberations passed on in this case; but **after all**, thus they concluded together, to let it proceede on.
(1620 Florio, *The decameron* [EEBO])

In (12) and (13) the 'in the end' use of *after all* is still largely lexical and semantically contentful, but is used with inferential pragmatics in clause-initial contexts where coming to a conclusion at the end of a series of discussions is appropriate. I concur with Lewis (2000, 2007) that the 'in the end' reading was a crucial step in the development of DSM meanings, as these involve reasoning. The development of this reading may have been accompanied by some loss of morphological compositionality and univerbation. By hypothesis development of 'in the end' exemplifies constructionalization of *after all* as a Conjunct connecting two segments of discourse. It is largely monofunctional, but in examples like (12) a concessive can (with hindsight) also be inferred because the 'worst of all' is a polar choice that is contrary to what might be expected.

During the 1630s and 1640s *after all* appears in a number of examples in modal contexts (e.g. after conditional *if*) or in final position. Here, even though it can be understood as meaning 'in the end', and may have been intended with that meaning, it may also have been understood as signaling a stronger inferential relationship, depending on context. For example, clause-medial use in (14) after modal *he might have stayed with us* can be inferred to implicate justification for the position put forward in D1 ('why did we need Christ's ascension?'):

(14) what needed then (say wee) the ascension? hee might have now stayed with us, **after all**, for our comfort.
'why did we need (we say) the ascension? he might have stayed with us, after all, for our comfort.' (1635 Austin, *Devotionis Augustinianae flamma* [EEBO])

And clause-final use in the conditional clause in (15) can be inferred to have been understood as inviting the AD/R to re-process what precedes ('save them') with the concessive implicature 'though X might have thought otherwise/despite everything':

(15) if all the gods of the world can do for you what he hath done, can pardon sinne by giving his son, can heale your soules, and save them **after all**, follow them.
(1646 Peters, *Gods doings* [EEBO])

This 'despite everything' meaning can also be inferred after the subordinator *if*:

(16) to what purpose were all their watchings, … their so frequent martyrdomes, of what excellency or availe, if **after all**, they should be hurried out of this world and all their fortunes and possessions, by untimely, by disgracefull, by dolourous deaths, (1647 Taylor, *Of the sacred order* [EEBO])

Examples like (14)–(16), which probably exemplify the inferential 'in the end' reading enriched by justification or concession interpretations, are found in rhetorical texts that are intended to explicate hidden meanings in Biblical texts. They are not found in practical instructions, even if D2 contains a modal. For example, (17) occurs in a set of instructions for planting trees. These instructions are of the form *you may V*. In (17) *after all* is unambiguously used as a CircAdv meaning 'after all the other actions':

(17) you may lay some dung upon the uppermost face of the hole **after all**, if you please ('if it please you'), in light and sandy ground.
(1650 Hartlib, *A designe for plenty* [EEBO])

In the early 17thC we find several instances of *for after all these things*, *yet after all this* where the adverbial is in clause-initial position. It is a topicalized temporal adverbial argument of the clause. Toward the end of the century pronominal *after all* is attested in the context of *for* and *yet*. Both *for* and *yet* were already extant DSMs used to signal inferential relationships between D1 and D2. Although pronominal *after all* can be understood as Conjunct 'in the end', the preceding DSMs probably strengthened the inferential meanings: justificational with *for*, concessive with *yet* 'however':

(18) a. but o the mischief of common-place-books![27] which make men write what they find, and not what is to their purpose: **for after all**, dr: field doth but seem to advance another principle in his opinion.
(1673 Stillingfield, *Answer to several late treatises* [EEBO])

27. A 'common-place-book' was a book with lists of quotations from famous people.

b. although a person useth his sincere endeavour by all moral helps, and the divine grace assisting him to find out in these writings the things necessary to salvation, **yet after all** he can not certainly understand the meaning of them (1673 Stillingfield, *Answer to several late treatises* [EEBO]; Traugott In Press: 73)

In sum, during the 17thC a cluster of replicated contextual uses is attested in EEBO that enriched the inferential meaning of *after all* 'in the end'. They were:

a. use in contexts of reasoning,
b. association with a modal clause (in religious texts),
c. collocation with an inferential DSM.

By the end of the century SP/Ws began to use *after all* independently of such contexts. This use was a profile shift such that the pragmatic interpretations available in context came to be understood as the main contribution of *after all* to a [D1 __ D2] sequence, as in (19):

(19) he understood the humour of the athenian people, who were easily disgusted with the merit of extraordinary persons, whom he set his wit to abuse, that he might please that people: **after all**, he often is no ('not') otherwise pleasant than by his Buffoonry.
(1674 Rapin, *Reflections on Aristotle's treatise of poesie* [EEBO])

Examples such as (19), in which the context is neither *for* nor modal, suggest that a justificational use was conventionalized by the 1670s and that *after all* had come to be linked by at least some SP/Ws in the network of constructions to the justification schema of which *for* was already a member.

Around this period *after all* also begins to appear in medial position and contrastive but not modal contexts as in (20), where *yet* signals the contrast. Here it does not function as temporal 'after everything', inferential 'in the end' or justification for saying D2, but can be understood as signaling that a certain situation holds despite expectations arising from prior discourse, specifically despite having indicated their discontent regarding the secretary's pay, they were willing to put that discontent aside under certain circumstances:

(20) [regarding perceived favoritism and corruption] they being of opinion, that more moderate fees ought to have contented a pen-man that was no better qualified than with the ordinary fruits of a writing-school: yet they assured his majesty **after all**, that if a reasonable supply would suit with his occasions; they were ready so far to testifie their obedience.
(1681 Cotton, *The field of bloud* [EEBO])

This kind of concessive meaning came to be entrenched in post-clausal use as SP/W's subjective assessment without contrastive *yet* or *but* during the 18thC, but still largely in conditional contexts, as in:

(21) a. [after lengthy expressions of suspicion about her daughter] Why, if she should be innocent, if she should be wronged **after all?** I don't know what to think. (1700 Congreve, *Way of the world* V [Traugott 2004: 558; 2018a: 36; In press: 74, 176)

b. [Lettice's soliloquy after a dream that her dead mother is alive] But let me think a little. If my Mother shou'd be Alive, **after all**. Ay, that wou'd fright me worse than seeing twenty Ghosts, for she'll force me to marry Ned Ploughshare.
(1730–31 Lillo, *Sylvia* [CLMET_3_0_1_14; Traugott 2018a: 36])

The new pragmatic uses around 1700 suggest that *after all* was regularly being used as a DM by the beginning of the 18thC.

Working with a much smaller corpus, in Traugott (2004) I suggested that concessive uses arose before justificational ones. The far larger EEBO corpus suggests that both may have arisen in the late 17thC, but in different contexts, the justificational meaning in the context of *for* and pre-clausal use, the concessive meaning in *yet*, modal, conditional and post-clausal contexts.

Lewis (2007: 97) interprets the development as a functional split, modeled in Figure 4.5. In this model she counts only justification as a "connective", because it is used in initial position in her data. I, however, count the concessive use as connective also, since it signals a connection between a discourse segment and an earlier one that is textually either available or accessible.

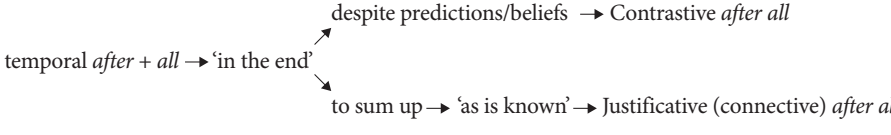

Figure 4.5 Functional split of *after all* (Lewis 2007: 129)

The use of *after all* as a concessive is consistent with pan-European use of *all* in concessive *Es* such as English (*although, all the same, for all*), German (*allerdings, bei all*), French (*toutfois*, *tout … que*) (König 1985: 266). It may be that there was French influence on the development of *after all* (cf. Fr. *aprés tout*).[28] In its concessive

28. For an example of a detailed study of the likely influence of French on a DM in the Late Middle English and Early Modern English periods, see Sorva (2007) on the rise and obsolescence of conjunctive concessive *albeit* 'although'.

use *after all* was linked with the contrastive schema that includes *however* and *though*, both of which are monofunctional paraphrases. Lenker (2010) shows that post-clausal position came to be associated with concessive use in the 18thC. However, over time, concessive uses of *after all* weakened and many clause-final uses of are best understood as justificational, especially in the context of a predication with *BE*:

(22) in solitude I often shuddered at my friend … On meeting him again, I was often filled with remorse, when his deep eyes beamed kindly upon me, as with the glow of a household fire that was burning in a cave. "He is a man **after all**," thought I; "his Maker's own truest image, a philanthropic man! –"
(1852 Hawthorne, *Blithedale romance* [COHA])

Although a concessive reading seems possible in (22) given that the cognizer is describing a change in her perspective, it is more plausibly used as an explanation of her remorse at shuddering when she saw her friend. The two readings can also be merged.

A further, somewhat infrequent, use is an epistemic 'of course, everyone knows' reading found in medial position with a form of the verb *BE*. This use is an extension of justificational meaning and developed in the 19thC. It usually occurs in the context of a statement that S/W wants the AD/R to take as a general truth or well-known fact:

(23) they have not convinced me of the incorrectness of my opinion; because that which is founded in truth is, **after all**, the only thing that is 'good and nourishing' to the understanding. (1820 Ballou, *Series of letters* [COHA])

In PDE this has come to be the typical use of *after all* in medial position (see Example (9) above in Section 4.5.1).

Around 1900, use of *after all* is prototypically justificational in pre-clausal position, concessive in post-clausal position, and epistemic in medial position if the predicate is *BE*. This division of labor has, however, been somewhat relaxed, especially in colloquial usage, and the justificational reading has come to predominate in all positions in speech, blogs and represented speech, according to the 2020 COCA data, and a small count of the COHA data for the decade 2000 presented in Table 12.4 in Chapter 12.3.1. Lewis (2007: 98) states that concessive use "is no longer possible in PDE", and this is borne out by the COCA data. It is, however, available intuitively.

Position does not modulate the meaning of most DMs as dramatically as it has in the past for *after all*, as will become evident in later chapters. Nevertheless, as we will see in Chapters 7 and 12, a weak correlation of post-clausal position and concessive meaning appears to be characteristic of English.

Finally, a word about subjectivity, intersubjectivity and increases in them with respect to *after all*. In PDE all DM meanings of *after all* are subjective in that they are used to express SP/W's perspective. All index SP/W's reasons for saying D2. To the extent that justification and concession anticipate AD/Rs objections and counterarguments, they are used intersubjectively, but in several examples like (22) justification and concession engage SP/W's own arguments rather than those of the interlocutor and any implied intersubjectivity appears to be mainly a default of procedural status and the speech situation rather than a new conventional meaning. However, the 'of course' reading used mainly in medial position as in (23) signals that AD/R should recall some fact and this use does appear to have arisen by inter-subjectification, the process whereby SP/Ws come to use an expression to convey more attention to AD/R than it did at an earlier time. These kinds of issues will be further discussed in Chapter 11.

The changes to use of *after all* up to about 1900 can be summarized in Figure 4.6 using Croft's (2001) constructional model cited in Figure 2.2. in Chapter 2.4. Only the barest outline of the most relevant factors is provided in Figure 4.6. Contexts are distinguished graphically by italics between stages of development. If a feature such as "Conjunct" does not undergo change, it is not repeated; "alt", "(pre)", "(post)", "med", "infer", and "justif" are short for 'alternative', 'pre-clausal', 'post-clausal', 'medial', 'justificational', and 'inferential' respectively.[29]

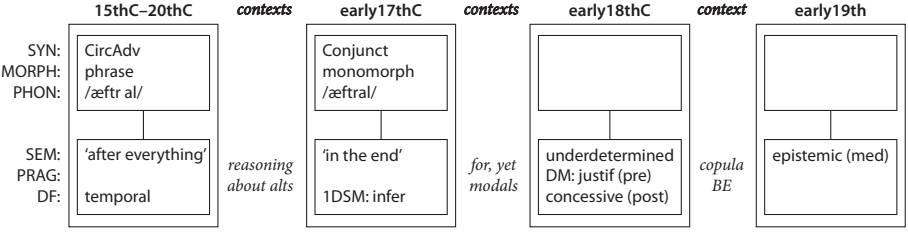

Figure 4.6 Outline of the rise of DM uses of *after all* up to c1900

Looking at the features of each grid in Figure 4.6, one can see that the shift from CircAdv to Conjunct was a constructionalization because in the early 17thC both the syntax and the semantics of the construction are different. There are no further form changes. However, there are several changes in the meaning component of the construction, and these are constructional changes. Because a DM is a kind of DSM, no new microconstruction was added to the constructicon when *after all* was neoanalyzed as a DM. The dating in the top line states the approximate

29. An earlier version of Figure 4.6 appears in Traugott (In Press: 78).

time when the use in question is attested with some frequency and across texts by different authors, suggesting that it was conventionalized by that time; in the case of the first grid, the notation 15thC–20thC indicates that CircAdv uses of *after all* became almost obsolete in the 20thC.

A brief history of one Inferential DM, *after all*, was presented above. Treatment of other Inferentials in constructional terms like *for, so, therefore* awaits further study, but see Lenker (2010) for analysis of Inferentials with sources in adverbials that signal result, e.g. *therefore*.

4.6 Summary

The main points of this chapter are:

a. DSMs are syntactically Conjuncts,
b. DMs are a highly pragmatic, multifunctional subset of DSMs,
c. Historically, the development of Conjuncts out of CircAdvs is a case of constructionalization as both form and meaning change, and a new construction is added to the constructicon,
d. The development of DM uses is a case of reshaping the pragmatics and discourse function of an extant [[Conjunct] ↔ [1DSM]] and is a constructional change.

In the next and final chapter of Part I we will look briefly at some alternative interpretations of the rise of DMs, specifically the hypotheses that DMs arise by processes of grammaticalization, pragmaticalization, cooptation into the thetical grammar component of Discourse Grammar, or diachronic construction grammar.

CHAPTER 5

Alternative hypotheses about the rise of Discourse Markers

5.1 Introduction

Before moving on to the case studies and evidence for the need to introduce more pragmatics into construction grammar, it may be useful to position the approach in this book vis-à-vis some non-constructional functionalist proposals regarding the type of morphosyntactic change that is involved in the rise of DSMs. It should be noted that the proposals refer to DMs in the broad sense of Schiffrin (1987) and I will retain that terminology here when quoting the proposals. (By way of reminder, Shiffrin's DMs are what I call PMs, a largish set of markers with diverse functions: social, epistemic, and connective. DSMs are connective only, but as 6 of Schiffrin's 11 DMs are connectors: *and, but, now, or, so, then*, the proposals under discussion are intended to cover DSMs as well as other types of PM).

Three proposals have received particular attention: the rise of DMs (in Schiffrin's sense) is a case of an extended view of grammaticalization (e.g. Traugott 1997[1995]), Brinton 2008); it is a case of pragmaticalization (e.g. Erman and Kotsinas 1993), and, more recently, it is a case of both grammaticalization and cooptation (e.g. Heine et al. 2017; Heine et al. In press). In this chapter I will outline and evaluate these proposals briefly in the light of the constructional perspective developed for DSMs here. A fuller discussion would identify many more subtle points of agreement and disagreement.

The grammaticalization proposal is discussed in Section 5.2 and the pragmaticalization proposal in Section 5.3. The third proposal, that DMs are coopted into the "thetical" component of Discourse Grammar, sometimes after, and sometimes before grammaticalization, is discussed in Section 5.4. I argue in Section 5.5 that a diachronic constructional perspective offers the most adequate account of the rise of DSMs (and by extension, the larger set of PMs). Section 5.6 summarizes, with focus on comparing the questions asked in each research paradigm and what the main issues are for the rise of DSMs.

5.2 The grammaticalization hypothesis

The hypothesis that the rise of DSMs is a case of grammaticalization presupposes that DSMs are part of grammar. Narrowly conceived, work on "grammar" has until relatively recently been focused on syntax, morphology and phonology and has excluded most DSMs. However, DSMs are encompassed by grammar provided it is conceptualized broadly to include discourse-pragmatic elements (see Waltereit 2012; Degand and Evers-Vermeul 2015: 61). Fraser (1988: 32) explicitly says DMs (in his sense of Connectors) are part of grammar.

As outlined in Chapter 3.5, the processes involved in the kind of change known as grammaticalization have been considered to include various kinds of reduction, including what are widely known as:

a. "bleaching" (loss of "semantic integrity"),
b. loss of syntactic freedom; this includes fixing not only of word order, but also of units in phrases ("univerbation" in the grammaticalization literature, "chunking" in the construction grammar literature),
c. syntactic scope reduction.

It has been noted that "bleaching" or loss of content meaning is usually accompanied in the course of grammaticalization by gain of procedural meaning (Sweetser 1988) and that therefore a "loss-and-gain" model is needed (Heine et al. 1991; Brems 2011). While monofunctional 1DSM sources of DMs are only partially bleached, *Es* come to be used in less and less contentful ways as they are recruited to multifunctional pragmatic DM status. Acquisition of new procedural functions is a gain that is typical of grammaticalization. DSMs are univerbated if the source is phrasal. However, as is discussed below, the development of DSMs counterexemplifies (c) scope reduction and some aspects of (b) loss of syntactic freedom.

The fact that most DMs (and PMs in general) originate in lexical items and over time are used as procedural items cueing SP/W's construal of relationships among segments of discourse led me to propose in the mid 1990s that DMs arise by the process of grammaticalization (Traugott 1997[1995]). I took Meillet's characterization of grammaticalization as "the attribution of grammatical character to an erstwhile autonomous word" (see Chapter 3.5) as my starting point and defined grammaticalization as: "The process whereby lexical material in highly constrained pragmatic and morphosyntactic contexts becomes grammatical". This is a unidirectional statement and in that respect congruent with then-current characterizations of grammaticalization.

The strictest interpretation of grammaticalization at the time was Lehmann's (2015[1995]) book *Thoughts on Grammaticalization*, an updated version of his 1982 working paper in the University of Cologne's UNITYP project on typology. Lehmann

proposes 6 parameters of grammaticalization (Lehmann 2015[1995]: 174) ("grammaticality" would be a better name for them as the parameters are synchronic) and 6 grammaticalization processes that affect these parameters. Of these processes, two are clearly relevant in an account of the development of [[Conjunct] ↔ [DSM]] pairs, notably:

1. Attrition, the process related to the parameter "loss of integrity" (Lehmann 2015[1995]: 134–141). This encompasses loss of semantic substance or desemanticization ("bleaching"), which can be understood as loss of contentful semantic features. All [[Conjunct] ↔ [DSM]] pairs can be said to have lost semantic substance; those that are multifunctional (DMs) more so than those that are monofunctional (1DSMs).
2. Bondedness (Lehmann 2015[1995]: 157–167). This concerns coalescence or chunking. Conjuncts typically become units morphologically, even though they derive from phrases (*after all, by the way, all the same*). Some combinations of Connnectors may themselves become units (*now then*, and partially, *Oh, by the way*), as discussed in Chapter 10.

   Another two of the processes that Lehmann mentions are partially relevant:
3. Paradigmatization (Lehmann 2015[1995]: 141–144). Items that are loosely related in a semantic field come to be organized in a small, tightly integrated paradigm. Diewald and Smirnova (2012) have identified paradigmatization as a fourth essential stage in grammaticalization, the first three being "untypical" contexts (associated with conversational implicatures), "critical" contexts (associated with multiple opacity), and reorganization and differentiation (associated with polysemous or heterosemous[30] items) (Diewald and Smirnova 2012: 128). The history of DSMs is a history of paradigmatization within a schema: sources of the DSMs discussed here are for the most part adverbial, and mostly spatial, temporal, or manner. Only a smallish subset of adverbials underwent neoanalysis as Conjuncts and only a small subset of those came to be used with DM function. The Conjuncts came to be integrated into the [D1 Connector.Cxn D2] schema that licenses a relatively small set of Connector alternatives. This set consists of a small number of subtypes (Elaborators, Contrastives, Digressives, Inferentials, etc.). New DSMs come to be integrated into these subschemas (or are lost from them), as will be discussed in Chapter 6 in connection with Elaboratives. But these subschemas are not tightly integrated in the way paradigms in the strict sense of grammaticalization work are.

30. "Heterosemy" is a subtype of polysemy in which related *Es* instantiate different categories, e.g. *have* (main verb and auxiliary) and *bank* (noun and verb) (Enfield 2006).

4. Obligatorification, associated with paradigmatic variability (Lehmann 2015[1995]: 146–152). The speaker's choice of an item within a paradigm becomes constrained. Diewald (2011) points out that in written German the "modal particles" are optional, nevertheless in speech, they are virtually obligatory. Similarly in English, although core DSMs like *and* and *but* appear fairly regularly in writing, they are almost obligatory in speech, where they are often used in more pragmatic functions. However, this does not apply to more specific DSMs, some of which, as we will see, may appear more frequently in writing than in speech (e.g. *moreover*), or arise in mainly high style texts (e.g. *by the way*).

   The two other processes of grammaticalization that Lehmann proposes are highly problematic for [[Conjunct] ↔ [DSM]] pairings. Of these, only the first concerned me in the Traugott (1997[1995]) paper:

5. Condensation, associated with the parameter of structural scope (Lehmann 2015[1995]: 152–157). Items that relate to constituents of arbitrary complexity, such as main verb, undergo condensation to items that modify a word or stem, such as an auxiliary or tense marker. The rise of DSMs by definition concerns the neoanalysis of a clause-internal item (often a Circumstance adverbial serving as an argument of a clause) as a [[Conjunct] ↔ [DSM]] microconstruction in pre-clausal position, where it has scope over the following clause D2 or even a larger discourse segment. Use in post-final position usually entails scope over the preceding clause or larger discourse, not reduction of structural scope.
6. Fixation, or reduction of syntagmatic variability (Lehmann 2015[1995]: 167–170). This concerns the fixing of a relatively free item, e.g. word order. Most DSMs other than *and* can occur in more than one position relative to the clause and therefore do not show evidence of fixation or reduction of syntagmatic variability. This is true of *after all* and *by the way*. Aijmer (1986) showed that *actually* (used largely as an epistemic DM) can be used before virtually any constituent. And *but* has recently been used concessively in a new, post-final position (Mulder et al. 2009).

Brinton (2017: 27–28) proposes that Hopper's (1991) principles of grammaticalization are more pertinent to the development of pragmatic markers than Lehmann's parameters, specifically:

a. Decategorialization: loss of the morphological and syntactic characteristics of the source, e.g. an adverbial source can be modified by an adjective (*by the winding way*), but not the DSM use. In the case of clausal Pragmatic Markers, the pronoun is restricted to 1st or 2nd person (*I mean, y'know*). This is related to attrition in 1.) above.

b. Persistence: retention of some trace of the original meaning, e.g. *by the way* retains some sense of path, even if only metaphorically and marginally; *after all* retains some sense of reasoning 'in the end'.
c. Layering and divergence: old and new constructions co-exist as the result of functional split, e.g. *after all* meaning 'after everything' coexisted until the end of the 19thC with concessive *She was right after all* (however, the adverbial is no longer attested; in COCA we find only *after all this*).

As Brinton (2017: 29) notes, Pragmatic Markers do not reduce phonetically or fuse with the host form in the way that decategorialization predicts, but remain independent items (Brinton 2017: 29). Importantly, they become more independent. So even Hopper's principles do not account adequately for the development of the subset of PMs that have DSM function. The main factor that remains to be accounted for is scope expansion: Conjuncts have scope over D2, but their source has only local clause-internal scope. This type of development runs directly counter to prototypical processes of grammaticalization.

Although they do not address the rise of DSMs, a further avenue of linking DSM development with grammaticalization might be found in Boye and Harder's (2007, 2012) proposal that grammaticalization is essentially a profile shift from "discursively primary" to "discursively secondary" meaning. In their analysis, a lexical *E* like *table* is conventionally focalizable (*it is a table that I want*), addressable (one can ask *Is that really a table?)*, and "discursively primary" (it is prosodically stressable). By contrast, a grammatical (and also a procedural) *E* is neither focalizable nor addressable except when cited (e.g. *There are too many 'after alls' in this article*). Grammatical *E*s are "abstract" and "discursively secondary" by convention (Boye and Harder 2012: 13). Furthermore, there is synchronic gradience between "primary" and "secondary", "contentful" and "procedural" status. Boye and Harder conclude that this synchronic gradience is the result of gradual grammaticalization. They suggest that the process of grammaticalization involves strengthening of the "situating" (contextual) relationship to the text or speaker and that "the 'situating' meaning … is naturally secondary to the meaning it situates" (p. 38). Once conventionalized, this situating meaning may come to be associated with the *E*, and a "part-time secondary" meaning may become a "full-time secondary" meaning via gradual bleaching or loss of contentful semantics. This approach to grammaticalization might appear to apply to the shift from CircAdv to Conjunct and of some monofunctional 1DSMs to DM status. However, like Hopper's principles, it does not account for gain of procedural meaning or for scope expansion. Furthermore, the stability of Conjuncts that would be labelled as "part-time secondary" would be difficult to explain. The term "partially secondary" seems more appropriate.

It is unquestionably true that certain changes long associated with grammaticalization, such as changes from lexical to procedural status, decategorialization and loss-and-gain phenomena are relevant to the rise of DSMs. An approach to the development of DSMs from the perspective of grammaticalization is valuable in spotlighting processes undergone in the shift from semantically contentful, truth-conditional *E*s that are arguments (typically CircAdvs) in a clause to non-truth-conditional, pragmatic *E*s that are "extra-clausal" and cue the relationship between D1 and D2. But while there is directionality toward pragmatic, procedural meaning, there is no UNIdirectionality in the sense of a steady shift toward increased reduction of the original signal.

There are too many aspects of the development of DSMs that are not consistent with grammaticalization as it is generally understood for an account in terms of that theory to be adequate. However, it should be noted that Degand and Evers-Vermeul (2015: 78) see "grammaticalization as the evolution of linguistic expressions from a more referential, lexically meaningful state to a more functional, elusive state, in which these expressions start to mark the clause, sentence or wider context in which they occur". Such a shift in thinking about grammaticalization is welcome, and the kind of approach I had in mind in 1995, but now that we have a constructional model for thinking about change that encompasses not only the changes Degand and Evers-Vermeul identify, but also more abstract schemas and networks, there is no longer a need to rethink grammaticalization in this way. Instead, reduction processes of grammaticalization can be regarded as typical of processes of change undergone by developing procedural constructions post-constructionalization.

5.3 The pragmaticalization hypothesis

Some scholars, e.g. Erman and Kotsinas (1993), Aijmer (1997, 2002), Frank-Job (2006) and Dostie (2009) have argued that "pragmaticalization" should be recognized as a change-type separate from grammaticalization and that it accounts better than grammaticalization for the rise of PMs, including what they call DMs. Degand and Simon-Vandenbergen (2011b) compare the hypotheses, and Heine (2013) provides a detailed summary and bibliography of work on pragmaticalization, so only the barest outline of the arguments is given here.

One reason for postulating pragmaticalization is that DSMs were considered at the time to be "agrammatical" (e.g. J. A. Goldberg 1980, cited in Brinton 2017: 31), or not grammatical in the way that e.g. tense, aspect and case are. Erman and Kotsinas (1993: 81–82) summarize the difference between "grammatical" and "pragmatic" as the difference between:

a. functions concerned with decoding, that is "guiding the addressee as to the interpretation of the message proper" (grammatical and subject to grammaticalization),
b. functions geared toward "ORIENTATION of the discourse and ORGANISATION of the conversation" (caps original) (pragmatic and subject to pragmaticalization).

Another reason Erman and Kotsinas (1993: 79) give for postulating pragmaticalization is:

c. the hypothesis that it is possible for "a lexical element to develop directly into a discourse marker without an intermediate stage of grammaticalization", such as *I mean, you know* (p. 81).

However, as was shown in Chapter 4.5.2, in connection with *after all*, and will be shown in later chapters, the development of lexical phrases into Conjuncts and of 1DSM functions into multifunctional DM functions involves accumulations of contextual modulations and gradual development. This is represented in the DSM Trajectory Hypothesis put forward in Chapter 4.4.2, where I argue for an intermediate step between lexical (CircAdv) and DM use. The loss of internal boundaries, or chunking, is a necessary feature in this Hypothesis. With respect to *I mean* and *you know* there is not only the loss of options like use of a modal (*you might know*), but also restriction to the personal pronouns (*she means* is not a pragmatic marker; it is typically used in intentional contexts like *she means to X*). Both of these changes are typical of grammaticalization and counter Erman and Kotsinas' proposal cited above that there is no intermediate stage of grammaticalization.

As Brinton (2017: 34) points out, "[t]he choice between pragmaticalization and grammaticalization seems to hinge not on the process itself but on what is encompassed by "grammar"" (see also Degand and Evers-Vermeul 2015). Diewald (2011: 451) argues, as I have done here, that pragmatic functions are "genuine grammatical functions which are indispensable for the organization and structuring" of discourse and "the fundamental features of grammar itself are rooted in pragmatics". This is because grammar encompasses deictic signs and indexicals, which "locate the utterance or elements of it with respect to the speaker" (Diewald 2011: 458). "Grammar itself may be seen as frozen pragmatic anchoring" and pragmaticalization as an integral part of grammaticalization (p. 461). Pragmaticalization, in Diewald's view, is a "sub-process of grammaticalization" (p. 458). In my view it is a name for procedural development the outcome of which is an *E* with mainly pragmatic function and with scope over discourse segments, not a separate process, and is therefore ultimately not theoretically valuable. The papers in Degand and Simon-Vandenbergen (2011a) almost unanimously reject pragmaticalization as an adequate account of the development of DSM functions. The perspective from

pragmaticalization has, however, been important in the history of work on the rise of DSM functions in throwing the spotlight on the orienting and organizational aspects of their meaning.

5.4 The hypothesis of cooptation to thetical grammar

In the last decade a theory of a dual-leveled Discourse Grammar has been proposed (see e.g. Kaltenböck et al. 2011; Heine et al. 2017; Heine et al. In press). It differs from construction grammar and many other models of grammar in that it has two main components: sentence grammar, and thetical grammar. Heine (2019) provides a detailed account of several other dual-level grammar proposals, including the distinction between micro-grammar and macro-grammar developed in Haselow (2013, 2016 and elsewhere). Heine (2019) also discusses evidence for the dual-level distinction in neurolinguist terms of left-right hemisphere asymmetries.

Sentence grammar, as conceptualized in Discourse Grammar, is restricted to the syntax of internal constituent structure. Kaltenböck and Heine (2014: 349) conceptualize sentence grammar as addressing speakers' need for a grammar that is "sufficiently fixed and constrained in order to work as system" of word order and other conventions that allow for "cohesive linear flow" and planning. According to Heine et al. (2017: 816) sentence grammar "is organized in terms of parts of speech or constituent types such as sentences, clauses, phrases, words, and morphemes, plus the morphosyntactic machinery to relate constituents to each other". More succinctly, "Sentence grammar is organized in terms of propositional concepts and clauses and their combination" (Heine 2019: 418).

Thetical grammar, by contrast, is conceptualized as providing material for spontaneously occurring discourse events and concerns "external" phenomena such as non-restrictive (appositive) relative clauses, imperatives, comment clauses, and DSMs. Theticals are "linguistic units beyond the sentence that are syntactically, semantically, and typically also prosodically detached from expressions of Sentence Grammar" (Heine 2019: 418). By and large they are the kinds of expressions that are regarded as "supplements" in Huddleston, Payne and Petersen (2002: 1350–1362). Key to Thetical grammar is the idea that linguistic units are "coopted to it", in other words units are instantaneously taken from sentence grammar by a "a cognitive-communicative operation whereby pieces of discourse located in one domain are transferred to another domain" (Heine et al. 2017: 813).

"Thetical grammar" is conceptualized as encompassing Conjuncts and DSMs, which the authors call "DMs". According to Heine et al. (In press: 2.3.1) a thetical has the following grammatical properties:

a. Its meaning is not part of the host sentence,
b. Its function is metatextual, relating the host utterance to the situation of discourse, more specifically to speaker-hearer interaction,
c. It is syntactically unattached as it is not a constituent of the sentence,
d. It is likely to be set off from the sentence prosodically,
e. It can be moved to other positions in the sentence.

Heine et al. (2017) group theticals into three types:

A. *Instantaneous theticals*: fully compositional units that can be spontaneously coopted any time from sentence grammar. Most of these are one-offs, e.g.:

(2) we did feel uh **union council was two weeks ago** when this was put to us that uh the increase from seven to sixteen uh was actually a very good idea.
(ICE:GB:s1b–075–68 [Heine et al. 2017: 819)

B. *Constructional theticals*: these are also compositional. They are coopted units, but unlike instantaneous theticals, they occur in "recurrent patterns" with "a schematic form or function" (Heine et al. 2017: 819), such as a reportative, quoting clause or a non-restrictive appositive relative:

(3) 'My name is Hindley Earnshshaw,' **slurred the drunk**, 'old Mr. Earnshaw's eldest son'. (2003 Fforde, *The well of lost plot* [Heine et al. 2017: 820])

C. *Formulaic theticals*: these are non-compositional, morphosyntactically unanalyzable, and usually short, e.g. *anyway, however, as it were*, and other DMs (Heine et al. 2017: 820).

Coopted units are said to be anchored in hosts. For example, the quotative in (3) (*slurred the drunk*) is anchored in the host quotation *My name is Hindley Earnshshaw, old Mr. Earnshaw's eldest son*, and the appositive clause *old Mr. Earnshaw's eldest son* is anchored in the NP, *Hindley Earnshaw*. Likewise DMs are anchored in the clause that they are associated with.

Formulaic theticals have the following typical properties (Kaltenböck et al. 2011; Heine et al. 2017: 817 and elsewhere):

i. they are syntactically independent,
ii. they tend to be set off prosodically from the rest of the sentence,
iii. they tend to be positionally mobile,
iv. their internal structure is built on principles of sentence grammar but they can be "elliptic",
v. their meaning is non-restrictive.

By "non-restrictive meaning" is meant meaning that is "not an inherent part of the structure of a sentence or its constituents". It relates to the "situation of discourse". Heine et al. (2017: 817) characterize "situation of discourse" as:

a. text organization,
b. source of information,
c. attitudes of the speaker,
d. speaker-hearer interaction,
e. discourse setting,
f. world knowledge.

The properties cited above are similar to those of PMs in general (see Chapter 4.2.1). Importantly for present purposes, Heine et al. (In press: Section 1.5) say that a thetical "has essentially the properties … proposed for DMs".

Heine (2013) points out that thetical grammar addresses many of the features discussed by proponents of pragmaticalization, such as increase in semantic-pragmatic scope, lack of fusion, and syntactic independence, and suggests that if pragmaticalization is "taken to refer to the mechanism of cooptation followed by grammaticalization … there is no problem in its use" (Heine 2013: 1239).

Most theticals are regarded as the result of transfer or "cooptation" from sentence grammar to thetical grammar. Exceptions are primary interjections such as *Oh*, and hesitation markers, presumably because these have no lexical or contentful cognates. In Heine et al. (In press: Section 2.3.1), drawing on several earlier works, cooptation is defined as follows:

> Cooptation is a fully productive operation whereby a chunk of sentence grammar, such as a word, a phrase, a reduced clause, a full clause, or some other piece of text, is deployed for use on the metatextual level of discourse processing, thereby turning into a thetical. Its functions are determined by the situation of discourse, serving (a) to overcome constraints imposed by linearization in structuring texts, (b) to provide the source of information, (c) to place a text in a wider perspective, e.g. by elaborating, proffering an explanation, a comment or supplementary information, (d) to describe the attitudes of the speaker, and/or (e) to interact with the hearer.

As this definition suggests, constraints on cooptation are essentially semantic: "What matters is that the speaker honors the principles of relevance to enable the hearer to … establish a cognitive link between the two parts of the utterance" (Heine et al. 2017: 830).

Cooptation is said to be a synchronic operation (Heine et al. 2017: 843). It is a mode of organizing discourse in "the semantic space provided by the situation of discourse within which the grammar operates" (p. 848). The proposal is that whereas grammaticalization is gradual, cooptation is instantaneous (Heine 2013: 1224; Heine et al. In press: Section 2.1). In many cases, there may be a tri-partite trajectory:

(4) internal grammaticalization – cooptation – further grammaticalization

If a DSM is derived from a phrasal adverbial with the structure [Prep NP], there will typically be some loss of integrity and decategorialization of the source and univerbation before the *E* is used as a DSM (e.g. *after all* attests to univerbation, and decategorialization of the quantifier *all* before use as a DSM). Further grammaticalization after cooptation can be seen in increased frequency and further loss of the original semantics. Reviewing Brinton's (2008) example of *I mean*, Heine (2013: 2031) finds "no convincing evidence … for a gradual transition" from sentence grammar to thetical grammar. Yet he cites examples provided by Brinton of "indeterminate structures" in Middle English in which *I mean* has its full content meaning used in initial position but is followed by a *that*-less complement. In such contexts, it is hypothesized, *I mean* can be interpreted either as a matrix clause introducing a complement clause or as the adverbial 'namely', as in (5).

(5) **I mene** Maister Geffrey Chaucer hath translated this sayd werke oute of latyn in to oure vsual and moder tongue.
'I mean/namely Master Geoffrey Chaucer has translated this said work from Latin into our usual mother tongue.'
(1477–84 Caxton, *The prologues and epilogues* 37 [HC; Brinton 2008: 125; Heine 2013: 1230])[31]

Several problems may be mentioned here with this particular analysis. Intermediate structures of "indeterminate structure" are among types of evidence for gradual change and therefore do not support the hypothesis of instantaneous cooptation. Further, Brinton (2008: 124–126) questions whether the PM *I mean* was in fact derived from a matrix clause introducing a complement clause since the latter use is very rare in the corpora she investigated. The source of PM *I mean* may well have been structures of the type *I mean X* where X is a "NP, VP, AP, PP, AdvP", but not a clause (Brinton 2008: 127). Brinton cites (4) out of context. In fact, putting (4) in its context, as attested in the Helsinki Corpus from which it is drawn, it is unambiguously of the type *I mean NP*, as can be seen in (6) (Traugott In Press: 96).

(6) Therfore the worshipful fader & first foundeur & enbelissher of ornate eloquence in our englissh. **I mene** Maister Geffrey Chaucer hath translated this sayd werke oute of latyn in to oure vsual and moder tonge.
'Therefore the worshipful father and first founder and embellisher of ornate eloquence in our English, I mean Master Geoffrey Chaucer, has translated this said work from Latin into our usual mother tongue.'

31. In EEBO the passage appears in Caxton's Boethius, *De consolatione philosophie.*

Here *Maister Geoffrey Chaucer* is appositional to *fader and first foundeur and enbellissher*, not the immediate subject of *hath translated* (the period is what would in present day punctuation practices be a comma).

The chief problem with the thetical hypothesis is that there is no evidence of instantaneous development. It is true that the first innovation was doubtless instantaneous, but so are all innovations, including those pertaining to "sentence grammar", e.g. the use of main verb with auxiliary function. In the case of *I mean*, the adverbial use involved gradual fixing/chunking of the 1st person singular present tense *mean*. In EEBO *I, you*, and *(s)he* are all attested collocating with *mean* from the mid-16thC on. The distribution and frequency are different in each case. For example, in the 1570s *I mean* (78 examples) collocates mainly with NPs and *that*-complements, *you mean* (11 examples) collocates with these same elements and purposive *to V*, whereas *he, she, they mean(s)* (7 examples for all third persons) collocates only with NP and *to V*. The difference in distribution would normally be considered to be a case of prior grammaticalization. As an adverbial meaning 'namely' it could be used, like Conjuncts discussed in this book, in a position that was ambiguously either initial, topicalized and lexical or pre-clausal and a partially pragmatic DSM.

Use in a position that is indeterminately initial or pre-clausal is precisely the kind of context in which change occurs gradually in the sense that there is the possibility of an imperceptible category shift. In Chapter 1.3, I mentioned the problem of distinguishing between initial and pre-clausal position and the impossibility in many cases of deciding as a researcher which interpretation was intended (though at the time there may have been different prosodic instantiations, which we cannot recapture). Desemanticization of *I mean* (use as a marker of self-repair), use in clause-medial and clause-final positions, and phonological reduction to e.g. [mi:n] are instances of later grammaticalization according to Heine (2013: 1230–1231). These can be interpreted as continuous processes that are typical, but not necessary, as DSMs become more procedural, more entrenched, and more frequent over time. In sum, the development of *I mean* points to gradual conventionalization, not cooptation, followed by changes that are consistent with continuous grammaticalization (and post-constructionalization constructional changes).

As mentioned above, in the view of Kaltenböck et al. (2011), Heine (2013) and others, "cooptation is in principle a unique, instantaneous operation" (Kaltenböck et al. 2011: 883, ft. 28) while grammaticalization, by contrast, is gradual. Another contrast that they invoke is that: "Whereas grammaticalization typically requires specific contexts to materialize, this is not the case with cooptation" (Heine et al. In press: Section 2.3.1).

With regard to instantaneous cooptation, in the case of "instantaneous theticals" as in (2) above, instantaneous, "on-the-fly" cooptation is abundantly clear. On

the other hand, in the case of "formulaic theticals", of which DMs are said to be a subgroup, it is not, because formulae take time to arise and be conventionalized in the sense of coming to be shared within a community of speakers. The same is true of "constructional theticals", which are defined as "recurrent patterns … having a schematic format and function" (Heine et al. 2017: 819). For a pattern to emerge is a gradual change process, even though it may be instantaneous for innovating individuals, and for it to be recognized and unconsciously entrenched within a schema also takes time (see Traugott and Trousdale 2013: 63–65; Petré 2019). Heine et al. (In press: Section 2.3.1) allow for conventionalization in their "Innovation model", which is essentially the same model for any innovation leading to change (innovation, propagation, conventionalization) as was proposed in Chapter 3.2.2 above and earlier in Croft (2000),[32] except that the particular kind of innovation that is cooptation is said to be "unique".

Even though there may be evidence of lateral hemisphere asymmetries such as are discussed in Heine (2019) and Heine et al. (In press: Section 8.1), even though instantaneous theticals such as (2) certainly do arise, and even though innovation by hypothesis is always instantaneous, there does not appear to be empirical evidence in the corpus data used here of any unique kind of instantaneous innovation independent of context in the case of DSMs. Nevertheless, the approach from Discourse Grammar is valuable in throwing the spotlight on the extent to which DSMs can function in a parenthetical way and, in the case of digressives, can signal that D2 should be interpreted as a parenthetical "aside" or as if it is parenthetical in the ongoing discourse.

5.5 The Diachronic Construction Grammar hypothesis

Because construction grammar is monolayered, and no distinction is made between sentence grammar and thetical grammar, there is no imperative to treat *E*s that conform to internal constituency as different in principle from *E*s that do not. Because constructions are assembled, different degrees of tightness of organization can be accounted for in a gradient from informal spoken conversation to formal, largely written, registers. In Chapter 2.3 it was mentioned that e.g. *Why didn't you go?* is conceptualized as the assembly of, among other constructions:

i. the proposition construction *You went for a reason*
ii. the negation construction
iii. the interrogative construction

32. Croft (2000: 4) identifies change with innovation followed by propagation. He goes on to characterize propagation as "essentially the establishment of a convention" (p. 98).

Example (3) above, *'My name is Hindley Earnshshaw,' slurred the drunk, 'old Mr. Earnshaw's eldest son'* could be conceptualized as the assembly of, among other constructions:

a. the proposition *my name is Hindley Earnshaw* (an instance of the predicate construction)
b. the report construction: *slurred the drunk*
c. the determiner construction: *the NP*
d. the apposition construction: *old Mr. Earnshaw's eldest son*
e. the possessive construction: *Mr. Earnshaw's son*
f. the adjectival modifying construction (twice): *old Mr. Earnshaw; eldest son*

There is no need to invoke a separate coopting process to account for the presence of the report (b) or the apposition (d).

From a Diachronic Construction Grammar perspective, the rise of DSMs is a case of the rise of procedural expressions. Not all procedural *E*s are "grammatical" items in the traditional sense of tense, aspect, modality and case. Even though there is evidence of bleaching, univerbation, and development of procedural meaning, DSMs are not usefully accounted for in terms of "grammaticalization" since they do not exemplify systemic reduction including scope reduction. They are pragmatic, but so are many other *E*s such as present tense used to narrate past events, so "pragmaticalization" is not a useful independent concept designed specifically to account for PMs. The difficulties in identifying instantaneous cooptation independent of context render "theticalization" a problematic account.

The concept of constructions and construction grammar is still being refined (e.g. Goldberg 2019), as is that of Diachronic Construction Grammar. Nevertheless, a constructional approach to the rise of DSMs accounts well for the changes observed in the corpora. It throws the spotlight on DSMs as signs with network relationships to each other and other sets of constructions and provides a way of specifying how particular features of the constructions involved are modified over time.

When an expression is first conventionalized as a [[Conjunct] ↔ [DSM]] it may be used in a way that is only marginally associated with its functional category, but over time its use may be better integrated into the category. This is particularly clear in the case of those Conjuncts that come to be used with DM function. The kind of shift in use over time from marginal to core status in a (sub)category described here is typical of new microconstructions in general, such as the *BE going to V* future, which was initially only a relative, not a deictic future, but later came to be used as a more central kind of deictic future. This kind of shift can be accounted for in terms of generalization over microconstructions and reorganization within a schema.

5.6 Summary comparison of the four approaches

In this section I briefly summarize and compare the four approaches mentioned above, with focus on the questions that research in each asks and what each brings to the issue of accounting for the rise of DSMs (and Pragmatic Markers in general). Because it is a summary, what follows will inevitably gloss over important points, but hopefully will highlight differences among perspectives and the potentials for further work.

Research on grammaticalization has traditionally asked: How do grams arise? For DSMs, the question is, are DSMs grams? The answer is No. However, like grams, they are procedurals and they belong to "grammar". That is, they cue SP/W's perspective on the relationship between the relevant units, in this case, D1 and D2. They function at the level of clause combinations. Historically, like grams, they undergo semantic loss of content (bleaching), decategorialization, and often univerbation, but unlike grams, most undergo very little morphophonemic reduction and, crucially, they involve scope expansion. They therefore are not well incorporated into grammaticalization studies, unless both the concept of grammar and the object of the research are stretched.

Research on pragmaticalization asks: How does a syntagma or word with propositional meaning come to be endowed with "essentially metacommunicative, discourse interactional meaning" (Frank-Job 2006: 397)? Proponents for the most part treat it as a process separate from grammaticalization. On the other hand, Diewald (2011) and Prévost (2011) have argued that pragmaticalization can be incorporated within grammaticalization. For DSMs the question is, is a separate process needed to account for them? Like Degand and Evers-Vermeuil (2015), I think not, provided grammar is understood to include conventionalized pragmatics.

Research on Discourse Grammar asks: How can linguistic units "beyond the sentence" that are detached from expressions of Sentence Grammar be accounted for? It is hypothesized that a dual-level grammar can account for them and specifically that they are coopted to Thetical grammar. For DSMs the questions are, do DSMs show evidence of instantaneous cooptation to Thetical grammar? The data discussed in this book provide evidence that they do not, given that the objective is to account for change understood as conventionalization.

Finally, research on Diachronic Construction Grammar asks: How do constructions come into being? As discussed briefly in Chapter 3.5, most phenomena that are researched in work on grammaticalization can be accounted for in a mono-level constructional model of grammar, except for some late-stage cases of phonological reduction. In addition, there is extensive concern for analogical processes, growth (and obsolescence) of patterns and schemas. Assuming with

Goldberg (2013: 16) that "semantics, information structure, and pragmatics are interrelated; all play a role in linguistic function", construction grammar embraces a perspective on grammar that can provide a rich account of DSMs (see also van Bogaert 2011).

The following chapters in Part II will provide extensive evidence in support of the Diachronic Construction Grammar position.

PART II

Case studies

CHAPTER 6

The development of elaborative markers

6.1 Introduction

One of Fraser's (1996 and elsewhere) classes of DMs is "elaborative markers", the core example being *and*. Biber et al. (1999: 80) analyze *and* as a coordinator and the other expressions discussed here (*also, further, furthermore, moreover*) as "linking adverbials" ("Conjuncts" in this book). Although the distinction is a fine one, Biber et al. point to two differences. For one, *and* and *but* are mutually exclusive, whereas this is not true of linking adverbials: **and but* is not a sequence except in cases of hesitation; however, *and also* is. As Biber et al. say, this constraint does not apply in the cases of *so* and *yet*, which can be preceded by *and* or *but*. A second difference that they draw attention to is that the position of coordinators "is fixed at the clause boundary" whereas this is not true for e.g. *also*. However, *but* has recently been used in post-clausal position in some varieties of English, as is exemplified in Chapter 7.2 and in Chapter 12.2.2. I follow the analysis proposed in Schiffrin (1987) that *and* and *but* can be used pragmatically as DMs in both her and my senses.

And and *but* are coordinators when they conjoin lexical categories such as nouns, verbs and adjectives, so a coordination construction needs to be recognized, but is not immediately relevant to the study of DSMs. Idiosyncrasies of positional use of some DMs can be specified as constraints on the micro-constructions in question, and for our purposes here, no categorial differences need be specified. Quirk et al. (1985: 635) characterize *also, further, moreover* as "reinforcing conjuncts". In Chapter 4.3.2 the general class of Connectors with [[Conjunct] ↔ [DSM]] structure was shown to license several subsets, including elaborative markers. They can be characterized as [[Conjunct] ↔ [Elaborator]]. In PDE the expressions *and* and *also* can be used as multifunctional elaborative DMs. However, *further, furthermore* and *moreover* are monofunctional and best categorized as elaborative 1DSMs.

The function of elaborative DSMs is to mark that D2 is to be understood as a continuation and expansion of D1. Discourse continuation can be made within the bounds of the argument, in which case it does not need to have attention drawn to it (see Mauri and van der Auwera 2012 on absence of clausal 'and'-markers in many languages). Fraser (1996: 340) suggests that elaboration signals "that the utterance following constitutes a refinement of some sort on the preceding discourse" and use of elaboration markers highlights this function. This is borne out by my data.

Among examples of elaborative DSMs that Fraser (1996: 340) cites (under the name DMs) are: *above all, also, alternatively, analogously, and, besides, by the same token, correspondingly, equally, further, further(more), in addition, in other words, more accurately, more importantly, more precisely, more to the point, on that basis, on top of it all, similarly, that is.* Members of Fraser's list cover a range of conceptualizations of elaboration. *And* and the set discussed in this chapter: *also, besides, further, furthermore, in addition, moreover* and *plus* are used to signal that D2 is in some way an addition to D1 and expands on it. Others express SP/W's rather specific mental mapping of the manner in which the relationship between D2 and D1 holds, e.g. *analogously, on that basis*, *more accurately*. In its earlier history, *also* did so as well, as discussed in Section 6.2. The focus in this chapter is on how markers of addition to and expansion of D1 come into being and elaborates on the methodological issue introduced in Chapter 4.4 of how to determine the status of an expression as a CircAdv or Conjunct on the form side or as a monofunctional 1DSM or multifunctional DM on the meaning side.

Historically many of the expressions Fraser cites include borrowed lexical items (cf. *alternatively, accurately, correspondingly, to the point, equally, in addition, on that basis, similarly*). Whether native (*also*) or borrowed (*alternatively*), many derive from the conceptual domain of similarity and difference (*also, alternatively, analogously, by the same token, correspondingly, likewise, in other words, otherwise, similarly*). As discussed below in Section 6.2, *also* derives from *eall-swa* 'in exactly the same way' and therefore is in origin conceptually akin to the set of similarity expressions. Some expressions suggest partial origins in quantifiers (*above all, furthermore, moreover*), and extension along spatial vertical and horizontal axes (*above all, moreover, on top of it all, further, furthermore, besides*).

The etymology of *and* is disputed but may be related to Latin *anti* 'in contrast, opposition' (see OED *and*, etymology). If so, in origin it is drawn from a conceptually different semantic space than the other Elaborators. In Old English it was used primarily to coordinate nouns and numerals, but it is also used as an elaborative clause Connector. Its history goes too far back and is too uncertain to be discussed here. Among other elaborative Connectors in Old English is *eac* 'in addition' (cf. German *auch*), as in:

(1) & on þam geare forðferde Raulf eorl on xiikalendas Ianuarii, & lið
and in that year died Raulf earl on 12calendars January, and lies
on Burh; **eac** gefor Heca biscop on Suðsexum, & Ægelric wæs on
in Burh; also died Heca bishop in Sussex and Ægelric was in
setl ahafen. (late 1000s *Chron D*.16. [DOE, *eac* 'in addition'])
camp captured
'and in that year died Earl Ralph on January 12th and lies in Burrow; also Bishop Heca died in Sussex and Ægelric was captured in his camp'.

Eac is spelled in a variety of ways. In examples that follow, it appears as *eac, oc* and, in Middle English, *eke*.

A further marker was *eallswa* 'likewise', the etymological source of *also*, which is the topic of Section 6.2. *Further and furthermore* are the topic of Section 6.3 and *moreover* of Section 6.4. Some other Elaborators (*besides, in addition* and *plus*) are mentioned briefly in 6.5. Section 6.6 concludes.

6.2 *Also*[33]

In PDE *also* is used primarily as a pre-clausal DM meaning 'in addition' and as a signal of a partial topic-shift that implicates D2 is a continuation of the larger topic at hand. It also implicates a small degree of emphasis on the following D2. Quirk et al. (1985: 635) say that in PDE, *also* used as a Connector is more specific than *and* and functions as a reinforcer. The strength of *also* is enhanced in (2a) by use of *do*, which intensifies *They use students*, and in (2b) by use of attention-getting *look*:

(2) a. So clearly there is an ongoing intelligence agency effort, and it is largely centered on journalists. And **also**, they do use students.
(2019 *Fox_Ingraham* [COCA])

b. That's the kind of precision I want to see from the Congress. I don't know if they'll do it. CUOMO. **Also**, look, I think, Mike, he can handle that pretty easily. (2019 *CNN-Cuomo* [COCA])

c. if the seas rise two, three, four feet, look at the real – trillions of dollars of real estate is in the way of that. **Also**, it's going to desertify.
(2019 *CBS_FaceNation* [COCA])

The history of *also* is somewhat different from that of the other elaborative DSMs discussed in this chapter as it originates in a manner adverbial rather than in a spatial or temporal adverbial (other manner adverbial sources are generally rather rare in English, but two manner digressives are mentioned in Chapter 8: *incidentally* and *parenthetically*.

Morphologically, *eallswa* is a combination of adverbial *eall* 'completely, exactly' with *swa* 'so'. *All* is still used as an intensifier, but with the weaker meaning of 'very'

---

33. For analysis of *also* in EEBO and later, searches were restricted to *also* followed by comma, as this provides the clearest potential evidence for DM use. Even so, the numbers are large (there are 89,774 hits of 'also,' in EEBO and 21,063 in COHA). As mentioned in the methodology section of Chapter 1, "post-clausal" is defined as occurrence at the end of a clause. This includes use preceding an *inquit*, or another clause, e.g. *and/but*-coordination, etc. Occurrence before phrasal modifiers was excluded. These are often participial as in *She stood up* ***also****, her lips tightening, under a rapid frown* (1920 Wharton, *Age of innocence* [COHA]).

in e.g. *She is all sad* (see Buchstaller and Traugott 2006). That *eall* and *swa* were originally independent and autonomous adverbials can be seen from examples in DOE where they are correlative, that is, *eall* introduces D1 and *swa* D2, as in (3):

(3) ða he cild wæs, **eall** hine man fedde **swa** man oðre cild
when he child was, exactly him one fed as one other children
fedeð (c1000 WHom 6, 164 [DOE])
feeds.

Usually the two adverbials appear combined as *eallswa* and are used as a manner adverbial meaning 'in exactly the same way'. In the manner adverbial use *eallswa* could be used in a variety of positions in the clause, such as clause-medial and clause-final. OED cites two examples with use in final position, one of which is (4):

(4) þet he hæfde læten his abbotrice for þet micelle unsibbe þet wæs on
that he had left his abbacy for that great unrest that was in
þet land .. oc hit ne wæs na ðe ma **eallswa**
that land … and it not was not the more in-exactly-same-way
'that he had left his abbacy because of the great unrest in the land … and it was not exactly in the same way any more as it had been'.
(1127 *Peterborough Chron* (Laud) [OED *also* A.1.a])

From about 1000 on, its main use is clause-initial, however, as in (5). To a modern reader it is ambiguous in this position between a manner adverbial 'in the same way' and a Conjunct meaning 'likewise'. Such examples typically attest to a constraint that D1 and D2 are syntactically and semantically parallel:

(5) Astriges … þe Bartholomeum ofslo awedde & on þam wodan
Astriges … who Batholomew-ACC slew went-mad and in that mad
dreame gewat; **ealswa** Egeus þe Andream ahencg, þærrihte on
frenzy died; likewise Egeus who Andrew-ACC hanged, right-away in
wodan dreame geendode.
mad frenzy ended.
'Astriges, who slew Bartholomew, went mad and died in his mad frenzy; likewise Egeus, who hanged Andrew, right away died in a mad frenzy'.
(c1000 Ælfric *ÆCHom* I, 35, 479.105 [DOEC])

During Middle English, phonological reduction to *also* occurred. MED points out that use as an adverb "connecting sentences" (Conjunct) and comparative conjunction were not clearly distinguished as *also* and the phonologically further reduced *as* until the 15thC. By the end of Middle English, s*wa* used on its own had been reduced and split into *as* 'like' and *so* 'therefore'. As a result, the original compositionality of *eall* 'completely' and *swa* 'in that way' became opaque and the form *also* came to be used as a Conjunct meaning 'similar to and in addition to D1' within

the frame [D1 Connector.Cxn D2]. The strength of *eall* 'exactly' was weakened on the semantic side and the requirement of strict parallelism between D1 and D2 was loosened on the syntactic side, presumably because of the loss of full compositionality. If a SP/W wished to draw attention to parallels, *eek* (OE *eac*, see (1) above) could be used. In (6) *and eek also* concludes a list of ways in which Melibee's wife Prudence tells her husband that he has erred:

(6) **Also ye han erred, for …** ye han ycleped straunge folk, yonge
Also you have erred, because … you have called foreign folk, young
folk, false flatereres, … and folk that doon yow reverence withouten love.
folk, false flatterers … and folk who do you reverence without love.
And eek also ye have erred, for ye han broght with yow to
And as-well also you have erred because you have brought with you to
youre conseil ire, coveitise, and hastifnesse,
your council anger, greed, and hastiness,
(c1390 Chaucer, CT, *Tale of Melibee* 1244–1246
[HCGW, translation from website])

All the same, as indicated above, *also* continues to be used with reinforcing pragmatics. The strength of *also* can be interpreted as a residue of the older meaning 'exactly' and as an example of the frequently observed phenomenon that old meanings may persist and constrain later ones (see e.g. Bybee and Pagliuca 1987; Hopper 1991).[34]

Unlike *and, also* can be used post-clausally. In fact, that is the preferred position in Middle and Early Modern English. In that position it no longer has manner adverbial meaning. There has been a split between the manner adverbial and the Conjunct uses. In post-clausal position *also* can have scope over a preceding NP or V, or, often ambiguously, over the whole preceding clause, as in (7), which is ambiguous between *also him, who was there* (narrow scope, conjunction), and 'in addition slew him that was there' (wide scope, Conjunct). Given the preceding clause 'come to London and slew hym that was ther', it is likely that wide scope and parallelism with the earlier construct was intended:

(7) but constantin come to london and slowe hym that was ther and after he went to wynchestre &; slowe him that ther was **also**, so that both his enemyes were deed.
'but Constantine came to London and killed him that was there and afterwards he went to Winchester and killed him that was there also, so that both his enemies were dead'. (1482 Caxton, *Chronicles of England* [EEBO]; Traugott In press: 105)

34. A residue of the original 'exactly' meaning can also be seen in its use as an adjunct adverbial that is a focus marker like *just, only, even, especially* (Quirk et al. 1985: 561; Biber et al. 1999; Pullum and Huddleston 2002: 586–588).

In earlier texts in EEBO *also* is used in pre-clausal position mainly in translations. (8) is an example from a treatise translated from French that provides lists of good behaviors for various professions, including knights and in this case surgeons. *Also* is typically used to introduce the last item on the list, thereby emphasizing it:

(8) the cyrurgyens ought also to be debonayr: amyable: &; to haue pytye of their pacyents: and **also** they ought not be hasty to launse and cutte apostumes and soores.
'the surgeons ought also to be courteous, amiable, and to have pity on their patients; and also they ought not to be in too much of a hurry to lance and cut ulcers and sores'. (1474 Cessolis, *To the right noble George Duc of Clarence*, trans. Caxton [EEBO])

By the 1540s *also* is found more frequently in texts written originally in English, like (9). (9) is structured in the same rhetorical way as (8): a list of good behaviors culminating in *also* D2:

(9) his maiestie commaundeth and streightely chargeth, all Persones, uicares, … to reade and declare to their parishioners, playnly, and distinctely, in suche ordre, as they stande in the boke, … **also**, hys maiestie commaundeth, that the sayde Ecclesiasticall persones, …
'his majesty commands and clearly charges all persons who are vicars … to read (the homilies) and explain to their parishioners plainly and distinctly in such order as they stand in the book, … also, his majesty commands that the said ecclesiastical persons …' (1547 Cranmer, *Certayne sermons* [EEBO])

Thomas Elyot, a widely read scholar, translated Latin *etiam* as '*also, more ouer, yea, ye forsothe* 'yes truly'' (1542 *Bibliotheca Eliota* [EEBO]). This suggests that *also* was understood in the 16thC as a Conjunct available to implicate epistemic certitude and confirmation of the content of D2. However, it rarely appears in initial/pre-clausal position in this meaning.

In her chapter on adverbial markers of addition, Lenker (2010: Chapter 11) finds that *also* was "not attested at all as an additive adverbial connector", i.e. in pre-clausal position, in her data from 1500–1920. (Her data are a small corpus of a million, two hundred words, based on the Helsinki Corpus (Lenker 2010: 11–12), whereas (9) is drawn from the 755 million word EEBO.) She says the Connector use was held in "low esteem" in dictionaries of the time, which often cite *furthermore, moreover, besides* as the central additives (Lenker 2010: 217). In COHA, *also* is preferred in post-clausal position until the 1900s. I searched the first 100 examples of 'also,' in COHA by decade from 1810–2000. In the first 100 hits of 'also,' as late as the 1890's there are only two examples in initial position, and they precede a phrase, not a clause. The frequencies changed rapidly in the first two decades of the 20thC

and by the 1920s pre-clausal and post-clausal uses are almost equal in number. Over the decades frequency of pre-clausal '*also*,' increases and by the 2000s we find 62 pre-clausal instances. In many cases, *also* signals a topic-shift that is nonetheless a continuation of D1. The same is true of COCA.[35]

To summarize the development of *also*, it originated in a CircAdv of manner specifying exact similarity of manner (*eall swa*). It came in Middle English to be associated with similarity of discourse content. The constraints on preciseness of similarity even of discourse content were weakened as *eallswa* was reduced in form to *also*. By the end of Middle English, *also* came to be used as a Conjunct meaning 'similar to and in addition to D1' within [D1 Connector.Cxn D2]. This constructionalization resulted in the addition of a new connective to the constructicon. Until the Late Modern period it was used primarily in post-clausal and clause-internal position as a focus marker that could specify a particular V or NP. However, it was available with low frequency and epistemic implicatures from Old English for use pre-clausally as a Connector to mark addition to D1. In the early 20thC there is corpus evidence that Connector use had become more frequent and by the 1920s this use appears to have become entrenched in pre-clausal position. *Also* is a DM because it is multifunctional and appears pre-clausally either to reinforce the coordination function of *and* or *but* or to express addition (on a positive upward scale) and topic-continuation. It has come to be a central member of the Elaborator subschema of the Connective.Schema, along with *and*.

The developments outlined above support the DSM Trajectory Hypothesis proposed in Chapter 4.4.2. The history of *also* is partially modeled in Figure 6.1. As in Chapter 4, only those properties that changed significantly are shown, and the contexts that enabled those changes ("assoc" is short for 'association'):

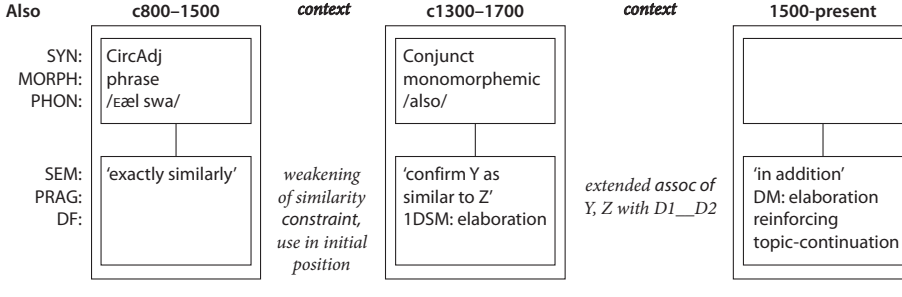

Figure 6.1 A constructional model of the development of DM *also*

35. In the BLOG 2012 section of COCA released in March 2020, almost all instances of 'also,' are pre-clausal.

As in the case of *after all*, modeled in Figure 4.6 in Chapter 4.5.2, the model shows that the shift from CircAdv to Conjunct syntax was accompanied by several meaning changes and was a constructionalization. The feature changes in the third template occur in the meaning component only, so extension to DM use is a constructional change that significantly reshaped the possible uses of the microconstruction but did not add a new construction to the inventory. What is not overtly shown, but can be deduced from the dates, is that the CircAdv and early non-multifunctional Conjunct uses obsolesced. That the DM use came to be the default use only during the 20thC is not shown. The frequency change that occurred in the 20thC can be thought of as a case of increased entrenchment, a strengthening of the symbolic link between Conjunction and DM pragmatics.

6.3 *Further* and *furthermore*

In this section I discuss the development of the CircAdvs *further* (6.3.1) and *furthermore* (6.3.2) into monofunctional 1DSMs.

6.3.1 *Further*

Further is the comparative of *forth*. *-Er* is the inflectional comparative morpheme (as in *warm-er*). The form *furðor* is used as a spatial CircAdv in Old English to refer to distance:[36]

(10) a. Hwæt wilt þu gesecan geond sidne grund feor oððe **furðor**
What, will you seek beyond spacious ground far or further
Þonne ðu geforþian miht (*Instructions for Christians* [DOEC])
than you accomplish might?
'What, do you wish to seek far beyond the spacious earth or further than you are able?'

b. He lædde me ða gyt **furðor**. and ic geseah þær ætforan us
He led me then yet further, and I saw there before us
mycele mare leoht. (c1000 ÆCHom II, 23 B1.2.26 [DOEC])
large bigger light.
'He then led me further still, and I saw a much larger light in front of us'.

36. MED notes in the entry for *further* that the alternate spelling *farther* does not appear until 1578. This spelling appears to have been the result of a kind of folk etymology after a phonological change that rendered *furth-er* non-compositional.

It is also used as a quantifier meaning 'more', especially in the context of *smeagan* 'to think, examine investigate', as in (11):

(11) we ne magon ne ne motan na **furðor** embe þis smeagan, gif we
we not may nor not must not further about this examine, if we
nellað us sylfe forpæran. (c1000 ÆLS (Christmas) B1.3.2 [DOEC])
not-want our selves to-destroy.
'we may not and must not examine this [the nature of God] further, if we do not want to destroy ourselves'.

OED cites use as a textual marker from around 1200 on in the context of verbs of saying:

(12) ȝiet hie seið **furðer** (c1200 *Vices and Virtues* [OED *further* 3])
still they say further

By hypothesis this use as well as that in (11) derives from the metaphor ARGUMENT IS A JOURNEY (Lakoff and Johnson (2003[1980]: 89–96) which was widely used in Latin source texts (see also *by the way* in Chapter 8.2).

In EEBO the first instance of pre-clausal *further* used as a textual Connector appears in a 1532 translation of Xenophon from Greek. Most examples are in translations from Latin, e.g. from Erasmus in 1538 and later. In the 1550s we find proclamations in which *further* is used as a Conjunct to combine lengthy legal requirements. (13) exemplifies use in religious argumentation:

(13) and this is Christes naturall bodye in the sacrament, a figure of his
and this is Christ's natural body in the sacrament, a figure of his
misticall bodye the Churche, and of the vnitye of the same: **and further**,
mystical body the Church, and of the unity of the same: and further,
where as our Lorde commaunded his disciples and all vs to doo the
where as our Lord commanded his disciples and all of-us to do the
same that he didde,…
same that he did…
(1558 Watson, *Holsome and catholyke doctrine* [EEBO])

Connective *further* can be strengthened by *also*. There are 5 examples in EEBO, coordinating complex clauses. Only one (14) meets my criteria for DSM status of preceding a finite clause:

(14) so that the true subject of Ecclesiasticall and civill justice can not rightly be divided: **further, also** it is as manifest that…no assured peace can continue.
(1641 Parker, *A discourse concerning Puritans* [EEBO])

PDE *further* is cited as "formal" in Quirk et al. (1985: 635). Although it is attested in COCA SPOKEN, it occurs in reports rather than in interactive talk, e.g.:

(15) But we do know another frontrunner is Brett Kavanaugh, who was a clerk for Justice Anthony Kennedy, who announced his retirement this week. **Further**, there's a dark-horse candidate that the president talked about during that gaggle. (2018 *CNN_Newsroom* [COCA])

Used as an Elaborator, *further* is relatively non-compositional and monomorphemic. It appears not be modulated by context, including position, so it is not versatile and is a monofunctional 1DSM. It is also available as a verb (*to further their development*) and as an adjective (*led to further unpleasantries*). In this regard it is unlike *furthermore* (see below), which can only be used as an adverb.

The development of *further* can be partially modeled as in Figure 6.2:

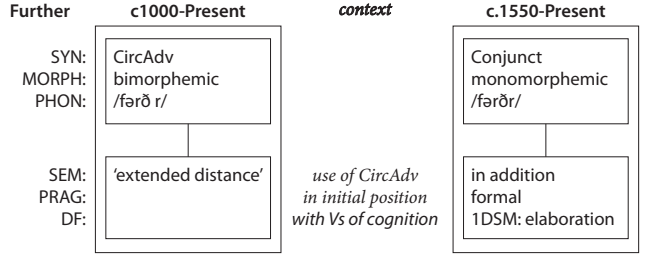

Figure 6.2 The development of *further*

6.3.2 *Furthermore*

Structurally, *furthermore* is in origin a double comparative consisting of *forth* + inflectional comparative *-er* + analytic comparative *more*. Like *further*, *furthermore* is characterized as "formal" in Quirk et al. (1985: 635), and as "academic" in Biber et al. (1999: 887).

Furthermore is attested first in Middle English as a double comparative spatial CircAdv. In (16) it has the form *forth* + InflectionalComparative and is followed by *more*:

(16) Þe sterrne comm rihht till þatt hus .. & flæh itt ta na **Forrþerr**
the star came right to that house .. and flew it then not further
mar, Acc heng þæroferr stille. (c1175 *Orm*. 7338 [OED])
more, But hung thereover still.
'The star came right to the house and then flew no further, but hung over it without moving'.

In later Middle English *furthermore* obsolesced as a spatial adverbial and was usually used as univerbated elaborative Connector by the end of the 14thC. This is a century and a half earlier than *further*. OED cites an example from Wyclif c1380, but provides no prior context. MED defines one of the uses of '*ferther-more*' as "introductory or connective adv.: moreover, furthermore, also". An example is:

(17) **And therfore seye I that it is good as now that ye suffre and**
And therefore say I that it is good for now that you suffer be
be pacient. Forthermoore, ye knowen wel that …
patient. Furthermore, you know well that …
(c1395 Chaucer, CT, *Tale of Melibee* 1480–81)

In EEBO, *furthermore* appears mainly in translations of Hebrew and Greek Biblical texts. This will have made it familiar and probably accounts for its use in a slightly wider set of genres than *further*. For example, it was used in Blount's *Natural History*:

(18) this strong liquor is ordinarily drank amongst the planters, as well alone, as made into punch: **furthermore**, when this juice hath so boil'd into the two first coppers,… (1693 Blount, *A natural history* [EEBO])

Like *further*, *furthermore* can be strengthened by a following *also* (5 examples in EEBO), e.g.:

(19) Parentes ought first of all [to shield their children] from all vnhonest games: **furthermore, also** their Parentes shall admonish them that they giue no credite to coniurations, witchcraftes, sorceries and such like vayne trifles:
'Parents should first of all shield their children from indecent games: furthermore too their parents shall warn them not to give credit to conjury, witchcraft, sorcery or similar trifles'.
(1581 Lowth, *The Christian man's closet* (trans. from Latin) [EEBO])

The development of Conjunct *furthermore* differs from that of *further* with respect to morphology, dating, and persistence of the CircAdv use, as shown in Figure 6.3:

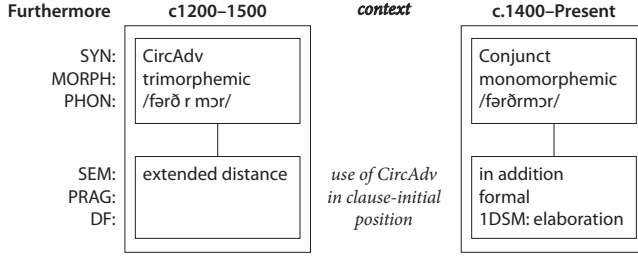

Figure 6.3 The development of *furthermore*

To summarize, both *further* and *furthermore* are preferred in formal registers. Both originate in the same comparative spatial adverb *further*. In PDE *further* is used as both a directional CircAdv and as a Conjunct in clause-initial position. However, *furthermore* is rarely used with spatial meaning after the 13thC (the last example cited in OED is dated 1554) and is now no longer used with spatial semantics; it is used exclusively as a textual Conjunct. Their histories support the first part of the DSM Trajectory Hypothesis proposed in Chapter 4.4.2.

6.4 *Moreover*

Despite *over*, *moreover* is not a spatial expression in origin. According to OED the oldest use of *moreover* is temporal (13thC), meaning 'for a long time, longer' (OED *moreover* adv. A.1.a). This use is said to be 'obsolete, rare' now. The temporal origin is confirmed in MED, which cites the same example as OED, but with more context:

(20) Quað pharaon to moysen, 'Nu ic rede ðat ge flen; For se ic gu
Said Pharaoh to Moses, 'Now I advise that you flee; for see I you
more-ouer nu, Dead sal me wreken ouer gu.'
longer now, death shall one wreak over you.
'Said Pharaoh to Moses: 'I now advise you to flee, because if I see you any longer now, someone will kill you".'
(1325 (c1250) *Gen and Ex.* (Corp-X 444) 3119 [MED *more-over* b])

Use '[a]t the beginning of the clause … indicating that it contains matter additional to what has already been stated" (OED *moreover* Adv 2.a) is cited in both OED and MED from the 1380s on, for example:

(21) And also, **more-ouer**, me þynkeþ … Men sholde constreyne no clerke
and also, moreove**r**, me thinks … men should constrain no cleric
to knauene werkes
to peasants' works.
'And also, moreover, I think … people should not force any cleric to do peasants' jobs'. (c1400 (?a1387) Langland, *Piers Plowman* C.vi.53 [OED, MED])

In (21) *moreover* is used to reinforce both *and* and *also*, itself a reinforcer, emphasizing the strength of the opinion. In EEBO it is found in Biblical works, such as Tyndale's homilies. In (22) he exhorts his listeners and readers to do deeds of mercy so that:

(22) our deades declare howe we love our neyghbours ad [*sic*] how much we
our deeds declare how we love our neighbors and how much we
have compastion on them at ther neade: **moreover** it is not possyble to
have compassion on them at their need: moreover it is not possible to
love except we se a cause.
love unless we see a reason.
(1528 Tyndale, *Fayth the mother of all good works* [EEBO])

Like *further(more), moreover* continues to be used as an elaborative Connector. OED comments that in PDE it is unusual in speech. It is, however, attested in blogs and in the relatively formal speech of the SPOKEN section of the COCA corpus. In (23) both examples are from radio interviews:

(23) a. They don't want to give up hope. KAUFMAN: **Moreover**, he argues that doctors have no incentive to see their patients die prematurely.
(1994 *NPR_ATC* [COCA])

b. This was pure propaganda in a war of words with the United States. **Moreover**, you know and everybody else here knows that the value of the hostages is the value to which the holder of those hostages can use them.
(1990 *CNN_Crossfire* [COCA])

Overwhelmingly, the register in which *moreover* is used in the corpora is academic or reportative.

Like *further* and *furthermore, moreover* undergoes little modulation in context and appears still to be a monofunctional Conjunct in all positions. Lenker (2010: 250) suggests *furthermore, moreover* (and also *overmore*) "do not signal cognitive complexity, but are iconic in their length" and "expressively highlight the semantic relation which is already operating in the linear sequence of text".

The development of *moreoever* can be partially represented as in Figure 6.4:

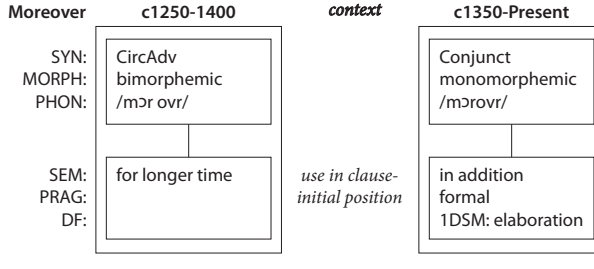

Figure 6.4 The development of *moreover*

Like *further(more), moreover* supports only the first part of the DSM Trajectory Hypothesis.

6.5 Other elaborators

The [[Conjunct] ↔ [Elaborator]] subschema has been available to SP/Ws from Old English times and before. This subschema licensed mainly *eac* 'also' and *and* in Old English. Over time *also*, *furthermore, further* and *moreover* were added (also *overmore*, Lenker 2010: 93), but *eac* was lost in standard English. Other additions to the subschema include *besides*, *in addition* and, recently, *plus*.

Besides, which OED says is the "usual" expression in speech by comparison with *moreover* (OED *moreover* Adv 2.a.), implicates that D2 is not especially important, a meaning that persists from origins in a spatial body part expression (*be* 'by' + *side* + adverbial *-es* (Rissanen 2004; Traugott and Trousdale 2013: 110–111)). In EEBO *besides* is used mainly as a preposition, or clause-finally as a spatial adverbial. Among prepositional complements are anaphoric pronouns (*this, that*) referring to prior discourse as in (24):

(24) [Citing St Austen] fro the wc heresie god deliuer his faythfull:
from the which heresy god deliver his faithful:
besides this s. austen doth saye …
besides this St. Austen does say …
(1533 Frith, *A book made by John Frith* [EEBO])

Besides came to be used occasionally without *this* as a Conjunct meaning 'in addition' in EModE. The first example in EEBO is:

(25) but it myght haue ben called a collusio, for prophaning of gods holy worde, and deceiuing the simple people: **besides**, whatsoeuer in their owne conceiptes they hadde taken for a veritie, the same wold thei proudly vtter and saye …
'but it could have been called a collusion to profane God's holy word and deceive the simple people: besides, whatever they had taken for a truth in their own minds, that is what they would proudly utter and say …'
(1556 Huggarde, *Displaying of the Protestantes* [EEBO])

By the end of the period covered in EEBO pre-clausal use predominates. Pre-clausal use predominates in COHA as well. There is, however, no evidence of DM use of b*esides* at any period.

In addition also began to be used in EModE (the first example in EEBO is from 1566). However, it appears very infrequently in EEBO. *In addition* is mainly used in a prepositional phrase (e.g. 1640 *in addition to her other miseryes*). A text-related lexical use is:

(26) thou shalt not forswear or perjure thy self, but shalt perform unto the lord thy vows: but **in addition to this** i say unto you, swear not at all in your common converse.
'you shall not swear falsely or perjure yourself, but shall perform your vows to the Lord; but in addition to this I say do not swear at all in your ordinary conversation'. (1681 Kettlewell, *The measures of Christian obedience* [EEBO])

In (27) and in most of the examples in COHA, *in addition (to this)* is used to link properties of the referential world rather than to signal the relation of upcoming D2 to D1 that SP/W wishes AD/R to understand:

(27) Her ready docility would of itself have been sufficient to surprise Lord James. But, **in addition**, there was a soft note in her voice.
(1911 Bennet, *Out of the primitive* [COHA])

In (27) the *soft note* is presented as perceived by Lord James in addition to her docility, not by the narrator of D1. Note that none of the Conjunct expressions discussed in this chapter: *besides, furthermore, moreover* can be coherently substituted for *in addition* after *but* in (27). This suggests that *in addition* has a lexical listing CircAdv function rather than a text-conjoining function in (26) and (27).

By the mid-20thC a few examples of *in addition* appear that can be interpreted as text-linking Connector uses, e.g.:

(28) But the best of him, his work, remains. **In addition**, those of us who were fortunate enough to know him have a legacy of rich remembrance.
(1966 Kanin, *Remembering Mr. Maugham* [COHA])

Here substitution by any of the Conjunct expressions discussed in this chapter allows for coherence.

A yet more recent elaborative Connector is *plus*, as in:

(29) This is WONDERFUL with the recipes all in one place (love that I didn't have to print out 8 different recipes) and the shopping list already done. **Plus**, everything looks really good! (2012 *Six Sister's stuff* [COCA BLOG])

In COHA *plus* is first attested as the mathematical term. Beginning in the 1900s it is used textually to mark NPs as members of additive lists, as in (30):

(30) this intensified competition will go on for a number of years until the amount of national production will be scaled down to the volume of what we ourselves can consume, **plus**, of course, whatever export trade we may have.
(1921 Filene, *Give us this day our daily bread* [COHA])

In the 1950s the NPs have been expanded to include verbal gerunds, as in (31): *having the fighters on you, and people shooting at you from the ground*:

(31) "All it would take to make it real would be to hang the big bomb on and swing the compass needle 0,.. 21 "a little," one weary pilot said, after a six-hour flight to a simulated target deep in Greece. "**Plus**, of course, having the fighters on you, and people shooting at you from the ground,"
(1957 Martin, *Cat Brown's kittens have claws* [COHA])

(32) is the first example in COHA of *plus* used to signal that a finite clause in D2 is an addition to and elaboration of D1:

(32) the mud-man on the ground waving his fist and swearing into zeppelin clouds would [b]e the same. **Plus**, we fished together, and were aficionados of a sort that way. (1977 Carlson, *Betrayed by F. Scott Fitzgerald* [COHA])

By the decade of the 2000s there is a large number of examples of *plus*, many of which combine clauses, suggesting that Conjunct use of *plus* had become well established. *Plus* is a good example of a 1DSM that came into being fairly recently in a gradual extension of contexts, in this case from NP referring to an inanimate entity to a gerund with verbal properties and finally to a clause.

6.6 Conclusion

This chapter has shown that an elaboration subschema [[D1 Elaborator D2] ↔ [D2 is presented as an elaboration of D1]] existed already in Old English. It exemplifies well the dynamic, changing nature of (sub)schemas because it has been relatively productive even though *eac* obsolesced and *overmore* was short-lived. In Standard English the subschema now licenses (in alphabetic order of those discussed in this chapter) *and, also, besides, further*, *furthermore*, *in addition, moreover*, and *plus* in some of their uses; these range from lexical Connectors like in *addition* to 1DSMs like *further*, and DMs like *and*. Table 6.1 shows in which periods the Elaborators mentioned in this chapter were or came to be used with any frequency, or obsolesced.

Table 6.1 Elaborative marker resources over the history of English

Period	Available in subschema	Loss from subschema
OE	*and, eac*	
ME	*and, eac, furthermore, overmore*	
EModE	*and, furthermore, further, moreover, besides*	*eac, overmore*
20thC	*and, furthermore, further, moreover, besides, also, in addition, plus*	

A subschema such as the Elaborator subschema may be conceptualized as a network that may expand and contract internally in constructional space. As we will see in later chapters, especially Chapter 13, this constructional space is linked externally with other (sub)schemas, including other DSM schemas.

The chapter has also shown that the constructionalization phase of the DSM Trajectory Hypothesis (CircAdv > Conjunct) is borne out in the history of all *E*s discussed, but only *also* underwent further reshaping to a DM through constructional changes.

As mentioned in Section 6.1. Fraser (1996) includes a considerably larger number of *E*s than have been discussed here among his examples of Elaborators that are DMs in his sense. Some of them have not been discussed here, because, although they can be used with a semantic connecting function to signal that the upcoming discourse continues and elaborates on what precedes, they are highly contentful and not Conjuncts, among them *alternatively, more importantly, on top of it all*. Similarly, structural differences among sets of *E*s have been discussed for Catalan. Cuenca (2015) distinguishes "grammatical connectives" of various types like *and*, "linking endophoric devices" that point back anaphorically ('that's why') or forward cataphorically ('namely'), and "lexical connectives" that "are functionally equivalent to grammatical connectives but still exhibit a proper syntactic structure and compositional meaning". She exemplifies the latter with English *to take just one example* and *first* and a wide range of Catalan markers. Among them are elaborative topic change markers like *més coses* 'more things', *deixin-m'hi afegir* 'let me add' (Cuenca 2015: Examples (12) and (13)). In Cuenca's terminology, *also* would be a "grammatical connective", *on top of it all* a "lexical connective", and *further* would be in between, relatively contentful, but nevertheless a DSM of the type that I have labelled 1DSM.

In the next Chapter I turn to Contrastives like *but*, *all the same* and *instead*.

CHAPTER 7

The development of contrastive markers

7.1 Introduction

When a speaker or writer uses a contrastive Connector such as *but, all the same* or *instead*, it is to signal that 'the following statement contradicts an inference that I would expect you to make' (Winograd 1976: 289, discussing *but*). "*But* presents two conjuncts which clash with each other in some way – it is contrary to our expectations to see the two presented side-by-side" (Sweetser 1990: 100). This clash is based in SP/W's (represented) belief about the real world, not in the real world itself (Sweetser 1990: 104) (see Chapter 2.2 for the distinction). The main purpose of the current chapter is to outline the development of the DM *but* (Section 7.2). Two other contrastive markers, *all the same* and *instead*, are more briefly discussed in Sections 7.3 and 7.4 respectively. They are monofunctional DSMs. Similarities and differences between the Connectors are discussed in Section 7.5. In the course of the chapter I will further illustrate the importance of discourse contexts in the rise of new uses.

7.2 *But*

Some general comments on functions of *but* are made in Section 7.2.1. The history of *but* is sketched in Section 7.2.2.

7.2.1 Background

In PDE *but* is used as a coordinator and a Connector. As a coordinator it is used to coordinate members of the same category and to implicate that some expectation is being cancelled. For example in (1a) *failure* and *aim* are both members of the category noun, and the expectation that is evoked and cancelled is that failure is a sin.[37]

37. An anonymous reviewer pointed out that the traditional view that *but* can be used to coordinate NPs has been shown to be problematic (see Chaves 2007).

(1) a. used to tell students that "not failure **but** low aim is sin".
(1992 *Ebony* [COCA])
b. He's low-key, **but** strong in "Raiders of the Lost Ark".
(1995 *CBS_Morning* [COCA])
c. Crombie wasn't working Sunday **but** started to get calls and texts asking if she was safe. (2017 *OregonLive.com* [COCA])

The topic of this section is use of *but* as a clause Connector instantiating a contrastive subschema [[D1] Contrast [D2]], where D2 is a finite clause with a subject, as in (2):

(2) A lot of people worry about crossing editorial and advertising lines, **but** I think it respects readers more. (2012 *digidday.com* [COCA WEB])

As Fraser (2009b) discusses, the kind of contrast conveyed may differ considerably in different contexts. Taking a largely Relevance Theoretic view of DMs and arguing that they are monosemous, Fraser (2009b: 293) suggests that there is one core meaning to *but* (Contrast) and 10 uses "signaled by context and pragmatic elaboration". An example of explicit Contrast is (3) (Fraser 2009b: 310):

(3) John likes to dance **but** *I* like to read.

Here *John* and *I* belong to the same set (people) and the contrast can be made explicit by adding *in comparison*, *by contrast*. Implicit Contrast is exemplified by (4) (Fraser 2009b: 313):

(4) We started late, *but* we will arrive on time.

Here *We started late* could be assumed to implicate late arrival; the implicature is contrasted and cancelled. After discussion of several different kinds of Contrast, the example in (5) is introduced (Fraser 2009b: 316):

(5) a. I had a lovely evening last night with Harry.
b. **But** did he repay you the money?

Fraser goes on to say "The fact that there is no contrastive result from comparing S1 and S2 follows from the fact that this case is not a DM use of *but*. Whereas a DM signals a semantic relationship holding between S1 and S2, the *but* in these examples is signalling a change in discourse topic, not a semantic notion".[38] In Fraser (2009a) he considers change in discourse topic to be discourse management marker use. As discussed in Chapter 4, I do not make a distinction between DSMs and

38. "S1" and "S2" are short for "segment 1" and "segment 2" (and equivalent to D1 and D2).

discourse management markers. Furthermore, DSMs and the subclass of DMs are in my view not semantic, though a DSM may be partially so, especially when it is first constructionalized.

What is important for this chapter is that *but* is multifunctional in contemporary use, and therefore a DM, unlike the other two contrastives discussed. From a constructional perspective, the meaning [contrast] is inherited as a discourse function from the [D1 Contrast D2] frame. Other functions are modulations in context that include self-correction, correction of others, objection, topic management and discourse management, such as introducing a summary (Lenker 2010: 207). These are illocutionary acts of various types, and contrastive *but* combines with them in ways that Fraser (2009b) lists as separate uses.

A constraint on form is that, unlike some other DMs, *but* is used in pre-clausal position only in Standard English. However, a conventionalized use has recently been observed in post-clausal position in conversation (e.g. Mulder and Thompson 2008; Mulder et al. 2009). This particular use will be introduced in Chapter 12.2.2.

7.2.2 The history of *but*

In PDE *but* is the default DSM used to signal Contrast, but in Old English the default contrastive was *ac*. *Ac* was replaced with *but* by the end of Middle English (c.1500). The etymological source of *but* is the Old English preposition *butan* 'outside' (< *be utan* 'on the outside'), which could be used as a spatial adverbial. It contrasted semantically with *binnan* 'inside' (< *be innan* 'on the inside').

In the context *ealle butan anum* 'all outside of/excluding/except one' *butan* came to be used as an exclusive scalar marker by the 13thC. It expresses an evaluation, putting upper bounds on the complement (Nevalainen 1990: 342), cf. *all but dead* 'nearly dead'. As an exclusive it is part of a schema that in PDE includes *almost, barely, nearly, scarcely* (Nevalainen (1991: 125–127). This use of *but* declined in frequency from the mid-17thC on (Nevalainen 1990: 345), but is still found in expressions like *Nothing but X will do*.

Used pre-clausally *but* could be used to mean conditional 'unless', a meaning that was optionally strengthened in Early Middle English by *if (that)*. In this use it is a subordinator:

(6) swore by Almyghty God and by his holy names, **but if þat** he wolde go
swore by Almighty God and by his holy names, unless he would go
wiþ ham, he shulde be dede. (c.1200 *Brut* p. 217.1333 [HC])
with them he would be dead.

Use as a coordinator is attested from the 1300s on, usually in the context of a negative D1, where contrast is expressed overtly, as in (7). This is what Fraser (2009b) calls an explicit contrast:

(7) Wolde þe erl nouth dwelle þore **but** sone nam until his lond
Wanted the earl not dwell there but straight-away went to his land
'the earl did not want to live there but straight-away went to his own land'.
(c1300 *Havelok* (LdMisc 108) 2929 [MED *but* conj., 6c])

By the 16thC the negative context constraint begins to be disregarded. In (8) *but* is used as a Connector to invite AD/R to understand that SP/W conceptualizes a contrast significant enough to be worthy of mention between the door posts and the door:

(8) then wente he in / and measured the dore postes, whiche were two
then went he in / and measured the door posts, which were two
cubytes thycke; **but** the dore it selfe was syxe cubite.[39]
cubits thick but the door it self was six cubit.
(1540 *The Byble in Englysshe* [EEBO])

The fact that the negative marker came to be regarded as redundant and unnecessary is far from unique to *but*. In the case of *but*, contrast is inherently negative ('not that but this'), so a negative marker is redundant. A recent example of non-use of a redundant negative is *I couldn't care less* > *I could care less*, where *less* is inherently negative in meaning. This use has been criticized as "senseless" or "illogical", but the change is entirely reasonable – redundancy avoidance is usually considered logical. It is in fact a key stage in the much discussed "Jespersen Cycle" of negation (e.g. van der Auwera 2009), as in French *ne* > *ne pas* > *pas* and English *ne* > *ne naht* > *not.*

This severely truncated history of *but* reveals several things, among them:

1. The source of *but* is a "real world" spatial ('outside'). This literal use has obsolesced, but persists pragmatically in so far as SP/W signals that D2 is metaphorically "outside of" what might be considered to be directly coherent with D1 or directly inferred from the pragmatics associated with D1.
2. As a preposition, *butan* prototypically had narrow syntactic and semantic scope over NPs, AdjPs, and VPs. Used as a Connector it has scope over the following clause.
3. *But* has been used discursively in a variety of ways to combine units, not only as a contrastive coordinating Connector but also as a conditional subordinator.

39. A 'cubit' is a measurement approximately equal to the length of the forearm, about 18 inches.

4. Used to counter a position expressed in D1, *but* signals that the topic is a kind of elaboration of D1, but from a different point of view. The shift in point of view entails a topic shift.
5. The conventionalization of *but* as a clause Connector is a case of a constructionalization as a microconstruction that is [−truth-conditional, +conventional]. Its conventional meaning is 'signals contrast'.

From a constructional perspective we may note that:

a. From Old English on, there have been a number of schemas, including spatial, scalar, conditional and contrastive and elaborative constructions. As a micro-construction, *but* was used in earlier English as a spatial or as a scalar preposition to instantiate spatial and scalar schemas. It came to be used with *if* as a subordinator, but since the 14thC its prototypical use has been to instantiate the Contrast.Subschema. Pragmatically weaker uses (e.g. (5)) show functional overlap with the Elaborator.Subschema.
b. The later meanings show evidence that the source meaning 'outside' persists; it constrains later uses in that contrastive expressions derived from *butan* can only be licensed by schemas that plausibly involve an abstract 'outside' meaning (note that 'outside' is not a source meaning for e.g. *instead, by contrast*).

These findings show that physical world *butan* 'outside' has over time been reconceptualized in terms of the discourse world (see Sweetser's 1990 metaphorical analysis of *and*, *but* and *or* mentioned in Chapter 2.2). In this world it marks:

i. a referent of a phrase as outside a set (*all but one* 'all except one' implicates 'one is outside the set X'),[40]
ii. a clause as conditional and subordinate (outside the circumstance X, Y),
iii. a clause as in some way outside normal coherence expectation (coordinator *but*).

Figure 7.1 provides a very partial model of the changes that led to the rise of DM *but*. It shows that the shift of meaning from 'outside' to 'except' was a neoanalysis and constructional change since the meaning not the form changed. It also shows that the shift from preposition to Conjunct was a constructionalization. The latter change resulted in the addition of a new Conjunct to the inventory, which eventually came to be preferred to *ac*; it also resulted in the restriction of the preposition *but* to a small niche of exceptive expressions. The further shift to a DM was a constructional change. The phonological changes from disyllabic *butan* to

40. Note that *except* < Latin 'out taken'.

monosyllabic *but*, and of 'short' /U/ to /^/ were systemic constructional changes in Standard English. They too did not result in addition of a micro-construction to the inventory and therefore are considered not to be constructional changes. "Prep" is short for 'preposition', "btw" for 'between'; as in earlier figures "1DSM" is short for 'relatively monofunctional DSM' and "DM" for 'relatively multifunctional DSM'.

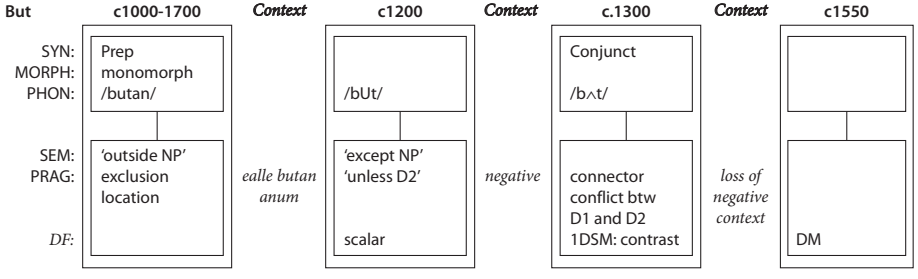

Figure 7.1 Partial model of the rise of DM *but*

As mentioned above, during its history *but* has intersected with several conceptual domains or schemas. The prototype *E*s they license at various periods that are relevant over time to the various uses of *but* are summarized in Table 7.1 ("Sch" is short for 'Schema', "scalar.exclus" for 'scalar exclusivity', and "elab" for 'Elaborator'):

Table 7.1 Some conceptual schemas instantiated over the history of English by *but*

	Spatial.Schema	**ScalarExclus. Sch**	**Conditional. Sch**	**Contrast. Schema**	**Elab. Schema**
OE	*butan* 'ouside of'	*butan* 'except'	*gif* 'if'	*ac*	*eac*
ME	*butan*	*but* 'except'	*but if*	*ac, but*	*eac, and*
EModE	*outside*	*but 'except'*	*unless*	*but*	*but, and*
ModE	*outside*	*but* 'except'	*unless*	*but*	*but, and*

A full constructional understanding of the development of *but* would necessitate a far more fine-grained analysis, but I hope that this sketch has shown the importance of a dynamic, flexible view of language. There are conceptual external network links between schemas such as Table 7.1 summarizes. Chapter 13 will provide further discussion of such network links. It is also important to look at the internal networks within a schema, and to that end I next look at the histories of *all the same* (7.3) and *instead* (7.4).

7.3 *All the same*

In PDE, in one use, *all the same* asserts that every item is equivalent. Here *all* is an adjective that specifies quantification over individuals, as in (9):

(9) His tall body had been shoved into **all the same** distortions as mine.
(2010 Osborn, *Distortions* [COHA])

All the same distortions here can be understood as 'each distortion was the same'. There is also a possible reading 'exactly the same distortions', where *all* is not the quantifier but an adverbial (see Chapter 6.2 on use of *all* meaning 'exactly' in *also*).

In another use, *all the same* is a Connector expressing concession ('despite that, even so'):

(10) a. So long as he stayed in that lonesome river, he would be all right. **All the same**, he remained watchful,
(1990 Matthiessen, *Killing Mr. Watson* [COCA]; Traugott In press: Lecture 7B, Example (1))

b. Chechen rebels say they've turned each of the nine-story apartments around it into fortresses. And one good sniper can halt the advance of an entire company. **All the same**, Russian generals continue to argue the campaign is going their way. (2000 *CNN_WorldNews* [COCA])

In (10) *all the same* can be substituted by *even so, in spite of that, nevertheless.* Like them, *all the same* is a monofunctional 1DSM and is not used multifunctionally. As a Connector *all the same* implicates concessive 'although one might have expected otherwise'. Specifically, (10a) is understood as 'although he thought he would be all right, he was watchful', and (10b) as 'although Chechen rebels have …, Russian generals continue to…' Syntactically, *all the same* is a Conjunct marker like *but*, and not a subordination marker like *although*. Unlike in the case of *after all*, concessive interpretations occur in any position. It appears with the same meaning in post-clausal position:

(11) Of course I was blameless, but there might have been a scandal **all the same**.
(1993 Ross, *Cut to the quick* [COCA])

In COCA contrastive pre-clausal *all the same* appears occasionally in the SPOKEN section, as (10b) above attests. However, it is preferred in FICTION and TV, where it also occurs post-clausally, and is equivalent to *even so, despite that, however*, as in (11).

All the same originates in an adjectival phrase [all the same N], where *all* is the quantifier meaning 'every' and the noun it modifies is typically plural, as in (12), although a few instances of singular Ns such a *night* are also attested:

(12) a. [about clothing for a tournament] lordis of the garther with chynes of gold and **all the same** sutes of hertes as is before said.
'lords of the Garter with chains of gold and all the same suits with hearts as was said before'. (1485 *Schort [and] breue tabull on chronicles* [EEBO])

b. those thinges that the Greeke chroniclers haue entreated vpo disorderly …, **all the same** thynges hath pompeius … settyng them in due order compyled in hystorye.
'those things that the Greek chroniclers have discussed in disorderly fashion … all the same things has Pompey compiled in history, setting them in proper order'.
(1564 Golding, *Thabridgment of the histories of Trogus Pompeius* [EEBO])

In (12a) we can paraphrase *all the same sutes of hertes* as 'every suit of hearts was the same'. The plural nouns *sutes* and *thynges* support the quantifier reading, but an adverbial 'exactly the same' reading is not ruled out. Although the N is typically unmodified, the head of *all the same* is an NP not an N, as attested by examples like *hence it is that all the same glrorious and sprituall effects in the scripture, are sometimes attributed to the father* (1656) and *are constituted of all the same substantial parts* (1682).

All the same without a nominal head is not attested in EEBO. OED (*same* C. Adv. 2) cites the contrastive meaning 'in spite of what has been mentioned, even if circumstances had been otherwise, nevertheless, notwithstanding' first in 1803. In that example from 1803 it is in medial position (*could all the same wait on she* [sic] *too*). The first example in OED of Connector use in pre-clausal position is dated 1845 and preceded by *but*. A pre-clausal example appears earlier in COHA in 1823 (13a), where the contrast is with "Wordsworth's finer moods were just those of which he never attempted to give a philosophic account", a long-distance D1. *All the same* is attested with more clearly contrastive meaning shortly after in a translation from French (13b). Pragmatically, *all the same* cues asymmetry between D1 and D2 and implicates 'not exactly the same':

(13) a. It seems to me that Wordsworth's finer moods were just those of which he never attempted to give a philosophic account, and that he did not refer to childhood in his Preface is an evidence of his inspiration when dealing with it. **All the same**, his treatment of childhood accords with his manifesto to the British public.
(1823 *Everett's new ideas on population* [COHA]; Traugott In Press: Lecture 7B, Example (6))

b. The wolf was incapable of an abuse of confidence, and behaved in society, that is to say among men, with the discretion of a poodle. **All the same**, if bad-tempered officials had to be dealt with, difficulties might have arisen.
(1833 Hugo, *By order of the king* [COHA])

To ascertain how *all the same* came to be used as a Conjunct with contrastive meaning in the data gap between 1700 (the end of EEBO) and 1810 (the beginning of COHA, version 2019), I searched CLMET3_0_1 (88 texts representing the years 1710–1780), and CLMET3_0_2 (99 texts representing the period 1780–1850). 9 of the 88 texts of CLMET3_0_1 (10%) provided a total of 21 hits of *all the same*, one of which was excluded on grounds that the string does not form a constituent, leaving 20 examples. None of these can be interpreted as *but*. 17 modify a plural noun, e.g. *all the same relations/objects/ variations*. Here *all* can be understood as 'every', e.g. *every relation the same*. In 3 cases the N is singular (*all the same respect/ company/ appearance*). Here the quantifier 'every' reading of *all* is excluded. Instead, *all* is understood as 'completely/entirely', the scalar adverbial use of *all* that can be found in Old English, and which is a possible reading in some of the early examples in EEBO, as mentioned in connection with (12) above.

In the CLMET3_0_1 corpus 2 examples of *all the same* occur clause-finally after predicate *be*. One with the singular antecedent *it* (14) implicates anaphoric reference to the preceding clause (*if they had been under lock and key*) and evokes an 'entirely/exactly the same' reading for *all the same*:

(14) [lamenting the disappearance of her workbasket] they were all in my workbasket, that I left upon the table in the sarvants-hall, when mistresses bell rung; but if they had been under lock and kay, **'t**would have been **all the same**; for there are double keys to all the locks in Bath.
(1771 Smollett, *Expedition of Humphrey Clinker* [CLMET3_0_1_50])

There is also one instance of initial use; this too evokes an 'entirely/exactly the same' reading, but is comparative (*all the same as if*) and not an instance of Conjunct *all the same*.

(15) "The doctor," said he, "is so much overshadowed with presumption and self-conceit, that his merit has no relief. It does not rise. There is no keeping in the picture, my dear sir. **All the same** as if I were to represent the moon under a cloud; there will be nothing but a deep mass of shade.
(1751 Smollett, *Adventures of Peregrine Pickle* [CLMET3_0_1_48])

In CLMET3_0_2 there are 56 hits of *all the same* compared to 20 in CLMET3_0_1. This suggests that the period 1780–1850 was the one in which SP/Ws may have started to use the phrase in a significantly different way. This is in fact true, but pragmatic uses that by hypothesis enabled PDE usage are limited to a small set of authors. 13 of the 56 examples (23%) are attested in one excerpt, the letters of Robert Browning and Elizabeth Barrett Browning (#185). This means that the change was less wide-spread than the raw numbers might imply. Nevertheless, the

Brownings, especially Elizabeth, who was nominated as poet laureate, were highly influential and their usage may well have been imitated.[41]

Another 13 of the 56 hits are followed by a noun (7 singular, 6 plural), a far lower percentage (23%) than in CLMET3_0_1 texts, where 85% are followed by a noun. In CLMET3_0_1 there is only one Example (12) of the formula *it BE all the same* (*to X*), a predicate expression used clause-finally. But in CLMET3_0_2 *it is all the same (to X)* is attested 19 times, e.g. (16). In examples with singular nouns or *it* antecedents, *all* cannot be understood as the quantifier. It is understood as the degree modifier 'entirely'.

(16) a. 'That's right, Townsend,' said Archer, 'laugh on, my boy! Friend or foe, **it's all the same to you**. I know how to value your friendship now. You are a mighty good fellow when the sun shines; but let a storm come, and how you slink away!' (1796–1801 Edgeworth, *The parent's assistant* [CLMET3_0_2_117]; Traugott In press: Lecture 7B, Example (9))

b. Let them show ever so little inclination, and men go down on their knees at once: old or ugly, it is **all the same**.
(1843 Thackeray, *Vanity Fair* [CLMET3_0_2_174])

The wider context of all but one example of (*It is) all the same (to X)* implicates that the center of consciousness is dismissive of the uniformity described. Indeed, Lenker (2010: 190) notes that "most of the new coinages" of contrastive adverbial Connectors in (Early) Modern English are dismissive, e.g. *at least, after all, anyhow, anyway, in any case, all the same*. 'Being all the same', where *all* means 'entirely, exactly', is largely negatively evaluated, as in (17). It is associated with the idea that 'all the same thing' is dull:

(17) We go the same dull round for ever; nothing new, no variety! **all the same** thing over again! (1782 Burney, *Cecilia* [CLMET3_0_2_96]; Traugott In press: Lecture 7B, Example (10))

When predicated of people it can be used to implicate lack of character or judgement, as in (16a) above.

In addition to the 19 examples of *it is all the same (to X)*, there are 24 examples of *all the same* with concessive or contrastive implicatures. Of these 15 are clause-final uses (18a), 4 are non-final uses (18b), and 5 are initial uses (18c). In each case the context is marked as contrastive (*but* in (18a), (c), *not the less* in (18b)):

41. See e.g. Pratt and Denison (2000) and Fitzmaurice (2010) on the influence of small groups of writers in the transition from the 18th–19th centuries.

(18) a. [complaining about having to sit and chat with neighbors and family] Well, such things must be, and our friends mean them as civility, and we must take and give the currency of the country. But I am _ diddled_ out of a day **all the same**. (1826 Scott, *Journal* [CLMET3_0_2_123])

b. I have no wish to excite your pity, gentlemen, or to gain your silence, by practising upon your feelings. Be silent. I am not the less ruined, not the less disgraced, not the less utterly undone. Be silent; my honour, **all the same**, in four-and-twenty hours, has gone for ever. I have no motive, then, to deceive you. (1826 Disraeli, *Vivian Grey* [CLMET3_0_2_165])

c. Too good, too, too indulgent you are, my own Ba, to 'acts' first or last; but **all the same**, I am glad and encouraged.[42] (1845 R. Browning, *Letter* [CLMET3_0_2_185])

(18a) is best interpreted as 'Although such things must be, …, nevertheless I am cheated out of a day', and (18b) as 'Even though you may be silent, nevertheless my honor has gone for ever'.[43] In (18c) *all the same* appears to strengthen the *but* ('but despite this').

The Brownings' 13 examples are notable in that only 1 instance of *all the same* is used to modify an NP, whereas NP contexts are the norm in earlier periods. The other 12 examples can be interpreted as Connectors signaling a concessive meaning roughly equivalent to 'however, even so, regardless'. 6 are post-clausal (17a), 1 is non-final (17b) and 5 are pre-clausal (19c) below and (18c) above. While (19a) occurs in a contrastive context (*yet*), the other two do not. In (19b) and (19c) *all the same* is being used as a Contrastive independently of contrastive context.

(19) a. It would have been cruel, you think, to reproach me. Perhaps so! yet the kindness and patience of the desisting from reproach, are positive things **all the same**. (1845 E. B. Browning, *Letter* [CLMET3_0_2_185])

b. With which conviction, … of your extravagance of kindness to me unworthy, will it seem characteristically consistent when I pray you not to begin frightening me, **all the same**, with threats of writing less kindly? (1846 R. Browning, *Letter* [CLMET3_0_2_185])

c. if the toad does 'take it into his toad's head to spit at you,' you will not 'drop dead,' I warrant. **All the same**, if one may make a circuit through a flower-bed and see the less of his toad-habits and general ugliness, so much the better. (1846 E. B. Browning, *Letter* [CLMET3_0_2_185])

---

42. Robert Browning called Elizabeth Barrett Browning *Ba* affectionately, an abbreviation she did not entirely appreciate.

43. I thank Kate Beeching for this interpretation.

In (19b) and (19c) *all the same* has been neoanalyzed as a contrastive Conjunct. One later writer may have interpreted *however* and *all the same* as not quite synonymous in clause-final position:

(20) He was game, **however**, **all the same**, Jack.
(1908 Smith, *Peter: A novel of which he is not the hero* [COHA])

(20) is a hapax legomenon, a one-off (there are no other examples in COHA or COCA) and *all the same* may simply have been intended as reinforcement of *however*.

As in the case of many Conjunct DSMs, several factors may have contributed to the development of adverbial concessive use out of adjectival *all the same*, some local, some more general. Likely local enabling factors include:

1. the potential for an adverbial 'exactly' reading in *all the same N-PL* contexts such as (12),
2. loss of the quantifier meaning of *all* in the contexts of *it* anaphoric reference as in (14) and (16) and of a singular noun as in (17),
3. the development of the abstract formula *it BE all the same (to X)* as in (16),
4. negative attitudes toward uniformity as in (17),
5. use in contrastive contexts (18),
6. use in initial position in the context of *as if* (15) and (21) below.

Anything that is 'as if' it is the same as something else is by definition not literally exactly the same, as the examples in (21) show. In both examples the context is negative: *forget … cannot read* (21a), *but … finds nothing* (21b)):

(21) a. 'give her your aunt Grafton's prayer book, and she will read as glib as a minister.' 'Jennet, said I, 'you are mad outright – you seem to forget that Nelema can not read any thing'. '**It is all the same as if** she could,' persisted Jennet. (1827 Sedgwick, *Hope Leslie*, Vol. 1 [COHA])

b. While he muses on the subject which occupies his thoughts and those of the community, the fire [of devout feeling] burns; but when he goes into the house of God he finds nothing to meet the peculiar state of things. **It is all the same as if** nothing had happened.
(1843 *The liturgy of the Protestant* [COHA])

In addition to the assembly of six local contextual factors listed above that may have contributed to the change to Conjunct status, two more general factors may also have played a role. As mentioned in connection with the development of concessive post-clausal *after all* in Chapter 4.5.2, Lenker (2010) identifies the late 17thC and especially the 18thC as a period in which post-clausal use of adverbial contrastives is attested, mostly in interactive texts such as dramas and letters. An example she gives is:

(22) I hope I han't ('haven't'] kill'd the Fool, **however.**
(1696 Vanbrugh, *The Relapse* [Lenker 2010: 196])

Another example is use of *after all* in final position around 1700 (see Chapter 4.5, Example (21a)). So the period when contrastive *all the same* was coming into being was a period in which post-clausal position had come to be the locus for concessive use of Contrastives. Half of the Brownings' use of the new concessive *all the same* appear in post-clausal position. And other authors in CLMET 3_0_2 also appear to prefer it in post-causal position, e.g.:

(23) they are spoiled children, and fancy the world was made for their accommodation; and though I humour both, I think a smart chastisement might improve them **all the same**. (1847 Brontë, *Wuthering Heights* [CLMET3_2_179])

Haselow (2013: 412) argues that in contemporary English post-clausal position is often associated with markers of concession or correction. They are relational and connect D2 to D1. They also express SP/W subjectivity. They may require AD/R to reprocess or even reinterpret the preceding assertion (Lenker 2010: 198). This is not a fully systemic positional context, however, as only contrastive markers used in post-clausal position are concessive, not other markers such as question tags (see Chapter 12.2.2).

Another likely factor in the development of the contrastive meaning of *all the same* is a tendency in semantic change for an expression of similarity or sameness to be interpreted as implicating dissimilarity in semantically negative contexts. In Traugott (1982) I drew attention to the way in which *þa hwile þe* 'during the time that' was reinterpreted as concessive *while* meaning 'although'. Here too, sameness (in this case of time) was understood in certain contexts as implicating non-sameness. A similar development is discussed in Beeching (2005) for French *quand même* 'all the same'.[44] Likewise *even so*, which can be used to paraphrase *all the same* in contexts such as (19) above (e.g. 'are positive things even so' in (19a)), originally meant *even* 'equally' + *so* 'that way'. The phrase *even so* was used from Old English times to express similarity and agreement. It meant 'likewise' and when used as a response meant 'Yes, exactly so'. According to the OED *even* began to be used in the mid 16thC in the contexts of conditional *if* and contrastive *though*. *Even so* appears in EEBO in the 17thC collocated with contrastive *yet*, as in (24):

44. Whether or not this *E* influenced the development of *all the same* deserves attention.

(24) [regarding a king's absolute power over the army] 't is true in case of Forraigne invasion, 't is expedient that the king be farre trusted, and **yet even so**, if the king should conspire with forraigne forces, … we might resume the common native posse.
'it is true that in the case of foreign invasion it is expedient to trust the king in many things; nevertheless, if the king were to conspire with foreign forces … we might call up the native posse'.
(1642 *Accommodation cordially desired* [EEBO])

In COHA we find *even so* in both the 'likewise' and 'nevertheless' meanings. The first is primarily used in comparative contexts (*as … as*), the second in negative, contrastive contexts (see *though … though … yet* in (25)):

(25) [regarding secession leading to the US Civil War] **Though** the thing was pre-determined, and most of the States committed by their public authorities before the people were called on to vote; **though** in taking the votes terrorism in many places reigned triumphant; **yet even so**, in several of the States, secession was carried only by narrow majorities.
(1862 Mill, *The contest in America* [COHA])

By the 21stC *even so* is no longer used in the 'likewise' sense (OED says this meaning is "archaic"). *Even so* can now be used in the meaning 'nevertheless, all the same' without a negative or concessive context. In other words, it has come to be used in ways somewhat similar to *all the same*:

(26) I've discovered energy, like blood, is a renewable resource. **Even so**, after my family is packed off to school and work for the day, I give first dibs on that resource to my writing. (2012 *Losing sanity* [COCA BLOG])

To summarize the histories of *but* and *all the same*, changes to two *E*s with very different sources, one spatial 'on the outside of X', the other adjectival 'all the same (N)', led to rather similar discourse-functional outcomes. Both came to be used to mark a contrastive relationship between D1 and D2, and in some contexts one can be substituted for the other.

As noted in Chapter 3.2.3, morphosyntactic change is typically gradual. Small-step changes in contexts and frequency over decades have been illustrated with the developments of *after all*, *also*, and *but*. In the case of *all the same*, the shift appears to have been preceded by small step changes such as increase in frequency of *all the same as this* and of semantically negative contexts. Potential for interpretation *all* as 'exactly' can be seen with hindsight from the 15thC on (12). However, as the contexts are always plural, we do not have evidence of Conjunct use emerging over a long period of time. Contexts in which the quantitative meaning 'every' is excluded and only the adverbial 'exactly' meaning is plausible emerged somewhat

rapidly and appear to have been strongly associated with a literary couple, the Brownings. The Brownings did not, however, innovate the new uses. Concessive and contrastive implicatures can be found in earlier writings such as those of Scott and Disraeli (Examples (18a) and (18b) respectively). The Brownings appear to have adopted such uses and other writers appear to have followed suit. Even though contrastive *all the same* was developed over a relatively short period of time, it has all the hallmarks of gradual constructionalization: constructional assemblies and conventionalized local contextual uses.

A partial model of the development of *all the same* appears in Figure 7.2. Key here is the constructionalization of *all* as an adverbial degree modifier 'exactly' in contexts of singular nouns and pronouns, followed by a second constructionalization to Conjunct. In Figure 7.2 "quant same" is short for a 'quantifier + the + same', "adv same" for 'adverbial + the + same', "equiv" is short for 'equivalent', and "ref" for 'reference', "deg mod" for 'degree modifier'. The "formula" referred to in the second set of contexts is *(It is) all the same (to X)*.

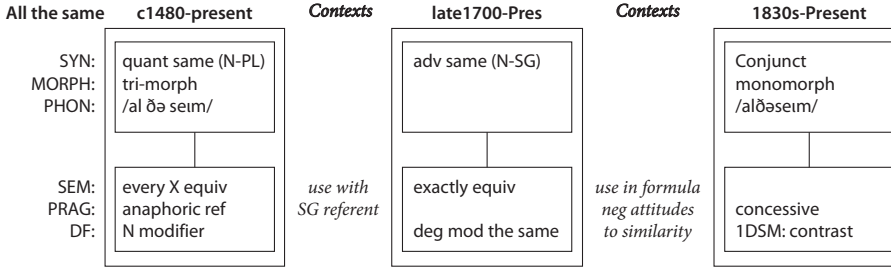

Figure 7.2 Partial model of the development of DSM *all the same*

7.4 *Instead*

Like *but* and *further(more), instead* is spatial in origin ('in place of'). It was used in the substitutive sense of 'X instead of/in place of Y' by late Old English. Although scattered examples of *in the stede of* occur, the article is for the most part not used (see also *in place of*). *Instead of* was used as a prepositional chunk from the mid 17thC on. At first restricted to substitution of one person for another (e.g. as *deputy*), it was extended to concrete inanimate objects (*turpentine instead of balsam*) (Schwenter and Traugott 1995).

Bare adverbial *instead* used as a topicalized Conjunct ('in place of that') is occasionally found in EEBO following *but* and specifying that a sharp type of contrast is intended:

(27) a. whenever the almighty intends any great event he puts it more or less in the people's minds before the execution, to induce them to holiness; for there is no other way to prevent his wrath: but **instead**, they continue to act as if they were never to die.
(1651 Love, *Strange and wonderful predictions* [EEBO])

b. and therefore had reason to expect some proof of it, but **instead**, he tells us how they look on themselves as obliged to …
(1686 Stillingfleet, *Doctrines and practices* [EEBO])

Bare *instead* meaning 'however' is not attested with any great frequency until the second half of the 19thC, as evidenced by COHA, e.g. (29).[45] The contrast is always sharp, and the source meaning 'in place of X' persists, as in (28):

(28) a. He was not permitted to know that she felt the keen humiliation, which a proud nature must suffer when it discovers that it has trusted an unworthy object. **Instead**, he was to feel himself the injured one.
(1884 Lynch, *Madeline Payne* [COHA])

b. Werner von Haeften, a collaborator, had felt too rushed to put a second bomb in the briefcase. Had they done so, Hitler would have certainly been killed. **Instead**, when the smoke cleared Hitler was still standing.
(2001 Beschloss, *The conquerors* [COHA])

As we have seen, *instead* was originally used as the complex spatial preposition *instead of*. Bare *instead* was used as a topicalized CircAdv as in (27). There is no need to think in terms of reduction from *instead of this* and invoke grammaticalization, as Lenker (2010: 177) does. For one, *instead of this* does not appear with any great frequency in the data used as an adverbial contrasting D1 and D2. More importantly, there is a general pattern in English of preposition > adverbial. In PDE there are several preposition-adverb pairs, such as *above, behind, within*, some of which have of NP complements, e.g. *off, outside* (e.g. *outside of the store, went outside*). In terms of meaning, in using *instead* as a Conjunct, SP/Ws in the ModE period appear to have construed texts abstractly and metaphorically as objects with parts that can be substituted one for another.

The development of DSM *instead* can be partially modeled as in Figure 7.3:

45. Lenker (2010: 98) finds that in her data *instead* used as a Connector is an even later PDE phenomenon. In her work, PDE refers to the period 1920–1990, whereas in this book PDE refers to the period 1970 – present.

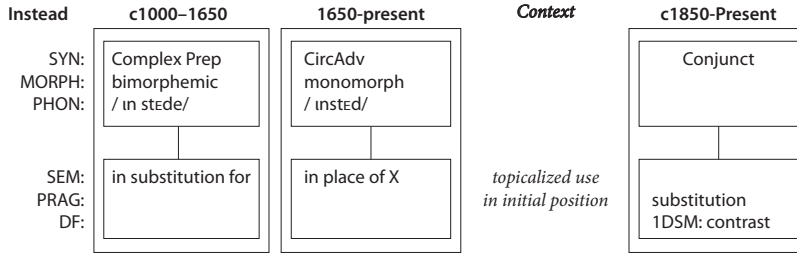

Figure 7.3 Partial model of the development of DSM *instead*

7.5 Conclusion

Of the three expressions discussed in this chapter: *but, instead*, and *all the same*, the first two were initially used in spatial expressions. *All the same*, however, originates in an adjective modifier expression, typically of NPs. We may conclude that the source constructional space of Contrast is less homogenous than that of Elaboration (see Chapter 6).

Given that not all Conjuncts in English derive from CircAdvs, the model of the DSM Trajectory Hypothesis in Figure 4.4 in Chapter 4.4.2 needs to be slightly modified. Although a large preponderance of [[Conjunct] ↔ [DSM]]s in English are derived from adverbials, "CircAdv" is too restrictive to be appropriate for a general model accounting for the rise of DSMs like *all the same*. "Adjectival phrase" (AP) may be rather too specific but is adequate to the data under discussion. Figure 7.4 represents the revised template with CircAdv and AP specified as alternative sources of [[Conjunct] ↔ [DSM]] constructions.

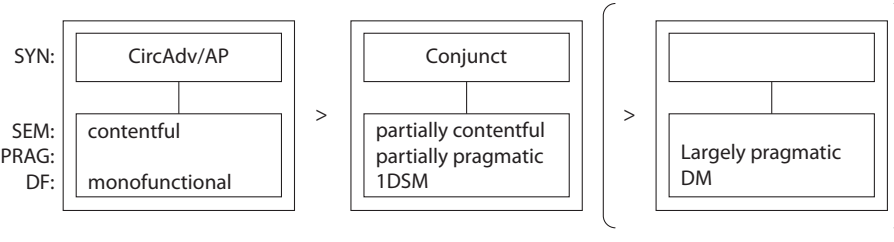

Figure 7.4 Revised Discourse Structuring Marker Trajectory Hypothesis

Given this revised hypothesis, all three contrastive markers discussed in this Chapter can be said to support the first part of the hypothesis, that contentful phrases can be reinterpreted as having the structure [[Conjunct] ↔ [1DSM]]. Like *further*, *furthermore* and *moreover*, discussed in Chapter 6, *all the same* and *instead*

did not undergo expansion of their pragmatic functions and reshaping of their discourse functions, in other words they have not (yet) been recruited to DM status. However, *but* came to be used as a DM.

The case studies in this chapter have further highlighted the importance of specific discourse contexts in the rise of new uses. The enabling factor of use of a source expression in clause initial position was illustrated by the use of exclusive *but* before conditional *if that* in Example (6), of topicalized *all the same things* in (12b), and of topicalized *instead* after *but* in (27). However, as in the case of DSMs discussed in other chapters, use in clause initial position is presumably necessary but not sufficient for use as a DSM to occur. Examples used in this chapter suggest that a second important context for the rise of DM *but* is use with a negative D1 as in (7); for *all the same* an additional context is use of the formula *it BE all the same (to X)* as in (16), evoking negative assessment of the content of D1; and for *instead* it is use with *but* as in (27).

Elaborators and Contrastives typically involve a topic shift by default. Some Connectors that mark topic shift by default signal digression or return to the prior topic. Digressive topic shift markers are the subject of the next chapter.

CHAPTER 8

The development of markers of "digressive" topic shift

8.1 Introduction[46]

The markers discussed in Chapter 7, *but, all the same*, and *instead* mainly signal that the relationship between D1 and D2 is a contrastive one. By default, D2 involves a topic-shift. There are various degrees of intensity of shift. Of the three contrastives, *but* signals the weakest topic-shift (indeed, Fraser 2009a: 893 lists *but* in one of its uses among topic-continuation markers). *All the same* can be used to reinforce and strengthen *but*, but the content of D2 need not contrast significantly with that of D1. *Instead* is the strongest of the three markers, implicating that D2 is different enough from D1 to warrant using it as a substitute for D1.

The present chapter concerns another type of topic-shift marker, the so-called "digressive" markers *by the way, by the by, incidentally* and *parenthetically* (Mittwoch et al. 2002). "Digression" implies stepping aside from the current topic and therefore always signals a topic shift (Pons and Estellés 2009). In conversation there is often no return to the original topic, but when and if SP/W does return to the original topic, it will be elaborated on and updated in a way that this move too signals a (relatively small) topic shift rather than simple continuation. The digressive markers together with the host clause serve as a parenthetical comment on, often an elaboration of, D1. Drawing on Biber et al. (1999: 137), I define parenthetical comments as clauses that "give additional information which is related to, but not part of the main message".

Four case studies are presented in this Chapter: the rise of *by the way* (8.2), which has been productive from the time that it was constructionalized as a DSM, and is currently widely used as a DM. The markers in the other three case studies: *by the by, incidentally* and *parenthetically* are 1DSMs with low productivity, only one of which, *incidentally*, shows potential signs of DM use (Section 8.3). Some other alleged digression markers are briefly introduced in Section 8.4. Section 8.5 concludes.

---

46. This Chapter draws on and updates Traugott (2020a, 2020b).

8.2 *By the way*

Like *but*, *by the way* originates in a spatial expression. By hypothesis, a prerequisite for eventual use of a spatial expression as a DSM is use in metaphorical conceptualizations of the development of an argument as an object with spatial coordinates. One possibility is to conceptualize the proposition in D2 as outside the space of the argument in D1. This is the kind of static configuration associated with *but*. Alternatively, one can conceptualize discourse as a journey from which the SP/W can take a side-trip. As mentioned in Chapter 6.3.1 in connection with *further*, Lakoff and Johnson (2003[1980]: 89–96) draw attention to metaphorical templates such as AN ARGUMENT IS A JOURNEY. This widely used metaphor probably enabled the Connector development of both *by the way* and *by the by* as well as *further* and the DSMs to be dicsussed in the next chapter on return to a prior topic.

In Old English *be (ðam) wege* may be used in situations that are static (1a) or dynamic (1b):[47]

(1) a. þa sæt þær sum blind man. **be ðam wege.**
then sat there a-certain blind man by that.DAT wayside.DAT.
(c.1000 Ælfric, *ÆCHom* I, 10 B1.1.1 [DOEC])

b. & þam he forgifð þone gastlican fodan. þæt hi ne teorian
and them he gives that spiritual food so-that they not faint
be wege. (c.1000 Ælfric, *ÆCHom I*, 12 B1.1.13 [DOEC];
along way.DAT. Traugott In press: 145)

It is the dynamic CircADv use exemplified in (1b) that underlies the DSM *by the way*. A context that I call 'talking *en route*' appears throughout the textual record from Old English on (2a). By the end of the 15thC the journey came to be replicated with some frequency as the background to an important event (2b) or statement (2c):

(2) a. Hi sylfe eac cyddon þæt hi crist **be wege** gespræcon,
they themselves also revealed that they Christ by way.DAT discussed,
and hu hi on heora gereorde hine oncneowon.
and how they at their meal him recognized.

b. it was told him ***by the way*** that his wyf was deed in trauayll of child.
(c.1000 *ÆCHom II*, 16 B1.2.19 [DOEC])
'it was told him on the road that his wife was dead in work of child (childbirth)'. (1482 Caxton, *Prolicionycion* [EEBO];
Traugott 2020a: 124; In press: 145)

47. The definite article had not been developed by this time and its source, the demonstrative, was optional (see Sommerer 2018).

c. and as they wente **by the way** abram said to his wif i fere & drede sore that … 'as they went along the way Abraham said to his wife "I fear and dread greatly that …"'(1483 Caxton, *Legenda aurea* [EEBO, Traugott 2020b: 5])

In examples like (2) the route is understood as the literal physical background to the conveying of the information that is the focus of the clause. In the 16thC the 'talking *en route*' collocation came to be used more frequently. By the time of Shakespeare (late 16thC, early 17thC), the concept 'make a salient statement while talking *en route*' had become quite common. 7 out of the 14 instances of adverbial *by the way* extracted from the *Open Source Shakespeare* collocate with a verb of saying or understanding, e.g. *recount* 'tell' in (3a) and *tell* in (3b). Note in both examples the use of *by the way* in clause-initial, topic position:

(3) a. Why, then, we are awake: let's follow him
And **by the way** let us recount our dreams.
(c1595–96 Shakespeare, *Midsummer Night's Dream* IV.i.202 [OSS; Traugott 2020a: 124; In press: 146)

b. Go with me to it, and I'll show you; and, **by the way**,
You shall tell me where in the forest you live. (c1599 Shakespeare, *As You Like it* III.ii.451 [OSS; Traugott 2020a: 124])

Particularly important for the development of the DSM use would appear to be the rise in the 16thC of a new figurative use where *way* is understood not literally as a road, but metaphorically as a textual journey and *by the way* means 'in the course of discussion, in passing'. In this use, *by the way* is still a CircAdv. Many examples appear in translations from Latin or French e.g. (4a) and in treatises (e.g. (4b)), sermons, and philosophical works. A verb of speaking is typically used, e.g. *say, touch, make mention*:

(4) a. many others the which at this time i omit: this much i will **say by the way**, that this straight passeth ouer the coast of afrike to the troppike of cancer, 'many others (details) which I omit at this time: I will say this much in passing, that this strait passes over the coast of Africa to the tropic of cancer'.
(1568 Hackett, *The new found vvorlde* [EEBO; Traugott 2020b: 5; In press: 147)

b. benedictus … vnpoped himself, and went to uelitras, liuing there more quietly the he shuld haue done at rome: here is to be **touched by the way**, the glose of the popes decrees futed:
'Benedict … revoked his own authority as pope, and went to Velletri, living there more quietly than he would have done in Rome; here is to be touched on in passing the exposition of the pope's decrees that are rejected'.
(1583 Foxe, *Actes and monuments* [EEBO])

This metaphorical extension was used from around 1600 in another enabling context, non-restrictive relative clauses. Here there is an element of subjectivity, as SP/W takes the stance of casually assessing D2 as worthy of notice:

(5) a. to vse al such meanes as may further their resolution, assuring themselues (which **by the way** is worth our obseruation) that …
(1616 Hakewell, *An ansvvere to a treatise vvritten by Dr. Carier* [EEBO])
b. had the unerring spirit of god, to direct them in all emergent occasions (which, **by the way**, renders their frequent grumblings no lesse prodigious, then blasphemous). (1656 Osborne, *Politicall reflections* [EEBO])

In (5) we understand *by the way* to mean 'in passing', as in (4). The appositive relative clause is not syntactically integrated with the host clause, and position after the relativizer puts *by the way* structurally close to the beginning of the clause.

Some early examples of Connector use of *by the way* in clause-initial position appear in the mid-1600s and are exemplified in (6), where *by the way* means 'casually in passing (in this text)'. Most are attested in translations and relatively formal genres, as in (6b). (6b) comments on a city in Wales called *Llan-baderne the Great*. Once a papal see, and an administrative center of the Catholic church, it had dwindled into a small village:

(6) a. sent him a plaine and cold answer, wash seven times in iordan: **by the way**, I dare boldly say, elisha in himselfe was not proud at other times.
(1640 Fuller, *Ioseph's partie-colored coate* [EEBO])
b. which city is now dwindled to nothing: reader, **by the way**, I observe that cities surnamed the great, come to little at last.
(1662 Thomas, *The history of the worthies of England* [EEBO];
Traugott In press: Lecture 8A, Example (13))

Here *elisha in himself* and *I observe that cities surnamed the great, come to little at last* are introduced by *by the way*, which implicates that background and somewhat unimportant information is to come. It invites the reader to think of D2 as a footnote on what precedes. However, in both examples, D2 is a generalization that is presumably meant as a contribution of some significance. In (6b) and in many other examples, the casualness and backgroundedness can be interpreted as somewhat false.

In (6a, b) a cluster of factors suggests [[Conjunct] ↔ [DSM]] use:

1. clause-initial position
2. there is no overt mention of a path, either literal or textual,
3. subjectivity: SP/W injects himself in first person and indexes the upcoming discourse as:
   i. a new discourse-topic (i.e. a topic shift)
   ii. only partially relevant

At around the same time, some examples of *by the way* used in clause-initial position appear to be ambiguous as in (7):

(7) *CAR.* I'le never be drunk agen.
LA. I hope you will say so, when you have heard all.
GEO. but ***by the way*** your late stock being spent, here are ten peeces towards a supply. (1653 Brome, *The mad couple well matched* [CED: D3CBROME; Traugott 2020a: 125])

Since Careless (CAR.) forfeited his stock of £500 on a drunken travel adventure described earlier, *by the way* in (7) could be understood as in earlier texts to mean 'on the way', in which case it refers anaphorically to the prior account of the journey and has local scope over *your late stock being spent*. On this interpretation it is a topicalized CircAdv in initial position. But from a present-day perspective it could also be interpreted as 'incidentally', in which case it is a DSM that refers forward cataphorically to D2. On this interpretation, *by the way* is a metatextual Conjunct in pre-clausal position. It has scope over all that follows (*your late stock being spent, here are ten peeces towards a supply*). It is relatively subjective in that it expresses the speaker's stance-to-text, downplaying the importance of the offer.

In many cases, SP/W uses *by the way* meaning 'in passing' to represent themselves as taking the stance that D2 is less important than D1 and is only partially relevant to it when in fact the content of D2 is salient information, as in (6a) and especially (6b). Narrators sometimes comment explicitly on the 'as if', falsely casual, 'suddenly-came-to-mind' meanings associated with *by the way*, as in (8). The background of (8b) is that the narrator has gone to breakfast in a café with the intention of finding out about a TV monitor that has been rumored to be suspicious and turns out to be used to spy on customers.

(8) a. After bidding good morning by way of farewell, he walked to the door, when suddenly turning, as if the thought had just struck him, he observed – "**By the way**, if anybody should happen to notice that I had called on you, I have no objections to your saying I had a talk with you about that case of Fanning's." (1856 Adams, *The lost hunter* [COHA])

b. When I paid for my breakfast I half turned away, then turned back casually. "**Oh, by the way**," I said. "Where's this wall TV place?" "This what?" she said. "You know," I said. "Color TV like a picture you hang on a wall." All the color faded from her face. Her eyes went past me, staring. I turned in the direction she was staring, and on the wall above the plateglass front of the cafe was a picture.
(1959 Graham, *The Gallery* [COHA; Traugott 2020b: 9]

Over time, association of *by the way* with partial relevance and casual delivery of D2 enabled interpretation of the DSM as an index of the interpersonal hedging

function. In using the term "hedge" here, I draw on the concept of a hedge as a discourse strategy that "attenuates … the full force of the speech act" (Fraser 2010: 16; see also Caffi 2013: 257 on mitigation as an "attenuation strategy rooted in metapragmatic awareness"). An elaborate narrative description of the how the potential to use hedging *by the way* is conceptualized is given in (9):

(9) "I might try to find something for you in my business." "It's very nice to hear, Sir," Jeff said, embarrassedly. "I'll remember it." "**By the way**," Oliver said, as though it would have been impolite to get to the main question too soon and he was casting about, looking for subjects of conversation, "that girl of yours you talked about the day I met you…" (1956 Shaw, *Lucy Brown* [COHA])

Use of *by the way* as a mitigating hedge on up-coming text that might be thought to be face-threatening is a case of intersubjectification. I will return to this topic in Chapter 11.

(10a) illustrates use of *by the way* to introduce an overt directive to do something, in this case take piano lessons that the addressee is not interested in. In (10b) it is used to shift topic and as if casually introduce a question (*where does this fellow Umholtz hang out?*), followed by an indirect directive to investigate Umholtz. In both cases, *by the way* attenuates the illocutionary force of D2:

(10) a. Verry still played. "Her talent is wonderful," said father, taking the cigar from his mouth. "**By the way**, you must take lessons in Milford; I wish you would learn to sing." I acquiesced, but I had no wish to learn to play.
(1862 Stoddard, *The Morgesons* [COHA])

b. Think you can get any of this stuff back?" "I hope so. **By the way**, where does this fellow Umholtz, the fabricator of spurious Whitneyville Walker Colts, hang out? I believe he ought to be looked into."
(1953 Piper, *Murder in the gunroom* [COHA])

This kind of hedging use is exploited in dramas at the turn of the 20thC, as in (11).

(11) a. CECIL GRAHAM. Hallo, Tuppy! Hear you're going to be married again; thought you were tired of that game.
LORD AUGUSTUS. You're excessively trivial, my dear boy, excessively trivial!
CECIL GRAHAM. **By the way**, Tuppy, which is it? Have you been twice married and once divorced, or twice divorced and once married? I say you've been twice divorced and once married. It seems so much more probable.
LORD AUGUSTUS. I have a very bad memory. I really don't remember which. [[Moves away R.]
(1892 Wilde, *Lady Windemere's Fan* [CLMET3_0_260])

b. RICHARD: So I hear you are married, Pastor, and that your wife has a most ungodly allowance of good looks. … [discussion of the propriety of saying this in front of her]… All the same, Pastor, I respect you more than I did before. **By the way**, did I hear, or did I not, that our late lamented Uncle Peter, though unmarried, was a father?
UNCLE TITUS: He had only one irregular child, sir.
(1897 Shaw, *The devil's disciple* [CLMET3_1_3_269; Traugott 2020a: 127; In press: 150)

Cecil Graham in (11a) is a troublemaker. "Tuppy" is Lord Augustus, who is clearly put out by Graham's topic of conversation and particularly by the question introduced by *by the way* (he gives an evasive answer and moves away). (11b) is striking for the multilayered, elaborate hedge. First there is explicit alleging of respect in *I respect you more than I did before*, then *by the way*, followed by fake prevarication *did I hear, or did I not*? I

The false hedging use of *by the way* is by hypothesis the source of the fairly recent development of an aggressive use of *by the way*, especially in combination with *Oh*. In examples like (12), *by the way* and often by implication the quoted their person user of it, is mocked and disapproved of.

(12) Now, what does it say to the entire world that you have a president who gets off the phone with Erdogan of Turkey, and then sends out a tweet and it's **oh, by the way**, we're deserting these people who have put their lives on the line to work with us in fighting against some of the worst terrorists in the world.
(2019 *ABC News* [COCA])

This use will be discussed in Chapter 10.4 on combinations of DMs.

Another recent use of *by the way* that was noted in Schegloff and Sacks (1973: 319) is what they call a "misplacement marker", as when one introduces oneself or someone else at an abnormally, embarrassingly late point in a conversation (*by the way, this is/I am X*). It appears in texts in COHA around the 2000s.[48] This particular use of *by the way* is preferred post-clausally in fiction, as illustrated in (13):

(13) My neighbor remained undaunted by my silence. "I'm Jane, **by the way**, Jane Mansfield." She laughed, showing off her blinding teeth again.
(2006 Sutherland, *The Red Hat Society's queens of Woodland Avenue* [COHA])

48. The thirty year time lag between Schegloff and Sack's observations about this use and its representation in texts may be due to absence of conversational interaction in the corpora used.

However, it is infrequent. In later decades, *Oh, by the way my name is X* is preferred, with *Oh*. In COCA NEWS shows, when a new participant is introduced, *by the way* has a different function. Here it does not implicate acknowledgement that the introduction is abnormally timed. Rather, it signals a topic shift associated with turning to a new participant in the show. In this case it is used pre-clausally, as in (14):

(14) We figured out the best way to take a beach photo. Okay. **By the way**, this is Willy and Juan. (2015 *NBC_Today Show* [COCA])

To summarize this section, the main change to the micro-construction *by the way* discussed above is:

(15) [[by the way]$_{\text{CircAdv}}$ ↔ [‘along the route’]] >
[[by the way]$_{\text{Conjunct}}$ ↔ [‘marks D2 as a digression’]]

This is a constructionalization as both form and meaning have changed. It occurred around 1575. The most significant form change is CircAdv > Conjunct change. There is also morphological change from the trimorphemic CircAdv phrase to a chunked single monomorpheme. A modified phrase like *by the long way* as in *by the long way of their descent* (1603), does not function as a DM, but as a CircAdv.

The chief contexts for the constructionalization are:

1. representation of talking *en route*; association of talk with the route by hypothesis enabled association of the path with the metaphor of text as journey, and development of the ‘in passing’ meaning,
2. representation of the talk as the main point of the contribution in D2; this contributed to interpretation of the route as relatively unimportant.

In addition, two constructional changes were discussed:

a. c1650 use as a DM signaling that D2 presents a casual topic-shift; this change was by hypothesis strengthened by association with non-restrictive relative clauses which tend to be presented as casual and relatively unimportant add-ons,
b. c1850 *by the way* came to be used to pragmatically invoke mitigation of face-threatening content in the upcoming D2.

Figure 8.1 is a partial constructional model of the changes (“morph” is short for ‘morphemic’, “b-g” for ‘background’, “nrRel” for ‘non-restrictive relative’, “repr” for ‘represented’, “digress” for ‘digressive’):

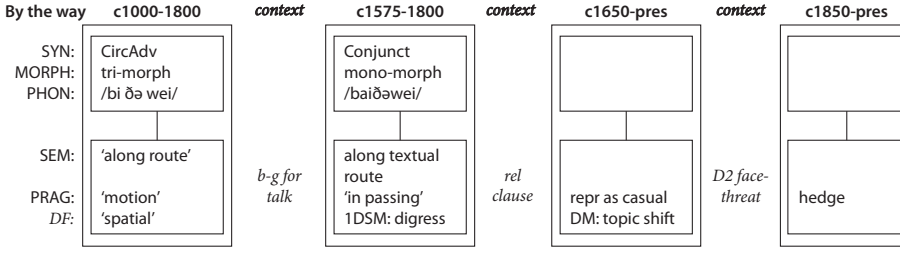

Figure 8.1 Partial model of the development of DM *by the way*

8.3 Three relatively unproductive markers of digression

As mentioned in Section 8.1, Mittwoch et al. (2002) list by *the way, by the by, incidentally* and *parenthetically* as markers of digression. I outline the development of *by the by* in 8.3.1, which appears to have some DM properties and of *incidentally* and *parenthetically* in 8.3.2. *Incidentally* is of considerable interest as it may be coming to be used as a DM, unlike *parenthetically*.

8.3.1 *By the by*

A somewhat infrequent expression *by the by(e)* was used in Middle English as a prepositional phrase (not to be confused with the temporal adverbial *by and by* 'in a while'). It is an example of a DSM that has unclear origins and has never been productive. MED points to the first *by* as a preposition meaning, as in Modern English, 'near, alongside, along' and the second *by* as a noun meaning 'village'. No examples of *by the by* in a spatial sense are provided. OED (*by* n^2) cites only "figurative" examples of *by the by* used with reference to text organization, along with *by the way*.

The phrase is used in EEBO as a textual marker in the 17thC around the same time as *by the way* came to be used as a DM, but it has never been token frequent like *by the way*. EEBO provides 1,126 examples in both spellings combined (compare 22,891 *by the way(e)*), COHA 268. In the 1 billion word 2020 release of COCA there are only 184 examples. This suggests that it has been obsolescing from the 18thC on.

In some scattered earlier entries in EEBO from the 16thC and 17thC *by the by* appears to be used as a temporal and to mean 'later', as in *by and by*. The first example in EEBO of use of *by the by* in a context referring to a linguistic activity (*testimonies produced*) (but not as a Connector) is:

(16) here is no mention of fathers, no testimonies produced from them, except three or foure at the most, which are brought in **by the by** at the latter end of the booke:
(1624 Casaubon, *The vindication or defence of Isaac Casaubon* [EEBO])

Here *by the by* means 'in passing', like *by the way* in (6) above. It is presented as one of several paraphrases in both (17a) and (17b). In (17a) *per accidens* is Latin for 'accidentally', and *concomitanter* for 'concomitantly, at the same time'. In (17b) *obiter* is Latin for 'incidentally' and *casu* for 'casually':

(17) a. when knowledge is but a circumstance in our sinnes … knowledge is indeed sinned against, yet but collaterally, and as a stander by, but as a circumstance onely, shot at per accidens, concomitanter, and **by the by**, as one that steps in to part a fray is smitten,
(1637 Goodwin, *Aggravation of sinne* [EEBO; Traugott 2020a: 129])
b. god should not be worshipped, obiter, or casu, at it were, in passing by, or, **by the by**, but to have the whole mind, intent upon the service,
(1670 Annand, *Pater noster* [EEBO]

By the by is found with *by the way* in:

(18) you may see then, how this mighty truth is hid under a mystery, and is couched in two or three words: how moses doth but (as it were) **by the way**, glance at it **by the by**; and if god open your eyes, you may see how this secret warp and woof runs through this web. (1657 Barker, *Gospel treasury opened* [EEBO])

In (18) *but* means 'except' and *but by the way* is probably to be understood as 'except in passing' and *by the by* more specifically as 'without full attention' (see (17b)). *By the by* is not used as a Connector in any of these examples.

The 1670s is the first decade in which there are over 100 examples of *by the by* in EEBO. Most are clause-final and not used as Connectors. The first pre-clausal Connector use in my data is attested from 1671 in a work said in the title to be "written originally in English":

(19) i will not insist upon their customes, which is a subject of which many have treated, however **by the by** let me say in few words, that in their division of the artificial day, they differ from the general practice of europe,
(1671 Gailhard, *The present state of the princes* [EEBO; Traugott 2020a: 129])

It is possible that the use in (19) was supported by the development of pre-clausal uses of *by the way* as in (6) above because the Connector uses of *by the by* are attested at about the same time as those of *by the way*, and in the same contexts (locutionary verbs and relative clauses).

In COHA version 2010 *by the by* is attested in only double digits from the 1810s through the 1890s, and only single digits after that, several in post-clausal position. All examples in pre-clausal position involve a topic-shift that is presented as casual. In the most recent 1 billion word edition of COCA, there are 179 hits of which 27 are post-clausal and 96 are pre-clausal, e.g. (20):

(20) they did pull back from some agreements we thought we had. And, **by the by**, that also includes all manner of enforcement to whatever conditions are made.
(2019 *CBS_News* [COCA])

There are scattered examples in which *by the by* is used like *by the way* as a hedge on a face-threatening D2. In (21) it appears to be used at the onset of a contribution to ratchet up the insults:

(21) This thread is about privilege, and it'd be kind of foolish to claim any just because you're a certain type of person. Whether that's the type of person you professed to be, or what Rebecca called Richard. "**By the by**, your arrogance in this situation is -outstanding- and I want to throw out a question".
(2012 *The privilege delusion* [COBA WEB])

In conclusion, *by the by* appears to have been used as a Conjunct from the 18thC on, perhaps on analogy with *by the way*. It is largely used in a monofunctional way. However, examples like (21) suggest that it is marginally multifunctional. It has been quantitatively marginal from the beginning in clause-initial position and appears to be obsolescing.

8.3.2 *Incidentally* and *parenthetically*

By the way clearly originates in a spatial CircAdv. The origins of *by the by* are less clear but appear to be spatial. The two other digressive markers cited by Mittwoch et al. (2002), *incidentally* and *parenthetically*, are drawn from a different conceptual domain, that of manner adverbs (like *also* in Chapter 6.2). All four expressions are CircAdvs in origin.

Incidentally is consistently used more frequently than *parenthetically* in the corpora, as shown in Table 8.1.

Table 8.1 Raw numbers of hits of *incidentally* and *parenthetically* in the corpora

	incidentally	***parenthetically***
EEBO	76	4
COHA (release 2021)	3,601	170
COCA (release 2020)	4,421	222

Table 8.1 gives raw numbers only; most hits are not used as Connectors. *Incidentally* and *parenthetically* are both used in EEBO exclusively as manner adverbials. Most are attested in the context of text-production and discursive practices, contexts that, as we have seen, enabled the development of *by the way*:

(22) a. many other scores there are, which are scattered here and there, by one or two, as **incidentally** he taketh occasion to write.
(1615 Worthington, *Whyte dyed black* [EEBO])

b. the first words seem to be put in **parenthetically**, and so the words following agree immediately with the 37 verse.
(1662 Lawson, *Exposition of the* Epistle [EEBO])

Among the hits of *incidentally* in COHA release 2021, most pre-clausal instances are Conjuncts with partial lexical and relatively objective properties. However, from the 1870s a few examples appear in which *incidentally* is grounded in the SP/W, suggesting that *incidentally* may be coming to be used as a stance marker, e.g.:

(23) there need not be much fear that any damage of this kind will be done. And, **incidentally**, it may be well to mention here one positive benefit which will result from the study of English literature as pursned in this manner.
(1870 Lounsbury, *Study of English literature* [COHA])

In (24) *incidentally* is used to introduce a new topic: the problem that the washer isn't working just when the group wants to sell it and to explore whether a repairman can be found who does not need to be paid. Misha uses a jokingly face-threatening preamble to this topic, and introduces it with *incidentally*. This, and 46 examples of the combination *Oh, incidentally*, confirm the possibility that *incidentally* may be coming to be used as a DM as was mentioned in Section 8.3 above, but examples are very sparse:

(24) "They talked about refrigerators, clearly." Misha rolled his eyes at Arkady. "**Incidentally**, you don't happen to know any murderer repairmen who owe you a favor?" (1981 Smith, *Gorky Park* [COHA])

Turning now to *parenthetically*, among the 170 examples of *parenthetically* in COHA release 2021, only 2 appear in initial position in the 19thC. In the 20thC use at the beginning of the clause becomes more common, especially in the context of a SP/W's explicit comment on his or her own discursive practice (25). Here, *parenthetically* is used as a Conjunct meaning 'in passing'.

(25) a. spread suddenly out over the entire and hitherto uninhabited central Pacific island area. **Parenthetically**, it becomes necessary here to call attention to the fact that… (1938 Boas, *General Anthropology* [COHA])

b. I just typed way too much so I'll stop here. **Parenthetically**, I saw a short interview you did. (2012 *Blog maverick* [COCA BLOG])

In COCA initial and pre-clausal uses of *incidentally* and *parenthetically* are attested mainly in the BLOG section, where they signal topic shifts. *Incidentally* has some of the characteristics of DMs: pre-clausal position independent of discursive verb contexts, subjective view-point, fairly extensive use in the SPOKEN component, and 26 examples of the combination *Oh, incidentally*. In some examples, it is used as a hedge and is therefore multifunctional, but only marginally so. There is no evidence that it has been conventionalized as a DM yet.

8.4 Some other alleged digressives

Fraser (2009a: 895) includes in his list of DMs that signal digression from the current topic a set of clausal formulae: *I almost forgot, I just remembered, that reminds me, to update you.* They occur infrequently compared to *by the way* and even *incidentally*. In the March 2020 release of COCA there are 916 hits of *I almost forgot*, 500 of *I just remembered*, and 968 of *that reminds me*, 130 *to update you.* By comparison, there are 38,959 hits of *by the way* (raw numbers, including manner adverbial use), and 4,426 of *incidentally*.

Of the four formulae mentioned above, only *I almost forgot* appears in EEBO, but without any topic digression or topic-shift function. In COHA release 2021 there are 407 examples. It is attested around 1900 in a use that could be paraphrased as *by the way*, introducing a new topic that is more important than the dismissive introduction implies (26a). In (26b) the veracity of *I almost forgot* is doubtful, given the importance of the information in D2, and it can be interpreted as being used as a hedge in an awkward situation. This use is infrequent, however.

(26) a. I think I'll take a nap. Oh, Baron, **I almost forgot**. I was in Asheville a few days ago, – Monday, Tuesday, – I don't know when," he went on, weakly, "and I met a man who said he thought he knew you".
(1903 Smith, *A tar-heel baron* [COHA])

b. Anyway, what brang you out on a day like this?" "Oh, **I almost forgot.** Wiley Dipp is coming to arrest you." (1993 *Outdoor life* [COHA])

I just remembered appears 257 times in COHA. Most examples in initial position are followed by a *that*-clause or an object (1868, *I just remembered who was with me*). Among the few examples without a following overt *that*-clause is (27). Here, *I just remembered* is a response to a question and does not appear to be used pragmatically as a DSM:

(27) Why do you just stand here and – (To the butler, who stands quietly awaiting recognition) Yes, Charles? CHARLES **I just remembered**, madam. Some bundles came for you this morning. (1935 Dayton, *First lady* [COHA])

Among the 712 hits of *that reminds me* in COHA, most are followed by a PP (*that reminds of somebody*) or a *that*-complementation. However, a few, e.g. (28), appear to be syntactically "outside the clause" and to be used to introduce a new topic in (represented) conversation that the center of consciousness has in mind to pursue.

(28) Mr. Landholm, you never come to see us." "I have so much else to see," he said, glancing at his book. "Yes, and **that reminds me** – Have you heard the news?" (1852 Warner, *Hills of the Shatemuc* [COHA])

These initial *that reminds me*'s function as topic-shifters and introductions to face-threatening topics. Digression markers signal a side-trip in an argument and presuppose some coherence with the main argument. However, *that reminds me* in (28) is used to introduce a completely new topic abruptly and is followed by several lines of about-the-bush teasing suspension about what the news is. *That reminds me* is not used as a digression marker, despite Fraser's categorization, so much as a marker of the type that he considers to introduce a new topic (Fraser 2009a: 895). Fraser's examples of the "new topic" category include *but* and *to change the topic*. Both of these, especially *to change the topic*, *on a different note* seem more appropriate paraphrases in (28) than either *by the way* or *incidentally*.

The last example, *to update you*, appears 130 times in the 2020 release of COCA. It is typically used in medial position following a clause such as *we want/need/continue*, or *I'd like/my goal will be*, but there are 11 pre-clausal examples, 4 of them introduced by the downtoner *just*. Since 10 of the examples are followed by *on X*, they do not qualify as potential DSMs that are independent of the structure of the clause. The one eligible example is:

(29) You remember the scene from when he went to Puerto Rico last week and threw paper towels at people. **To update you**, as of today 40% of the island does not have water, 85% does not have electricity. But what the president seems most concerned about and this may come as shock is himself. (2017 *CNN_Anderson-Cooper* [COCA])

The verb *update* presupposes prior discussion or at least information. In the case of (29) the topic, Puerto Rico and by implication the hurricane there, that has been mentioned earlier, is in the immediate proximity of the update. This update continues the prior topic, the president's response to the hurricane, introduces a 'shocking' reality to mock throwing paper towels, and then returns to the topic of the president's response to the hurricane, which Anderson-Cooper regards as irresponsible.

The clausal formulae discussed in this section are fully contentful and "addressable". They can be challenged. They are lexical. They are Connectors that are used to effect a topic-shift, but appear to have no, or at best minimal pragmatic function and there is no evidence that they are used like DSMs in the sense discussed in this book. They are so contentful and transparent in meaning that they are best characterized as "lexical Connectors", to modify the term "lexical connective" proposed by Cuenca (2015) for Pragmatic Markers in Catalan that are on the contentful end of the continuum from lexical to procedural use (see Chapter 4.3, Figure 4.1).

8.5 Summary

To summarize:

a. *By the way, by the by, incidentally* and *parenthetically* support the first phase of the DSM Trajectory Hypothesis, from CircAdv > [[Conjunct] ↔ [1DSM]] status but only *by the way* supports the second phase to [[Conjunct] ↔ [DM]] status. There is, however, marginal evidence that *incidentally* may be undergoing a shift to DM use.
b. Of the set of *E*s discussed in this chapter, only *by the way* was conventionally intersubjectified as a hedge on D2.
c. *By the way* has come since the 1680s to be the default way of expressing SP/W's stance that D2 is a digression and is being minimized in some way.
d. *By the way* appears to have been developed primarily in formal registers, many of which are translations.
e. *By the way* came to be used as a hedge on a face-threatening D2 (often a directive) during the 19thC.
f. *By the by* has been used from the 17thC on to connect clauses with the meaning 'in passing'. It has never been frequent and has been obsolescing from the early 19thC on.
g. Some formulae that have been analyzed as expressing a side-trip from the main trajectory of an argument, *I almost forgot, I just remembered, that reminds me, to update you* are used relatively infrequently and are very rarely found in the corpora in pre-clausal position. While there is evidence that they all can be used with a lexical connective function and are subjective in the sense that they express the point of view of SP/W regarding the shape of argumentation in progress, it appears that they have little to do with digression *per se*. Typically, they do not minimize the content of D2. Instead, they signal a shift to a new topic. Also, they are on the extreme lexical end of the lexical-procedural continuum, and are best thought of as "lexical Connectors".

The pragmatic differences between the various expressions discussed in this Chapter draw attention to the importance of nuanced text-building among cognitive skills. Furthermore, *by the way* shows that scholarly writing traditions can be highly influential in the development of DSMs.

An approach to expressions with Connector semantics that do not have the pragmatic DSM properties characterized in this book is elaborated on in more detail in the next chapter, in which the data are markers of return to a prior topic.

CHAPTER 9

The development of markers of Return to a prior topic

9.1 Introduction[49]

In Chapter 8.4 I suggested that some markers with Connector function are primarily lexical and do not have the pragmatic characteristics of DSMs as discussed in this book. The present Chapter elaborates on the idea of "lexical Connectors" and on comments regarding the problem of analyzing as DMs some of the expressions that are used to mark connectivity between D1 and D2, in the present case, the class of markers of the topic-shift Return to a prior topic. The main purpose of discussing this set of Connectors is to suggest how a CircAdv may come to be used as a 1DSM Connector.

Fraser (1996: 339) lists among DM topic change markers some *E*s that are DMs in my sense: *by the way, anyway*, the 1DSM digressives *incidentally*, *parenthetically*, and a set of relatively lengthy expressions including: *back to my original point, before I forget, returning to my point*. The class is recategorized in Fraser (2009a: 894) as a subset of discourse management markers. Marking of intention to Return to a prior topic serves a double function. Specifically, it is used to invite AD/R to reinterpret what precedes as a relatively lengthy digression and to interpret upcoming content as a continuation of a topic that was being addressed prior to the digression. This is an "endophoric" or "Janus-like" use that is both backward and forward oriented.[50]

Examples of markers that Fraser (2009a: 894) cites are: *back to my point, returning to my (previous) point, to return to the prior topic, I would like now to go back to what I was discussing, that point notwithstanding, if I might return to my prior point*. They serve to link D1 and D2 in the broad sense of larger, "global" discourse segments, and therefore serve a Connector function. As Fraser notes, they are typically long rather than short, sometimes even an entire clause. Most "have only one

---

49. This Chapter builds on parts of Traugott (2020c). In that paper I treat *back to X point* as a Conjunct in some of its uses. Here I hypothesize that it is a lexical Connector that may be undergoing a shift to Conjunct use.

50. See König (2020) on the endophoric character of many demonstratives. The metaphorically "Janus-like" character of some [[Conjunct] ↔ [DSM]]s has been identified in various works such as Fetzer (2014) and Mulder and Thompson (2008).

semantic meaning" and that meaning is "tangible" (Fraser 2009a: 894). They are "addressable" (Mittwoch et al. 2002; Boye & Harder 2012) and could potentially be refuted (*That's not a return to your point!*). Return to a prior topic is a subset of lexical *E*s that can be used for purposes of topic orientation. Cuenca (2015) calls Connectors like this "lexical connectives". In keeping with terminology in this book, I use the term "lexical Connector".

The subset of topic-orientation markers Return to a prior topic is distinctive in several ways beyond consisting of lexical Connectors. It is these differences that makes the class of special interest as it helps underscore some of the default properties of the DSMs discussed so far. Also, since *back to X point* and *back to X topic* may be undergoing incipient shifts from framing CircAdv to lexical connectives to Conjunct status, they provide some insight into the kinds of linguistic experiments SP/Ws use that may lead to constructionalization.

Return to a prior topic is the only subset of discourse management markers that Fraser (2009a) cites with no members that are DMs as defined in this book. However, in the other three sets that Fraser discusses there is at least one DM: *but* (Continuation with the current topic), *by the way* (Digression from the current topic), and *but, now, then* (Introducing a new topic).[51] Return to prior topic is also the only set in which all the members that Fraser (2009a) cites are compositional at the level of the phrase. Construction-internally, however, *notwithstanding* in *that point notwithstanding* is non-compositional. *Withstand* 'maintain a position against' has been a fixed lexical item since the late 9thC and *notwithstanding* has been used as a preposition and contrastive Conjunct since the 15thC according to the OED. *Withstand* without the *not* is moderately opaque compositionally, but analyzable, much like *withdraw* and *withhold*, where *with* is understood as 'from' rather than 'against'. However, *notwithstanding* is fully opaque (but analyzable) since it means 'taking a position against X' and *not* emphasizes the negative semantics of *withstand* rather than negating it.

In addition, the set of Return to prior topic markers is distinctive in that few of Fraser's examples have long and well-attested histories. Furthermore, few of Fraser's examples appear with any frequency in clause-initial position. This suggests that to date the discourse function that markers of Return to a prior topic serve is somewhat marginal in everyday language use.

In this Chapter I outline the histories of the expressions *to return to X point*, which is attested in EEBO (9.2), *back to X point* (9.3), and *back to X topic* (9.4). The possibility that *back to X point* and *back to X topic* are undergoing a shift in

51. Fraser (2009a: 895) cites *but* as both a discourse management marker signaling continuation with the current topic and as introducing a new topic. As mentioned in Chapter 7.2, in Fraser (2009b) he says that because *but* does not have the semantics of contrast in these functions it is not a DM (in his sense).

function is discussed in Section 9.5. Section 9.6 summarizes. Although there are some distributional differences, *to return to X point*, *back to X point* and *back to X topic* mean approximately the same thing when used with discourse structuring function. Like *by the way*, they are historically based on the metaphor AN ARGUMENT IS A JOURNEY, cf. *return, back*.

9.2 *To return to X point*

Both *return/retourne* and *point/poynt* are borrowed from French. OED cites occurrence of *point* from 1225 on and of *return* from 1325 on. *Point* is usually used in Middle English with the literal meaning 'starting point' (cf. *point of departure*), but in the expressions discussed here it is used in the figurative sense of 'matter under discussion, argument' (OED *point* 10).

According to Biber et al. (1999: 826), in contemporary English *to*-clauses like *to return to X point* are CircAdvs that are favored in academic prose and that tend to convey purpose (p. 828). In historical texts, *to return/retourne to X point* is a CircAdv and is largely restricted to formal writing. In the context of argumentation, when used in clause-initial position it may be regarded as a topicalized CircAdv that frames the upcoming D2. Because it also looks back to D1, it serves a linking function and may be considered to be a lexical Connector in this position. *To return/retourne to X point* appears 11 times in initial position prior to a finite clause in EEBO from the 1630s on. 3 examples attest *my*, 2 *our*, and 6 *the*. An expression with similar meaning, *to return/retourne to X purpose* is attested from the 1520s on in EEBO and is more frequent than *to return to X point* (123 examples in initial position in EEBO in all spellings compared to 11 *to return to X point*). However, *to return to X purpose* is not attested in either COHA or COCA and therefore appears to have obsolesced rapidly. Scattered examples appear with the *-ing* form, *returning to the X* (where X is a noun like *argument, claim, proposition, script* referring to a locutionary act), but a location is far more likely to occur in X than a discourse-related term.

Examples of *to return/retourne to X point* that appear to be lexical Connectors include:

(1) this is a plain discovery to me, that the whole was made and contrived by some jesuite, … , although it is possible that some honest men may be coosened (*cozened* 'tricked'$_{\text{ECT}}$) into the acting of it for them: but **to return to the point**, here we have a magistrate that hath neither compulsive nor restrictive power in matters of religion. (1648 Ashhurst, *Reasons against agreement with a late printed paper* [EEBO])

In 8 cases, *point* is modified by *main*, so the *E* is compositional and not fixed. In (2) the author, Gouge, comments that he has digressed allegedly for the purpose of keeping weak persons from despair (*by the way* is used in the Conjunct sense of 'in passing'), and then signals return to his main objective:

(2) this by the way i have noted to keep some weak ones from despaire, who from the forenamed text, numb: 15: 30, have inferred, that every presumptuous sin is unpardonable: but **to return to the main point**, the apostle useth such in non-latin alphabet:
(1655 Gouge, *A learned a very useful commentary* [EEBO])

In (2) we find 'noting', and 'inferring', the kinds of context in which spatial *by the way* came to be used as a DSM. *By the way* came to be used more frequently in clause-linking contexts from the 1650s on. By contrast, *to return to X point*, which was already infrequent in EEBO, came to be used even less frequently in COHA. Here it appears in initial position only with *the* in X (3 examples), all followed by a relative clause referring to the preceding digression, as in (3).

(3) we fearlessly repeat, our firm assurance that these bonds will yet be recognised and paid, from the teeming treasures which every year developes from their prolific soil. **But to return to the point** from which we have been thus tempted a little aside, it certainly was not generally understood at the time as so clear a case of gross unconstitutionality as …
(1842 U.S. Democratic Review, *Mississippi bond question* [COHA])

To return to X point used as a lexical Connector has essentially obsolesced by the time of COCA, where there is only 1 clause-initial example. Here X is *your*:

(4) There's a reason why two-thirds of young people now believe that socialism is a better solution for American economics than capitalism. It's because they've lost (INAUDIBLE). *Chen*: But, John, **to return to your point**, it would be interesting to see the first fault line between the Republican majority in the Congress and the president. (2017 CBS *Face the Nation* [COCA; Traugott 2020c: 10])

In sum, *to return to X point* is adverbial, contentful and compositional. Although it expresses SP/W's intention to go back to a prior topic after a digression, and there was the potential to use it as a Connector of finite clauses in the 17thC, SP/Ws did not conventionalize this use. It is infrequent in any use, but restricted to use with prepositional complements (e.g. *point of origin*), and restrictive relative clauses, as in (5):

(5) one inescapable conclusion is that no general empirical case can be made that progressive policy has harmed the United States economy … To the contrary, the shoe is on the other foot. # Nevertheless, **to return to the point** that opened this section, economic growth is hardly the sole driver of policy choices.
(2017 *Iowa Law review* [COCA])

In sum it functions, as Biber et al. (1999) say, as a CircAdv of purpose.

9.3 *Back to X point*

In EEBO *back to X point* is not attested in initial position. Rather, it follows a verb like *come*. In one text translated from French, *point* is modified by *principal*. One example each of clause-initial *to come back to the point* and *to get back to the point* appears in the 19thC, as evidenced by COHA. Like *to return to X point*, these *E*s are CircAdvs conveying SP/W's intention:

(6) a. The great body of people in every country are idolaters. They worship the image or form rather than God, the living principle of goodness. But, **to come back to the point**, I believe God does mark ('note') the conduct of men. (1843 *New Englander and Yale Review* [COHA])
b. But the way my blood pressure is these days because of all the aggravation I'll probably go first. Anyway, **to get back to the point**, I tell the both of them that I know they want to be alone.
(1981 Owens, *Chucky's hunch* [COHA; Traugott 2020c: 10])

There are no examples in the 2021 release of COHA with a personal pronoun determiner (*my, our*…) and none of initial *back to the point* until the end of the 20thC. (7) is from a play:

(7) In America it is reserved for top firefighters and military men – Men and women who deploy muscles to grapple with temporary problems like wild fires and war. Good. **Back to the point**, falsity should not be allowed to penetrate reality,
(2001 Omtatah, *Chains of junkdom* [COHA])

While there were only 2 examples of clause-initial *back to the point* in the 2019 COCA, there are 25 of *(to go/get) back to the point* in the 2020 release of COCA. 2 of these examples appeared in the earlier 2019 version of the corpus. The 23 new examples occur in the web and blog data of which (8) is an example, suggesting that this expression has come to be somewhat entrenched in those genres and may spread to others in future:

(8) I love the new cast members this year (not as much Vannessa Bayer, no offence). Anyway **back to the point**, Michaela and Casey will be missed.
(2012 *Exclusive* [COCA WEB])

In these genres the inherent semantic connectivity implicated by the lexical *E (to go/get) back to the point* appears to be foregrounded as it is preferred with a topic-resuming DSM preceding. These uses may point to the beginning of a shift from CircAdv to lexical Connector status and, as I suggest in what follows, possibly to Conjunct status.

The X in *back to X point* is usually *the*, but the personal possessives *my* or *your* also occur. *Back to my point* is mostly followed by a prepositional modifier (*back to my point about X*), or appears in the formulaic expressions *I go/get back to my point* or *It goes back to my point that*. Out of 69 hits of *back to my point* (raw count), 13 meet the criterion of use in initial position immediately followed by a finite clause that I have used for identifying Conjuncts. (9b) shows that it may follow *going* or *getting*:

(9) a. A mortgage on a marital homestead without the signature of both spouses is entirely void. But **back to my point**. People are deluding themselves if … (2012 *Foreclosure fraud for dummies* [COCA BLOG])
b. I guess time doesn't heal all wounds. Well, that's for later discussion. But **getting back to my point**, before Miss Southerlyn spoke with Dr. Alston the State didn't have enough evidence to charge Dr. Heinz with littering.
(2003 *Law and order* [COCA TV])

There are also 12 examples of *back to your point*, sometimes followed by a colon or a period (10), and some variants like *going/getting back to your point*.

(10) He was designing the B–1 bomber at Wright-Patterson Air Force Base. But **back to your point**. So customers are demanding of these authentic flavors.
(2014 *NPR Here and Now* [COCA])

In sum, it appears that *back to X point* is coming to be used in exploratory ways with different pronouns in X and various kinds of verbal expressions preceding. It is possible that these explorations could lead to neoanalysis as a Conjunct with 1DSM meaning. If so, a profile shift from lexical, contentful meaning to less truth-conditional, more pragmatic meaning can be expected. On the form side, fixing of the expression without adjectival or complement modifiers can also be expected. This is a potential change that could be tracked over the next few decades.

9.4 *Back to X topic*

According to OED (*topic* 3), the word *topic* is a 16thC borrowing from Greek *τοπικά* 'set of rules or maxims' and was not used in the sense 'subject of discourse, argument or literary composition' until the early 18thC. *Topic* is not to be found in COHA used in a way that can be analyzed as Connector use, such as (*to get/getting) back to X topic*. *To return to the topic* is used in initial position only with a complement, e.g. *to return to the topic of the original posting*. In the COCA there are two examples of *going back to the topic* (e.g. (11)), and one of *to come back to the topic*. Because *going* and *getting* may precede *back to X point*, it may be that a formula is currently be developing with *going/getting* + back to X N, where N is *point* or *topic*.

(11) you have put so much work into your blog, bless you. **Going back to the topic**, of course, on a more general level, people fear 'celebrities'.
(2012 *Michael Jackson justice* [COCA WEB])

In (11) *going back to the topic* introduces a D2 modified by the evaluative DM *of course*, and a framing CircAdv *on a more general level*, both of them pointing forward to D2.

As for bare *back to the topic*, it appears 13 times in blog and web data, as in (12a) and twice with *OK* as in (12b), marking the resumption of the prior topic:

(12) a. how dare you ignore American Idol like that?) **Back to the topic**. The DOE and UN predict capacity of around 700 quads in 2035. (2012 *What does population have to do with climate change?* [COCA BLOG])
b. I have to stop and catch my breath from laughing so hard. OK, **back to the topic**. Smelling a lawsuit with the facts posted for everyone to see, Harrison ordered the entire thread removed.
(2012 *Let's discuss those ideas* [COCA BLOG])

There is one example of *back to my topic* (13a), and two of *back to our topic*, one of them introduced by *getting* (13b):

(13) a. I'm horrified by the typos, some so bad that you loose [sic] the meaning of what I'm trying to say. So, **back to my topic**. I was thinking of the dangers of walls. (2012 *The Christian monist* [COCA BLOG])
b. This revolution … gave us access to things that were until then too expensive or even non-existent. **Getting back to our topic**, we're seeing increasing complexity in the world of Web design.
(2012 *The future of CSS* [COCA WEB])

The examples in COCA 2020 suggest that *back to X topic* may be coming to be used as a lexical Connector because it is cataphoric to D2. It is probably modeled analogically on *back to X point*, which is both older and somewhat more frequent in the 2020 COCA (23 *back to the point* vs. 13 *back to the topic*) (see Bybee 2003 on the importance of exemplar models for change). While *back to X point* and *back to X topic* are roughly equivalent, the presence of two instances of *OK, back to the topic* suggests that *back to X topic* may be associated with sharper topic shifts than *back to the point*. It has been attested only very recently (2012 in the 2020 release of COCA), so there has not yet been time for conventionalization and stabilization in the corpus.

9.5 Discussion

The type of topic-orientation labelled by Fraser Return to a prior topic is instantiated by mainly contentful expressions in English, if it is marked at all. The *E*s in this set that I have looked at above are at this point in their histories and as far as can be told from texts, [+truth-conditional] [–pragmatic] and [–conventional], and infrequently used. As we have seen, markers signaling Continue with the present topic (called Elaborators here) can be found in the earliest Old English texts functioning as Conjuncts. Markers of Digress from the present topic appear in the 17thC functioning as Conjuncts. Lexical connective expressions signaling Return to a prior topic are also attested in the 17thC. However, there is sparse evidence that these expressions have been conventionalized as lexical Connectors, and even less that they are used with Conjunct syntax or DSM characteristics. To meet the criteria of Conjunct syntax they would have to be used to directly introduce a finite clause, without a following prepositional complement or restrictive relative clause.

To return to X point, back to X point and *back to X topic* all were all initially used as CircAdvs. At first, the nouns *point* and *topic*, like *way* and *stead*, were largely used with adjectival and/or *PP* modifiers, typically *of NP*, e.g. *to the main point of the Controversie* (1681), *the fine topic of the conflict* (1822). Used in clause-initial position as framing adverbials referring to cognitive and locutionary acts, *back to X point* and *back to X topic* came to be interpretable as lexical Connectors marking the relationship between D1 and D2 in extended discourse. They are used to guide interpretation in context. Specifically, they are used to invite AD/R to reinterpret what precedes as a digression from the main argument and to interpret upcoming content as a continuation of a topic that was being addressed prior to the digression. In terms of genre contexts, COCA release 2020 suggests that, in present day American English, blogs and the web may be preferred loci for use of the Return

to a prior topic type of topic-orientation markers. *Back to X point* used in initial position and with possible incipient Conjunct properties appears in the data in the 1990s and *back to X topic* in the second decade of the 21stC. Return to a prior topic is therefore the most recent of the sets of topic orientation markers identified in Fraser (2009a) and discussed in this book. Examples are innovations that are only potentially coming to be conventionalized.

Various options appear to be being explored:

a. fixing of the goal *to X point/topic* without a modifier,
b. fixing of the X; this is usually *the*, but personal pronouns, especially first person singular, are found as well,
c. introducing the formula with such expressions as *going/getting/returning to*, or *to come/go/get back to X topic.*

This kind of gradual process of sifting and replication is found in the data for earlier examples of the rise of [[Conjunct] ↔ [DSM]]s. It gives contemporary support to the hypothesis that the procedurals discussed in this book arose gradually. Which of the options becomes conventionalized (if any does) can potentially be tracked during the next few decades. So can whether any of them comes to be used as a Conjunct/DSM. This would require evidence of some loss of compositionality.

Since some of the examples cited in Fraser (2009a: 894) turned out not to be attested in either COHA or COCA, e.g. *that point notwithstanding, to return to the prior topic*, a reminder about the limits of both intuition and corpus data may be useful.[52] If an *E* such as *to return to the prior topic* can be constructed, this shows that assemblies of the relevant individual micro-constructions are possible. Absence in the data shows that the construct has not been used in the data (it is, however, usable and replicable). Corpus data and constructed data can usefully complement each other (Schönefeld 2011). Knowledge of a language goes beyond what can be found in corpus data precisely because constructions can be assembled.

52. There are, however, some examples of *notwithstanding/not withstanding* in COCA used as a contrastive meaning 'in spite of that', e.g.:

(i) As such, the panel concluded, "there may be no margin of safety." **Notwithstanding**, the scientists acknowledged that … (2000 *Science news* [COCA])

9.6 Summary

The key points of this Chapter are that:

1. Markers of Return to prior topic are highly lexical.
2. They exemplify multiple exploratory shifts in use of CircAdvs that could lead to conventionalization as Conjunct/DSM use.
3. To be analyzed as a Conjunct DSM, an *E* expressing the concept Return to prior topic must be used with some loss of referentially and compositionality. This does not appear to have occurred yet.

In the next Chapter I consider the phenomenon of constraints on combining two or more DMs.

CHAPTER 10

The development of combinations of DMs

10.1 Introduction[53]

One of the topics that has been of considerable interest recently is what constraints there are on the linear ordering of combinations of what I call Pragmatic Markers. Combinations of the "discourse markers" that Schiffrin (1987) identified preceding a clause (*and, because, but, I mean, now, Oh, or, so, then, y'know, well*) have been studied in detail from a synchronic, quantitative perspective in e.g. Koops and Lohmann (2015), Lohmann and Koops (2016) drawing on the Fisher corpus of American telephone calls; see also Aijmer (2002) on combinations ('collocates') with *now, Oh, ah*, and *actually*. Fraser (2015) and Fraser and Traugott (2017) investigate combinations of DMs in Fraser's sense (i.e. combinations of DSMs that link D1 and D2) from a qualitative perspective, also in PDE. Haselow (2019) addresses initial and final combinations in conversational turn-taking and proposes that the order of DMs (in Schiffrin's sense) reflects the order of communicative tasks to which speakers are "cognitively oriented at turn-beginning and -endings" (Haselow 2019: 2).

In this Chapter the main focus is on the distinction between combinations of independent DMs (e.g. *after all*, *but*) and combinations that suggest fusion or chunking of a combination that results in a new construction. For example, *now then* is a unit distinct from independent uses of pragmatic *now* and *then* (see Schiffrin 1987: 230). In other words, there are three separate constructions in the constructicon: *now, then* and *now then*. As discussed in Section 10.3, *now* and *then* have been used pragmatically in discourse structuring ways since Old English; the unit *now then* came into being in the mid 16thC. The question is how this happened and what the evidence for it is.

I consider combinations to consist of the co-occurrence of two or more DMs (and sometimes 1DSMs) that can be used independently and are prosodically integrated, at least to a degree (Lohmann and Koops 2016: Section 2.2), for example:

---

53. Many thanks to Bruce Fraser for inviting me to think about what DM combinations are possible in English and whether the positions for these combinations are meaningful.

(1) you are to abound ['be rich'] in all love and ready submission unto the doctrine of your faithfull shepherds, **and yet also**, not so to admire any person, as thereby to be drawn into errours or falshoods.
(1659 Burgess, *The scripture directory* [EEBO])

In the data, combinations of discourse structuring DMs that I have investigated are pre-clausal. In post-clausal position, combinations are very scarce. The only combination I have noted is *after all* followed by a possible epistemic PM *you know* in (2). However, as Mrs. Brent is addressing herself, *you know* could be a main verb here:

(2) A few minutes earlier, Mrs. Brent had been saying to herself inwardly: "She has not much heart **after all, you know**." (1907 Burnett, *The Shuttle* [COHA])

I do not consider combinations in final position in this Chapter. Nor do I consider combinations with subordinating Conjuncts, e.g. *for, because, although*.

There are several proposals regarding combinations in the literature to date. Among them two are mentioned here, one by Koops and Lohmann and Lohmann and Koops, the other by Haselow.

Lohmann and Koops (2016: 433) proposed an optimal "sequencing hierarchy" for the Schiffrin DMs:

(3) *oh* > *well* > *and* > *so* > *or* > *but* > *because* > *then* > *you know* > *now* > *I mean*

(3) is based on rank ordering work reported in Koops and Lohmann (2015), in which the frequencies of all 110 theoretically possible combinations of the 11 DMs were calculated. Called a "hierarchy", in keeping with implicational hierarchies in typology (e.g. Comrie 1989[1981]), (3) is a statement about the linear order in which adjacent markers are most likely to occur: "The predictions … are that any given DM precedes all DMs to its right in the hierarchy, while following all DMs to its left", in other words, *oh* before *well*, before *and*, before *so*, and so on. Most of the combinations were found and analyzed in pairs, but 3 may collocate, as in (1) above (*and yet also*). From COCA an example with *Oh* is:

(4) See what you've done. **Oh, well, so** you'll just write her a bigger, fatter cheque.
(1995 *Frasier* [COCA])

A basic assumption of Koops' and Lohmann's research is that "DMs in sequences have related functions, or functions that are at least compatible" and not random (Lohmann and Koops 2016: 430). Koops and Lohmann (2015: 252) propose an abstraction based on (3) above:

(5) coordinators (*and, but, or*) > subordinators (*so,*[54] *because*) > adverbs (*now, then*) > matrix clauses (*I mean, you know*)

54. It is not clear why *so* is labelled a subordinator. Schiffrin (1987: 191) explicitly contrasts the "marker of main ideas" function of *so* with the "marker of subordinate ideas" function of *because*.

"Matrix clauses" is a term that reflects an assumption about sources, specifically that PM *I mean* derives from the main clause *I mean that*, and PM *I think* from the main clause *I think that* (see Thompson and Mulac 1991).[55] This assumption has been challenged in Brinton (2006, 2008), where it is argued that some PMs that might appear to derive from matrix clauses are actually derived from appositional expressions like *as I mean*, or subordinate clauses like *as I think* rather than from complement-taking main clauses. The term "comment clause", adopted from Quirk et al. (1985) in Brinton (2008), is the most theory-neutral for this type of PM.

By contrast, Haselow's (2019) proposal is that in turn-taking combinations are constrained as much by the logic of communication as by that of compatibility. He presents the Discourse Marker Sequencing Hypothesis (Haselow 2019: 6) in (6):

(6) "The sequential order of discourse markers reflects the temporal logic of communicative tasks to be performed at a particular moment in turn production".

Haselow proposes that turn-initially the tasks are, among other things:

a. interactional, e.g. getting attention, responding to prior talk (*Oh, well*), providing interpretive cues for upcoming talk (*yeah, well*) (social PMs),
b. structuring: integrating with surrounding discourse (*and, but, so*) (discourse structuring DMs in the sense adopted in this book),
c. cognitive: interpretation in terms of certainty (*I think, you know*) (epistemic PMs).

Turn-finally, the tasks are, among other things:

d. interactional: facilitating addressee response,
e. cognitive: providing a last interpretive clue before transitioning, interpretive fine-tuning of the preceding message, e.g. in terms of epistemic value, or illocutionary force, cancelling possible implicatures,
f. structuring: retrospective integration of the turn into ongoing discourse.

Since my data provide minimal evidence for turn-taking, I do not pursue this important hypothesis here. However, the structuring and cognitive functions are helpful for thinking about the function of pre-clausal and post-clausal position discussed in Chapter 11 and will be revisited there.

A constraint on ordering that was first noted in Oates (2000), and has been frequently echoed since, is that more general DMs in the discourse structuring sense precede more specific DMs. Specifically, Oates concludes that "weaker", less explicit DMs tend to precede "stronger", more explicit ones, e.g. *but nevertheless*, not *nevertheless but* (she does not make a distinction between 1DSMs and DMs). Use of weaker before stronger DMs is confirmed by examples such as (1) and (4)

55. Following Schiffrin (1987), Koops and Lohmann (also Haselow 2019 and Oates 2000) refer to social and epistemic markers as "DMs", whereas I refer to them as "PMs" (see chapter 4.1).

above. In many cases, the "stronger" *E* is a 1DSM like *nevertheless*, which follows from the fact that 1DSMs tend to be more contentful than DMs. The constraint "weaker before stronger" does not appear to hold, however, if PMs that do epistemic or social but not discourse structuring work are considered. For example, *you know* is intuitively weaker than either *however* or *on the other hand* (both 1DSMs) in:

(7) **but however, on the other hand, you know**, this is as I said earlier, just an idea.
(1991 *PBS_Newshour* [COCA; Fraser and Traugott 2017])

All the studies mentioned above concern synchronic Modern English. According to Lenker (2014: 19) although scattered combinations are attested in Old English texts, "the beginning of Middle English saw a rapid increase in the number of sentences which are introduced by a conjunction (*and, but* or *for*) in a collocation together with an adverbial connector". Lenker associates this with attempts to develop a new English prose style. Combinations have increased in frequency over time since Middle English, partly because of the emergence of new DMs and no doubt in part because of the increased "colloquialization" of written English, that is, increased acceptance of informal linguistic options in many types of writing in later periods (Mair 1997).

The combinations I am concerned with in this Chapter are exemplified by *E*s that either were already DMs in the early data (e.g. *and*) or came to be used as DMs during the history of English (e.g. *but*). The constructions investigated are a subset of Schiffrin's DMs in (3): *but*; metatextual, largely inferential *now*; metatextual inferential *then*; inferential *so* 'therefore'; *Oh* used as a Conjunct; and in addition *also* (see Chapter 6.2). Combinations with initial contrastive *yet* are also considered, since they occur quite frequently. Over time, the functions of some of the DMs considered have changed. In the 1500s *so* was usually used in the comparative sense 'likewise'. In PDE *so* is an inferential; this inferential meaning came to be preferred in the 1600s. In Old English *ȝiet* 'yet' meant 'in continuation, still' except in contrastive contexts where it is associated with the meaning 'but, however'. The latter contrastive use of *yet* is the prototypical meaning of the DM in PDE. The history of *Oh* remains to be studied, but in the data, it is used mainly to express surprise in EEBO. Use as a Conjunct DM with topic-shifting and other discourse structuring functions is attested first in the 1840s in COHA.

Questions about combinations that are addressed in this Chapter are:

1. Do some combinations function as fused units ("chunks" in constructional terms)? Lohmann and Koops (2016: 429) cite *and then* and *but then*.
2. Are the meaning changes undergone by developing DMs like *so* constrained in any predictable ways when they are used in combinations?

3. Are combinations of DMs used in a predictable order?
4. Are sequencing slots "meaningful in themselves"; that is, does placement in a particular position "imply a particular level of abstractness" (Koops and Lohmann 2015: 256)?

The Chapter is organized as follows. DM combinations with the elaborative Connector *also* are discussed in 10.2. Section 10.3 is a study of the univerbation of *now then*. DM combinations with the digressive DM *by the way* are the topic of Section 10.4, with focus on combinations with Conjunct uses of *Oh*. Other combinations with *Oh* are briefly touched on in 10.5. Section 10.6 discusses the findings and Section 10.7 concludes.

10.2 DM combinations with *also*

In this section I explore combinations of DMs with the Elaborator *also* < *eallswa* 'likewise' (see Chapter 6.2). It is of interest because, although DM uses of *also* are attested from the 1500s on, it was not entrenched as a Connector until the 20thC. Nevertheless, it occurs in combinations throughout its history. Because the meaning of *so* changed over time, *also* provides a window on the ways in which combinations may show evidence of developments that are different from those of the Connectors used independently. Furthermore, *also* was itself a combination of two constructions that were independent, as evidenced by the correlative *eall X swa Y* in Chapter 6.1, Example (3).

In this section three sets of combinations with *also* are explored: *and* + *also* in Section 10.2.1, *so* + *also* in Section 10.2.2, and combinations of two DMs with *also* in 10.2.3.

10.2.1 *And* + *also*

In Chapter 6.2 Example (3) provides a Middle English example of *also* meaning 'in addition I confirm'. It is preceded for emphasis and highlighting of parallelism by two markers also meaning 'in addition', *and* and *eek*. Part of that example is repeated here as (8):

(8) **Also** ye han erred, for … **And eek also** ye have erred, for …
'Also you have erred, because … And in addition also you have erred, because …' (c1390 Chaucer CT, *Tale of Melibee* 1244–1246)

The combination *and eek also* is a one-off and was presumably used for rhetorical effect and possibly also to cue the relatively new meaning of *also* 'in addition,

confirmatory'. In early Early Modern English use of *and also* tends to be preferred for joining rhetorically parallel clauses, as in (9):

(9) for Saynt austyn sayth they that wene to knowe trouthe and lyuyth
for Saint Augustine says they that think to know truth and lives
euyll &; viciously it is folye yf he knoweth hit not **and also** he sayth
evilly and in-vice it is sinful if he knows it not. and also he says
in an other place that …
in another place that …
'for St Augustine says it is sinful for those who think they know the truth and live evilly and wickedly if they don't know it (that they live evilly), and also in another place he says that…'
(1474 Jacobus de Cessolis, *To Prince George Duke of Clarence* [EEBO])

This preference for syntactic parallelism between D1 and D2 persists throughout the data bases used, but the preference was weakened over time, as (10) shows:

(10) All he had to prove was that the world was worthy of him. Which is quite a task, when you think about it. **And also**, if you look at it right, it's a very humble, self-effacing trait. (1983 Arthur, *Beyond the mountain* (COHA)]

And also is preferred for indexing parallelism of finite clauses, as in (9). By contrast, for the most part, *but also* is preferred as a marker of coordination of Ns or Vs or clauses with null subject in D2 throughout the data. *But also* is used only rarely to introduce a D2, but when it is, it is typically preceded by *not only* in Early Modern English. Therefore the constraint that D1 and D2 are parallel is even stronger than with *and also* in that period:

(11) he pursued soo formosus that not only he withsayde his doyng and his dedes **but also** he lette take vp formosus whanne he was deede oute of the erthe.
'he pursued Formosus so intensely that he not only contradicted his actions and deeds but he also had Formosus taken up out of the earth when he was dead'. (1482 Higden, *Prolicionycion* [EEBO])

However, over time *not only* comes to be omitted and the contrastive pragmatics of *but* is weakened, as in (12):

(12) first because there's a lot of Secret Service between you and him, **but also** because that's the nature of the process.
(2012 *want-to-take-a swing-Romney son apologizes* [COCA BLOG])

We may conclude that as *also* was coming to be crystallized as a marker of Elaboration, SP/Ws found no need to use *and* or *but* as strengthening Elaborators.

10.2.2 *So + also*

I turn now to combination of *also* with so, an *E* that underwent considerable change during the history of English. The combination is not attested with great frequency in the data. In COCA 2000 there are 573 examples of *so also*, and 440 of *yet also* compared with 41,266 of *and also* and 64,512 of *but also*. These are raw counts, so many of the examples of *but also* and even of *and also* combine units other than clauses, but nevertheless the numbers are suggestive of considerable differences in use.

In Old English *swa* was a Conjunct with two meanings. One was comparative (see *swa* as a component part of *eallswa* 'likewise', the etymological source of *also*), usually in the paired construction *swa X … swa Y*. In contemporary English this is usually *as X … so Y*. The other use was resultative 'therefore'. This has been the usual meaning from the 1600s on. However, the *so* in the combination *so also* is comparative and might appear to be functionally partially redundant with the *so* in *also.* However it was probably not considered redundant as *also* was doubtless not compositional by Early Modern English. In EEBO *so also* is usually paired with *as* in the antecedent clause and can be considered to be part of a comparative formula 'as X … so likewise Y', as in (13):

(13) for fayth **as** it hath eyes, **so also** it hath eares:
'because just as faith has eyes, so also/likewise it has ears:'
(1538 Erasmus, *Preparation to death* [EEBO]

In some cases antecedent *as* is no longer required, as in (14). Here what Paul said is compared with what John said and implied to be parallel with it (note it is not parallel to the *therefore* that marks the immediately preceding clause):

(14) the servants of god are not to clap peoples heads, or indulge them in this inclination of theirs to error, they are not to humor them in it; but on all hazards they are to testify against it: therefore, sayes john, believe not every spirit, **so also** the apostle paul does very sharply inveigh against people, when they are beginning to tamper with error. (1692 Fergusson, *A brief refutation* [EEBO])

In (14) the comparative is implicated by *so also*, without any comparative marker. This is a case of "context-absorption" (Kuteva 2001: 150) in which an *E* comes to implicate the older construction. I have not found examples in EEBO in which the "result" reading of *so* is plausible. This suggests that *so also* is conceptually a unit meaning 'likewise' in EEBO with retention of the older meaning of *so*. *So also* occurs relatively infrequently (17,381 raw count as compared with 78,644 *and also*, raw count and with 89,492 *but also*, raw count).

So also is used in similar ways in COHA, again mainly with *as* (15a), but also implicating comparison in the absence of *as* (15b):

(15) a. **As** the world is one of the first enemies which Christian faith has to encounter, **so also** it remains one of the last.
(1829 Wilks, *Christian Essays* [COHA])

b. Subsequently they [the Greeks] had their "Tactici," – professors of the military art. **So also** among the Romans there were the "Campi Martii".
(1832 D.M., *West Point* [COHA])

In COCA *so also* is mostly used the same way as in EEBO and COHA, and mostly in religious texts and educational works. These uses suggest that *so also* continued to be processed from Early Modern English on as a conceptual unit. It has, however, come to be used less frequently over time, as the raw counts suggest: in EEBO 17,381 instances of *so also* is 22% of 78,644 *and also*; however, in COCA 753 instances of *so also* is only 1.3% of 41,266 *and also*.

In sum, we may conclude that SP/Ws of English have used *and* and *but* as independent *E*s that may combine compositionally with *also* in its various uses. *So also* has, however, been used as a non-compositional unit from the 16thC on, but its use has been on the decline since the 19thC.

10.2.3 Combinations of two DMs with *also*

So far, the focus of discussion has been on combinations of two DMs. But on occasion three may be attested in EEBO, as in (8) (*and eek also*). There are 4 examples of *but yet also*:

(16) there is no one blessing but ('except') he is the sole author and worker of it: **but yet also** he maketh use of others, severally, in various degrees of usefulness.
'There is no blessing of which he is not the sole author; but he also makes use of others, separately, in various degrees of usefulness'.
(1680 Owen, *Exposition of Epistle of Paul* [EEBO])

In (16), *yet* is understood in its contrastive sense 'however', coherent with *but*, not in its continuative, elaborative sense 'still, in addition'.

And yet also is well represented in EEBO, unlike other combinations with *yet also*. There are 171 hits, 46 of which meet the criteria for potential pre-clausal Connector use. While most examples attest to use of *yet* as an Elaborative, coherent with *and*, over the decades represented in EEBO there is an increase in use of *and yet also* with contrastive meaning ('and nevertheless also'), e.g.:

(17) our god himself describes himself what he is, &; what his name is, exod: 3: 14: i am that i am; **and yet also** that expression comes infinitely short of what god is.
(1657 Everard, *The Gospel treasury* [EEBO])

And then also and *but then also* are (like *then also*) typically used to express a temporal relationship in EEBO, but they may also be used for reasoning, especially in the later decades and in the context of conditional *if*:

(18) a. for **if** [it] be that the white apparell of the minister haue any force, eyther to moue the people, or the minister vnto greater purenesse, … then it is that which ought to be commaunded … although the contrarie were forbidden: **and then also if** there be a vertue in a whitegarment … it were meete that order were taken.
'because if it is the case that the white apparel of the minister has any power to move either the people or the minister to greater purity … then it is what ought to be commanded … even if the contrary is forbidden: and then also if there is virtue in a white garment … it would be appropriate if that order was taken'. (1574 Whitgift, *Defense of the aunswere* [EEBO])

b. your disputation must needs end in a clamour, where the voyce of god is of no authority: **but then also if** scripture should be the Iudge, you might well say, … (1640 Buxton, *Replie to a relation* [EEBO])

Of 7 hits of *so then also*, 3 meet the criterion of being followed by a finite clause (one is a later edition of an earlier text). In these *so* is comparative (as in *so also*), not 'therefore', and *then* is temporal. Note in (19) *as … so*, and *when … then*:

(19) now, you know, **when** dear and intimate friends are to part, **as** their love then runs strongest, and their affections are most intire and vehement, **so then also** they especially discourse of those things, wherein most of all they desire to be satisfied. (1651 Dell, *Christ's spirit* [EEBO])

Moving on to the period represented in COHA, triple combinations are rare in that data base. Of 23 hits of *and then also*, 5 are textual Connectors used to represent shifts in the center of consciousness's thinking process, e.g.:

(20) Or he would [b]e the base of a very large human pyramid or portage a war canoe by himself for a quarter of a mile. **And then also** there was the head counsellor, (1932 Kirstein, *Nothing at all* [COHA])

In (20) *then* is to be understood textually as 'in addition'. Of 8 hits of *but then also*, 4 are textual Connectors; *then* is understood inferentially. There are no instances of *so then also, so yet also* and *or then also* preceding a clause with a subject. There is one instance of *and yet also*, attributed to a foreign count in a somewhat incoherent interaction. COCA shows the same results as COHA for triples: a few relevant examples of *and then also* and *but then also*, one example of *and yet also* but no instances of the other combinations used pre-clausally.

Conclusions from the study of combinations of DMs with *also* are:

1. Possible combinations are limited to *and, but, yet, then. Or* and inferential *now* do not combine in the data with *also*.
2. *So also* meaning 'likewise' appears to have been used as a non-compositional unit from Early Modern English on. However, it declined in frequency from the 1840s and is relatively infrequently used in PDE as represented by COCA.
3. Combinations that come to be chunked as units are understood differently from their parts, i.e. they come to be minimally compositional. Therefore, changes affecting one of the contributing constructions independently, for example *so* in *so also*, will not affect the meaning of the unit.

10.3 The combination *now then*

In PDE *now then* is a chunked DM. Its spelling as two words reflects older spelling conventions (cf. *after all, by the way*, *so also*). As CircAdvs, *now* and *then* are temporal and can be paraphrased as 'at this time' and 'at that time', respectively. *Now* can also be used alone meaning 'now, at this point in the text' and with a number of DM functions discussed in Schiffrin (1987) (see Example (19) above). *Then* can be used as inferential marker of conclusion ('therefore'). Combined as a unit they are non-compositional. "Where *now* signifies 'make a fresh start', *now then* signifies something like 'take a fresh look' – at something that has gone before: the *then* is the 'then' of consequence resulting from some intervening consideration that alters the view of things" (Bolinger 1989: 293; cited in Aijmer 2002: 650). Fraser (2009a: 893) classifies *now then* as an attention marker. However, *now then* also marks "transitions in topic, argument, activity" (Schiffrin 1987: 230) and can be considered to be a marker of topic-shift with inferential (reasoning) function. I explore the history of the unit *now then* in this section.

Using the cover term NU 'now' for the *E* of proximal temporality in various Germanic languages, Auer and Maschler (2016) show that it is quite old and has been used throughout Germanic not only as a temporal deictic (CircAdv), but also as a Conjunct meaning 'at this point in the argument, therefore'. Defour (2007: 72–76) discusses use of *now* in a textual sense from Old English on, noting that it is found mainly in the context of verbs of locution. She quantifies collocations with verbs of speaking in the Helsinki Corpus and other corpora, and cites Quirk et al.'s (1985: 640) observation that "the succession in time or space conveyed by the adverbial is converted into the logical succession of discourse when there is the implication of a verb of speaking". We have seen this principle operating in the case of spatial *by the way*.

In Old English *nu* mostly appears on its own, but in Ælfric's writings from the later Old English period it is sometimes attested with *for þi* 'therefore' and occasionally with *þonne* 'then'. There is no evidence that they form a unit at this time. As in PDE, *þonne* was used as both a distant temporal deictic and as an inferential Connector. Kemenade and Links (2020) provides detailed discussion of *þa* 'then', *þonne* 'then' and *nu* in Old English and of their

> functioning at the interface between syntax (in a fixed position as a grammaticalised functional head), information structure (marking the boundary between discourse-given and discourse-new information), pragmatics (expressing pragmatic meanings) and discourse management (linking the particle to the common ground between speaker and hearer). (Kemenade and Links 2020: 19)

The combination of *þonne* and *nu* is not favored in initial position in Old English but a scattering of examples can be found. The order is not fixed, and both *nu þonne* (21a) and *þonne nu* (21b) are attested. However, *þonne nu* is the dispreferred order, except after a DSM. The combination *swa þonne nu* 'so then now' occurs 5 times in DOEC and reinforces the textual argumentation:

(21) a. **Nu þonne** gif maria unbewedded wære …
Nu then if Mary not-married were …
(c1000 *ÆCHom* I,13 B1.1.14 [DOEC])

b. **swa þonne nu** gegongeþ þam mannum, that …
likewise then now befalls to-those men … þæt …
'likewise therefore it will happen to men, … that…'
(by end 10thC *HomS* 17 (BlHom 5) B3.2.17 [DOEC])

In Middle English and EModE, as in PDE, *now* and *then* may be used as temporal adverbials combined with a number of potential DMs, e.g. *and* or *but* + *because* or *therefore*. In these combinations *now* and *then* are used as literal CircAdv temporals, not as inferential Connectors.

However, *now* and *then* combined are used metatextually in Old and Middle English, and can be paraphrased as 'if we consider this to be true, then …' (Defour 2007: 189) or 'now therefore', as in (22). This is an inferential relation:

(22) **Nowe þen**, semeþ þe, be not þees sufficient resone?
now then, seems to-you, is not peace sufficient reason?
'Now therefore, does it not seem to you that peace is sufficient reason?'
(c1450(a1400) *Orolog.Sap*. (Dc 114)332/34 [MED *nou* 7.g])

The combination *now þen* "introduces a new step in the speaker's argumentation, based on the speaker's consideration of previous discourse steps" (Defour 2007: 76). Defour considers *now* to be "semantically bleached" in the combination.

MED [*nou* 7.g] cites translations from Latin *ergo* 'therefore' and a dictionary entry for *iam tunc* 'now at this point in the discourse'. The combination *now then* may also have been calqued from Latin *nunc ergo* 'now therefore' and *nunc igitur* 'now then, therefore'.[56] These were used for discourse structuring purposes in Late and Medieval Latin, especially in Biblical translations and works by Boethius and Bede that were translated and widely read in earlier English, e.g.:

(23) unde fit ut … vox illa permaneat. **Nunc ergo** quoniam
whence happens that …voice that persists. Now therefore since
secundum se divisionis differentiae dictae sunt, …
per se division-GEN different sayings are, …
'whence it happens that … that voice persists. Now therefore because the differences of the division itself have been stated …'
(505–510 Boethius, *De divisione* (CPL 0887))

In her discourse-pragmatic study of the Latin particles *nam, enim*, *autem, vero* and *at* (Kroon 1995) discusses many combinations that may have served as models for translators and writers in English. Kroon shows that the traditional classification of *nam* and *enim* as "causal" and *autem, vero*, *at* as "adversative *but*" is inadequate. From a discourse perspective, *nam* and *autem* are "presentational" Connectors favored in monologic texts, while *enim* and *vero* are "situating", interactional particles, and *at* is an "interactional" Connector favored in dialogal texts. The effect of Latin particles on English writing traditions deserves detailed study.

In EEBO 1470s-1550s there are a few instances of *now then* (also spelled *nowe then*): a total of 285 (raw count). These combinations are compositional. The relatively low frequency in the early part of the corpus suggests that *now* and *then* were separate Connectors that could be combined. However, in the decades of the 1560s-1640s there are 4,149 instances of *now(e) then* (raw count). The higher frequency in the later part of the 16thC suggests *now then* had come to be constructionalized as a DM unit. In pre-clausal position it is usually used to express topic-shift. It has been morphologically univerbated and pragmatically enriched. In other words, from 1560 on SP/Ws had three related choices of topic-shifting inferential expressions: *now, then*, and *now then*, each implicating a different perspective on the text. *Now* is cataphoric and draws attention to the upcoming D2, which is presented as a conclusion following from D1 (24a). *Then* usually marks D2 as the conclusion in the context of a conditional *if*-clause in D1 and signals an anaphoric logical connection with D1 (24b):

56. Many thanks to Piera Molinelli for this suggestion and for Example (23).

(24) a. neither could polycarpus perswade anicetus, to keepe the custome and tradition of asia: **now**, gentle reader, what neede more to be said, for the vncertentie of traditions?
'neither could Polycarpus persuade Anicetus to keep the custom and tradition of Asia; now, gentle reader, what more needs to be said about the uncertainty of traditions?' (1593 Bell, *Thomas Bels motiues* [EEBO])

b. for if there had bene a law geuen, which could haue geuen life, **then** no doubt righteousnes should haue come by the law.
'for if a law had been provided that could have given life, then no doubt that law would have produced righteousness'.
(1570 Gough, *The aunswer of Iohn Gough, preacher* [EEBO])

By contrast, *now then* shifts attention to an upcoming topic in D2 that is presented as of immediate relevance to what was said in D1, as in (25). When *now then* was univerbated, the anaphoric orientation of *then* was weakened.

(25) which he discouered vpon that triall: **now then** to speake a word or two of the action as it passed,
'which he discovered at that trial; now then to speak a word or two about the action as it occurred,' (1604 Parsons, *A relation of the trial made before the King of France* [EEBO])

The combination *but now then* is attested in the 1530s, and by hypothesis at this period is a combination of *but* with two Conjuncts, not of *but* + the unit *now then*. After 1560, in EEBO there are few tokens of combinations of DMs preceding *now then* + finite clause: *and* (24 examples), *but* (6), and *so* (6). (26) is an excerpt from a long sarcastic pun on monks who desire to wear a papal *bull* (edict on parchment issued by the pope) for show and eating meat of a *bull* (animal) on fasting day. It implicates that both are sinful desires:

(26) they must haue their Bulles well sealed, &; hanging downe with silken tassels: yea, and such as they themselues can well both read and vnderstand, otherwise it were all not worth a strawe: **so now then**, flesh (on fish dayes) is vncleane, to them that haue no Bulles.
'they must have their papal edicts well sealed and hanging down with silk tassles: yeah, and these edicts must be both readable and understandable, or it wouldn't be worth anything: so now then meat (on fish days) is unclean to those who have no papal edicts'. (1579 Gilpin, *The bee hiue of the Romishe Church* [EEBO; translated from Dutch])

Since *now then* has to do with reasoning and logic and is an inferential DM, *so* in (26) can be understood in its logical Connector meaning 'therefore' (contrast the 'likewise, similarly' meaning noted associated with *so also*). The possibility of

combinations of *and, but* and *so* with *now then*, is consistent with the Lohmann and Koops's (2016: 430) finding cited in 10.1 above that "DMs in sequences have related functions or functions that are at least compatible".

The unit *now then* has survived into PDE, and is quite frequently used. It appears pre-clausally in COCA where it is no longer as strongly associated with textual commentary as it was in the EModE data. Its use is reminiscent of Bolinger's (1989) "take a fresh look" cited above and implicates 'think again about what you may have assumed':

(27) a. There is no subsidy; it's a tax credit. **Now then**, who ultimately foots the bill for the taxes that the big oil companies pay?
(2012 *A series of logical* steps [COCA BLOG])

b. Well, now a story that made headlines because it exploded a few myths, and for the first time you're going to really understand what happened. **Now then**, do you think that the world would be a different place if women were in charge? (1993 *ABC_2020* [COCA])

The developments of the DM combination *now then* outlined above from Old English on can be modeled as in Figure 10.1. Since the source *Es* were already Conjuncts, the form change is use as a univerbated unit. The changes were of course far more nuanced than the Figure suggests. As discussed above, the inferential Conjunct meanings varied according to the discourse context and argumentation. The notation 'c1000–1600' indicates that combinations of the individual micro-constructions *now* and *then* preceded and presumably overlapped with the unit *now then*. "pdeictic" is short for 'proximal deictic' and "ddeictic" for 'distal deictic'. The D-F "attention-drawing" for *now* and *now then* is short for 'drawing attention to the textual organization and relevance of D2 to D1':

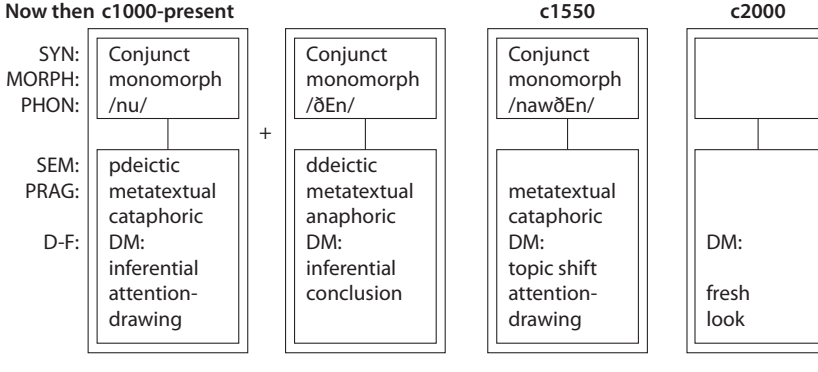

Figure 10.1 The development of DM *now then*

This study of the rise of the topic-orientation unit construction *now then* has shown that when combinations of individual micro-constructions come to be used as a unit the resulting meaning is not the same as that of the micro-constructions used separately. It is a constructionalization because form has changed (two monomorphemes have been united as one; nothing can intervene, and frequency of use has increased). So have pragmatic and discourse functional meanings. The newly formed unit occurs with greater frequency than the separate micro-constructions combined when it is first formed. Similar conclusions can be drawn about *so also* discussed in Section 10.2, although in the latter case the early history was harder to identify in the data because of the phonological changes and complex meaning changes undergone by Old English *swa* and *eallswa*. In both cases there was a marked increase in use when the new unit was formed, but use declined in frequency during the later Modern English period.

10.4 The rise of the combination *Oh, by the way*[57]

In EEBO there are several examples of Connector use of *by the way* preceded by one or two DMs. These are *and, but, for, if, so, therefore, yet*, the same set of DMs as collocates with locational CircAdv *by the way.* There are also scattered examples of *thus, howbeit, moreover*, which are compatible with Conjunct 'in passing' use. Connector *by the way* preceding a clause may also be preceded by two DMs: 35 instances of *and yet*, 21 of *but yet*, 14 of *and therefore*, and 3 of *and so*, e.g.:

(28) a. if after our rising we fall againe into the same or the like offence, this will cost deare: **but yet by the way** i affirme, that this may befall the child of god. (1633 Capel, *Tentations* [EEBO])

b. (Re ordinance to use the Book of Common Prayer] demonstrated the no-necessity, but plain unreasonablenesse of the change, **and so by the way** insisted on most of the defects of the directory. (1646 Hammond, *A view of the nevve directorie* [EEBO])

These are the typical of kinds of combinations discussed earlier in this Chapter. Note that *so* here means 'therefore'.

In this section I focus on combinations of *by the way* with *Oh*, as in (29):

57. I investigate only constructs with comma as those without comma are relatively infrequent. In COHA release 2021 there are 24 pre-clausal examples without comma to 493 with comma, in 2020-COCA 177 without comma, 100 of them in constructs found on the web and in blogs as against 1,104 with comma.

(29) a. Rick's a good, strong supporter, good man. And **oh, by the way**, for the other elected officials here, here's my word, in three short days, help is on the way. (2000 *CNN Event* [COCA])

b. We definitely were not treated the same way we'd been. We were told things like, **Oh, by the way**, you can't have your competitions on the weekends because … (1998 *Ms.* [COCA])

I posit two uses, in Section 10.4.1 a relatively non-evaluative use as in (29a), which I call OBTW1, attested from the 1840s on, and in Section 10.4.2 an evaluative one as in (29b), which I call OBTW2, attested from 1925 on.

10.4.1 OBTW1

There are no examples of *Oh* combined with *by the way* in EEBO. In EEBO *Oh* is used mainly as an exclamation of surprise. This is consistent with Culpeper and Kytö's (2010: 238–243) finding that in the *Early Modern English Dialogues* (1560–1760) *Oh* is largely used to express surprise and exasperation. Culpeper and Kytö also identify use as a politeness strategy, for examples in acceptance of thanks, usually before a vocative in dialogues. Most of their examples are from the 1700s, later than the data in EEBO, e.g.:

(30) **Oh**, Sir, I ask your Pardon, you're the Captain he told me of.
(1707 Farquar, *The Beaux Stratagem* [Culpeper and Kytö 2010: 279])

From the 1840s on in COHA release 2021 there are 480 examples of pre-clausal *Oh, by the way*, mostly in fiction. Combined with *by the way*, *Oh* appears typically to signal topic-shift and change of state (see Heritage 1994 et passim on use of *Oh* in conversational turn-taking) and hedging of a proposition that might be negatively assessed by AD/R, e.g.:

(31) a. we shall want it to take along for luncheon – you must get something! **Oh, by the way**, you may let the girls pick half a dozen quail and broil them, if you choose!" "Quail! do you say? and where'll I git quail, I'd be pleased to know?" (1845 Herbert, *Warwick Woodlands* [COHA; Traugott 2020b: 9])

b. This is a new comet, quite a stranger, they say wonderful, wonderful! I saw it last night. **Oh, by the way**, Jim, "turning again to the messenger," I want you to go down into the lower vaults today." The messenger followed the president silently. Of course, they wanted him to go down to the lower vaults. It was too dangerous for more valuable men.
(2004 Thomas, *Dark matter* [COHA])

(31a) is the first example of *Oh, by the way* in COHA. It appears at a turn continuation, and in a context that suggests that a hedge is intended. D2 is an indirect request (note that *if you choose* is appended at the end). Here *Oh* functions as a Conjunct DM. It is an attention-getter and marks change of state (shift from directive to indirect directive, and from indefinite *something* to specific *quail*) and change of topic. It is used to present D2 as something relatively trivial, as a hedge on the indirect order that follows. However, the grudging reply suggests that the mitigation was ineffective. Similarly, in (31b) *Oh, by the way* introduces an indirect order, this time followed by representation of the messenger's grudging thoughts. (31b) is a good example of the way in which *Oh, by the way* can be used as an attention-getter. It marks change of addressee as well as a topic change

Both examples in (31) exemplify an "as if" casualness in use of *Oh, by the way*. Fake casualness is commented on in (32), in which the protagonist fantasizes a scheme. This time the combination is *Oh, and by the way*:

(32) I have it all worked out. It's going to be perfect." Have fun at the dance tonight, Virginia, "I'll say. "**Oh, and by the way**, this is for you." Very casual-like. No big deal. (1999 Drake, *Virginia* [COCA])

There are 380 hits of *Oh, and by the way* in COCA. Of these only very few are evaluatively neutral; most are, like (32), embedded in a mocking or aggressive context. Of the 380 hits, 291 appear in data from movies and TV.

10.4.2 OBTW2

Over time, there is increasing use of *(oh,) (and) by the way* to introduce a challenging topic. The hedge appears to have come to be perceived as false and therefore an aggressive locution. In the 2020 COCA there are 4 explicit mentions in blogs of an "*Oh, by the way" routine*, as in (34):

(33) Everything the person says is suspect in my eyes from their [sic] on. It's that old "**oh, by the way …." routine**. I agree, it seems like more of this is cropping up on MIA. (2012 *It takes more than a* pill [COCA BLOG]]

All 4 mentions of the routine are critical of some aspect of the content of D2. One blog is about freelancing and the writer's perceived need to write up an "Oh, by the way" post. Another is in a blog caricaturing use of *Oh, by the way* and arguing that it should not be used when asking volunteers to do something. The point being made is that indirect directives preceded by *Oh, by the way* are perceived to be presented in a falsely casual and therefore unacceptable way:

(34) The leader handed her a huge manual and said, "**Oh, by the way**, we ask all volunteer leaders to read this." Lauren took it home and started to read it when the phone rang. The leader said, "**Oh, by the way**, we are giving a test on the manual. We want all our leaders to know the philosophy and what we expect from our volunteers." Lauren said, "I'm happy to read the manual" –she was being polite, "but I'm more the behind the scenes type of person. Can I be a helper?" The leader then said, "Well, we really want the parents to be leaders. **Oh, by the way**, we are having an eight-hour training session next Saturday and want all our leaders to attend."
(2012 *Volunteer power: Oh, By the Way* [COCA BLOG; Traugott In press: 155])

Such examples suggest that what was a combination of two DMs (*Oh + by the way*) may be becoming a unit DM in its mocking function. It's unit status is conceptual and functional, not segmental, as *and* can follow *Oh* (see (38) below), and prosodically there is a slight pause after *Oh*, reflected in writing by a comma.

Because in (33)–(34) SP/Ws mock styles of expression, the examples overlap functionally with a striking use that was attested first in the 1920s in which a third person ('Other') is alleged to have said, thought, or done something that SP/W considers shocking or at least inappropriate (Traugott In press 2020b). SP/W uses *Oh, by the way* to index his/her disaffiliation from the proposition (see e.g. Drew and Heritage 1992 on disaffiliation). In (35) the Other is a collective (the US Senate)

(35) [about congressmen in the US senate voting themselves a significant raise]. The whole affair was carried off in the most offhand manner, as if the Senate were remarking: "**oh, by the way**, of course we deserve more pay".
(1925/03/03 *Time* [Time Magazine Corpus: Traugott 2020b: Example (10)])

D2 is "pseudo-enacted" (Vandelanotte 2009: 138). Readers are not expected to think that all or even one senator actually said '*oh by the way, of course*…').

Key characteristics of OBTW2 are:

a. D2 is an alleged quotation.
b. The "as if" nature of the quotation.
c. D2 (and often the contradictions between attitudes expressed in D2 and D1) is normatively unacceptable in society.
d. Because OBTW2 marks an alleged quotation, it cannot be used to begin a turn, although it may claim to represent the beginning of a turn.

An example of OBTW2 in COHA is (36).

(36) Then I went to the Orthodox rabbis and said, "Will you come to the table?" And they said, "We'd love to come to the table." Then they added, "**Oh, by the way**, if the Reform are coming to the table, we're not coming to the table."
(1997 *SF Chronicle* [COHA]; Traugott In press: 155)

Many of the examples are politically incorrect and mock political figures who are perceived to have changed their position. 8 of the 16 hits in COCA SPOKEN for the year 2019 are of this type, e.g.:

(37) a. they had about 90 percent of the deal done, and then China said, **oh, by the way**, this idea of intellectual property theft and forced technology transfer that we've agreed to stop doing, now we're going to stop doing that.
(2019 *Fox – Ingraham* [COCA]; China is alleged to have said they will stop stopping intellectual property theft, etc.)

b. Now, what does it say to the entire world that you have a president who gets off the phone with Erdogan of Turkey, and then sends out a tweet and it's **oh, by the way**, we're deserting these people who have put their lives on the line to work with us. (2019 *ABC_This Week* [COCA])

Although the "*Oh, by the way* routine" cited in (34) suggests that the combination is coming to be univerbated, it has not yet been fully conventionalized as a unit as there is a variant *Oh, and by the way*:

(38) President Obama will come back one day and say, great news, Iran does not have nukes. **Oh, and by the way**, we are now the United States of Iran.
(2015 *Fox_The Five* [COCA])

However, OBTW2 is fairly advanced on the "cline from simple juxtaposition to the integration of co-occurring DMs under a more holistic combination serving one single function" (Haselow 2019: 1). This use is topic-oriented in so far as it typically elaborates on a topic in D1, but unlike OBTW1, it is not used to hedge and mitigate imposition on the negative face of an addressee. Instead, it exemplifies mock impoliteness (Beeching and Murphy 2019), not directly to the AD/R, but rather to the Other.

In conclusion, there are now two constructions with the form [*Oh, by the way*], one of which can be used either neutrally or as a hedge, the other signaling that SP/W distances themselves and disaffiliates from the content of D2 or, most recently, marking what follows as a deliberate face-threat and intended rudeness. While the prosody associated with each construction is, according to a small pilot study, variable, there are differences: in OBTW1 *by* usually receives primary stress, whereas in OBTW2 it is *way* that receives the stress.[58] This supports the two-construction analysis. While OBTW1 is a combination of *Oh* (attention-getter or topic-shift) + casual *by the way*, OBTW2 appears to be coming to be conceptualized as a relatively fixed unit. However, structurally it is not yet a unit because there is a small prosodic break after *Oh* and because *and* can be used after *Oh*.

58. I am grateful to Will Somers Clapp for providing some spectrographic analysis in the small pilot study.

10.5 *Oh* combined with other DMs

Oh is probably the most frequent DM in the corpora (other than the core Connectors *and, but, or, so*). But it is attested in combination with only a few of the DSMs discussed in prior Chapters.

a. Of the Elaboratives, *Oh* occurs relatively frequently with *and*. Of the Elaboratives discussed in Chapter 6, only *also* occurs with any frequency with *oh* (112 hits of *Oh, also* in the 1 billion word COCA, and 16 of *Oh also*). Most examples occur in the TV and Movies section of 2020 release of COCA, or in discussion of shows. There is 1 instance of *Oh, besides*, and none of *Oh* combined with *further(more), moreover, instead* used as DMs.
b. Of the Contrastives, only *but* collocates with *Oh*.
c. Of the Digressives, as discussed above, *Oh* combines with *by the way*. This is by far the most frequent combination with *Oh*, but the numbers are still low (1104 hits, raw count). There are 6 examples of *oh, by the by*. It also combines 26 times with *incidentally*.
d. For the clausal markers of return to prior topic that precede a clause, there are 146 examples of *Oh* preceding *I almost* forgot, 13 preceding *I just remembered*, and 89 preceding *that reminds me.* Many of these are punctuated as independent *E*s.
e. There is no example of a *Oh* combined with inferential *after all*.

This suggests that *Oh* is disfavored with Connectors other than *and* and Digressives.

10.6 Discussion

This section seeks to answer the questions at the end of Section 10.1.

Question 1: Are some combinations actually fused units ("chunks" in constructional terms)? The answer is positive. I have proposed that the independent constructions *so* and *also* and *now* and *then* were originally independent but came to be constructionalized as the single units *so also* and *now then*. These coexist with the independent constructions and are distinct from them by being non-compositional. I have suggested that OBTW2 may be undergoing such unification currently. A fully fused combination can be recognized by the main characteristic of univerbation: nothing can be used between the two items, but as is typical of morphosyntactic change, univerbation is gradual. *Oh, and by the way* in Example (38) by hypothesis indicates, along with prosody, that formal univerbation has not yet occurred (and of course it may never occur, because change is not determined). More importantly, in a fused combination the meaning of the unit is not compositionally that of the

Es combined. Furthermore, an older meaning may be retained, es exemplified by *so also*. Here the meaning of *so* persists from the time in Middle and Early Modern English when *so* meant 'likewise' rather than 'therefore'. This confirms Lohmann and Koops' (2016: 429–430) suggestion that univerbation may be accompanied "and in fact may be driven by, the development of idiomatic meanings or functions". OBTW2 has become idiomatic because it is identifiable with sarcasm, disaffiliation from D2 and rudeness, which the combination *Oh + by the way* in OBTW1 does not convey.

Question 2: Are the meaning changes undergone by developing DMs such as *so* and *yet* constrained in any predictable way when they are used in combinations? Here the answer appears to be negative. The only predictable factor appears to be that the pragmatics of the unit are stronger than those of the combined DMs. This is particularly clear in the case of *now then*.

Question 3: Are combinations of DSMs used in a predictable order? The answer is necessarily limited by the focus in this Chapter on the developments of unit functions. The unit combinations *so also* and *now then* are not attested in COCA with a DM preceding.[59] This means that only the independent DMs *also* and *by the way* can be drawn on to assess what slots need to be posited for maximal combinations of DSMs. For the triplet combinations that I discussed above 4 positions need to be posited, as in Table 10.1.[60] These are replicated orderings and appear to be conventionalized in PDE as represented by COCA:

Table 10.1 Slots associated in COCA with *also* and *by the way* in their DM functions

Slot 1	Slot 2	Slot 3	Slot 4
	and, but	*yet, then*	*also*
Oh	*and*		*by the way*

As mentioned in Section 10.1, Koops and Lohmann (2015: 252) posit the order in (5), repeated here as (39) for convenience. This is a syntax-based sequence:

(39) coordinators (*and, but, or*) – subordinators (*so, because*) – adverbs (*now, then*)

59. There is one example in COCA of pre-clausal *and now then* used at a turn, but since there is no pattern DM + *now then* this must be considered to be a hapax legomenon. *And* may be being used used to claim the floor here:

(i) decided to escalate rather than deescalate the situation. B: **And now then** we talk about what then happened to her in that jail cell three days later.
(2015 *CNN Burnett* [COCA])

60. Combinations in Table 10.1 are not exclusive. Others such as *now also* can be found, but they have not been discussed above (*now* is often temporal in this context).

However, Koops and Lohmann do not investigate DSMs as a category and subordinators were not investigated in the case studies above. Therefore (40) does not account for the findings summarized in Table 10.1. It is in fact difficult to associate the slots in Table 10.1 with DSM categories: *and* is a member of the Elaborator. Cxn, *but* of the Contrast.Cxn. *Yet* is also a member of the Contrast.Cxn, but is not used in the same slot as *but*.

This suggests a negative answer to question (4): Are sequencing slots "meaningful in themselves"; that is, does placement in a particular position "imply a particular level of abstractness" (Koops and Lohmann 2015: 256)? If slots were meaningful in themselves, rather than syntactic as in Koops and Lohmann's hypothesis in (40), one might expect constraints such as:

1. members of the same DSM type can be used as alternatives in the same slot,
2. members of the same DSM type are not expected to appear in combination.

If hypothesis 1. were supported by the data, *but* and *yet* would be expected to be alternatives available in the same slot. *And* and *but*, on the other hand, would not be expected to be used as alternatives in the same slot. If hypothesis 2 were supported, the sequence *but yet* would not be expected. Lenker (2010: 16) cites a combination of 3 contrastives: *but yet nevertheless*. There are 32 examples of this combination in EEBO, among them:

(40) the kings anger was much appeased; **but yet nevertheless** blank charters were brought into the city. (1682 Gough, *Londinum triumphans* [EEBO])

Fraser and Traugott (2017) found several examples of *and furthermore* (2 Elaborators) and *but yet* (2 Contrastives) in COCA. Such examples where combinations of the same type of DM occur are evidence against the hypothesis that slots have meanings on the assumption that there is one meaning per slot. In Example (40) slots 1, 2 and 3 are each filled with a Contrastive. This gives even stronger evidence than the combinations investigated here that slots are not meaningful.

10.7 Conclusion

Which DSM may combine with which other DSM and how frequently appears to depend on the individual DSM and its function. Lohmann and Koops (2016: 440) suggested that for what I call the "Schiffrin set" of DMs, "each DM is characterized by its own combinatory preferences". A more-usage-based way of phrasing this is that a SP/W's knowledge of DSMs includes knowledge of which combinations are conventional. One would expect that, because DSMs have conventionalized

meanings, if 2 DSMs from the same functional set are combined, they would be (a) compatible in meaning, (b) understood slightly differently, and this is borne out by the data. The data investigated for this study support Oates's (2000) finding that when DSMs from the same functional set collocate, the first is weaker than the second, as in the case of *but yet nevertheless* above, or *and furthermore*. Given the expectation that communication will be coherent, it is not surprising that incoherent combinations appear to be avoided. One would not expect *yet* to combine with *now then* since contrastive and inferential do not cohere well semantically, and indeed the DSM combination *yet* + *now then* is not attested in EEBO,[61] COHA or COCA. Absence of a particular combination does of course not imply that it could never have been used, or that it cannot be constructed, only that it is unlikely in the kinds of texts represented in the corpora.

This concludes Part II and the empirical evidence for some of the changes in the development of Discourse Structuring Markers in English. In Part III I turn to the question how to think about three open theoretical issues for Diachronic Construction Grammar: (i) subjectification and intersubjectification in the context of constructionalization (Chapter 11), (ii) whether the clausal positions in which the Connectors can be used are constructions (Chapter 12) and (iii) the concept of networks and nodes (Chapter 13).

61. The string appears once in EEBO:

(i) thou hadst much wealth; *yet now then* i no more. (1635 Swan, *Speculum Mundiae*)

However, this is not an example of the sequence in question as *then* in (i) functions not as inferential 'therefore', but as the standard of comparison that is usually spelled *than* ('yet now you have no more than I do').

PART III

Three open issues for a historical constructionalist perspective on pragmatics

CHAPTER 11

Subjectification, intersubjectification and the rise of DSMs

11.1 Introduction

Quirk et al. (1985: 440) define Conjuncts as adverbial linkers that "express the speaker's assessment of the relation between two linguistic units". I have analyzed DSMs as procedurals that do more than express SP/W's assessment of the relationship between D1 and D2 adopting Hansen's characterization of procedurals as meanings that "provide instructions to hearers on how the conceptual meanings expressed in an utterance should be combined and processed" (Hansen 2008: 20). From a usage perspective, DSM procedurals are *E*s that SP/W uses to guide AD/R's interpretation of the relationship between D1 and D2 (see Chapter 2.5). The hypothesis developed in the present Chapter is that because sources of Conjuncts are typically CircAdvs, they are usually relatively neutral with respect to SP/W assessments. However, when a SP/W uses a CircAdv as a [[Conjunct] ↔ [DSM]] the source *E* will by default be used in a way that is both more subjective (Speaker-oriented) and more intersubjective (Addressee-oriented) than the source. It will also by default be used in a way that is more "textual" than its source because Conjuncts are Connectors that link D1 and D2.

Subjectification and intersubjectification were briefly introduced in Chapter 3.4, where I mentioned that in previous work I have thought of them as types of mechanism of change. But my current constructionalist perspective suggests that this is mistaken. They are not 'hows' of innovation that can lead to change like neoanalysis and analogization. They are processes that lead to orientation of meanings toward grounding in the Speaker-Addressee dyad. They lead to conventionalized aspects of what Verhagen (1995: 125) calls "communicative function". This kind of function is not included in Croft's (2001: 18) model of a construction, but is consistent with his view of the importance of bringing together communication and cognition (Croft 2001: 364). Communicative function is pertinent to a wide range of constructions, not only DSMs but also procedurals in general and illocutionary uses of speech act verbs. I therefore propose to add communicative function to the meaning characteristics of Croft's model (abbreviated as CF). If a construction is a unit of knowledge in the constructicon, that knowledge must include aspects of

communicative function that are coded, for example, in procedurals. A low (weak) degree of (inter)subjectivity is characteristic of all procedurals since procedurals by definition cue AD/R to the relationship that SP/W is establishing between constructions, be they case relationships, aspectual or modal relationships, or discourse connecting relationships (see Section 11.4). In the case of [[Conjunct] ↔ [DSM]], SP/W cues the relationship between D1 and D2 for AD/R.

In my view, it is important to distinguish subjectivity and intersubjectivity, which are synchronic notions, from subjectification and intersubjectification, which are diachronic notions. They intersect because language use is dynamic and because markers of (inter)subjectivity typically arise as the result of (inter)subjectification.

In Section 11.2 I outline how the terms subjectivity and intersubjectivity have been used, and turn to subjectification and intersubjectification in Section 11.3. In particular, I consider how subjectification and intersubjectification have been conceptualized as being on a directional trajectory that includes "textual" meaning, and question the validity of that concept. Section 11.4 focuses on (inter)subjectification and textualization as evidenced by the development of DSMs exemplified by Digressives. Consequences for assignment of default processes at the time of constructionalization not only of DSMs but of procedurals in general are outlined in Section 11.5. Section 11.6 summarizes.

11.2 Characterizing subjectivity and intersubjectivity

All language use is both:

1. subjective to some degree: SP/W chooses what to say
2. intersubjective to some degree: SP/W chooses whom to address and how (politely, aggressively) and AD/R interprets.

In other words, subjectivity and intersubjectivity are ambient in the context of verbal interaction. This was pointed out by Benveniste (1971[1958]) over sixty years ago. Negotiation of meaning always involves some degree of intersubjectivity, see especially Verhagen (2005: 4) who regards "tailoring" of points of view with respect to other interlocutors (also referred to as "building common ground") as foundational to language.

The topic of subjectivity goes back in European linguistics at least to Bréal (1964[1900]). It has a rich history in Japanese linguistics. Shinzato (2014) points out that in the 1950s the Japanese linguist Minoru Watanabe distinguished subjectivity and intersubjectivity much as Benveniste (1971[1958]) did: SP speaks, AD interprets. In the 1980s attention was paid to the ways in which subjectivity arises

in the act of speaking in a speech event and how this is encoded linguistically, for example in deictics (Lyons 1982). Subjectivity soon came to be an essential topic in work on semantics.

Langacker (1987, 1990) and elsewhere proposed that subjectivity is fundamentally a matter of construal:

> The speaker (or hearer) by choosing appropriate focal "settings" and structuring a scene in a specific manner, establishes a construal relationship between himself and the scene so structured. (Langacker 1987: 128)

The definition of subjectivity in the glossary to Langacker (1987) reads: "An entity is subjective to the extent that its role as observer is maximized, and its role as object is minimized" (p. 493). The idea is exemplified by various scenarios, of which (1) is particularly well-known (Langacker 1990: 20):

(1) a. Vanessa is sitting across the table from me.
b. Vanessa is sitting across the table.

In (1a) SP/W puts herself on stage and refers to herself as "an objectively-construed participant" (Langacker 1990: 20), but in (1b) SP/W is "off-stage"; her role as object is minimized, and she is construing herself in the scene subjectively. For Langacker this construal is a case of "subjectification". He defines subjectification as: "the realignment of some relationship from the objective axis to the subjective axis" (Langacker 1990: 17). For him subjectification is a synchronic notion. The definition as realignment from the objective to the subjective axis appears to assume that the objective axis is basic. This kind of perspective on subjectification is elaborated in Athanasiadou et al. (2006).

11.3 Characterizing subjectification and intersubjectification

By contrast, I consider subjectification and intersubjectification to be historical concepts. They are processes that over time enable the development of conventionalized overt expressions of subjectivity and intersubjectivity. López-Causo (2010) provides a valuable comparison of Langacker's and Traugott's interpretations of the two concepts.

My initial proposal was that there was a unidirectional path of the type in (2) (Traugott 1982: 253):

(2) propositional > textual > expressive

In formulating (2) I drew on Halliday and Hasan's (1976) distinctions between "ideational" and "interpersonal" functions, but chose different terms: "propositional" to draw attention to contentful meaning, and "expressive" to draw attention to the evaluative aspects of negotiated meaning. By "textual" I meant 'pertaining to the construction of text', as in the development of *while* as a concessive subordinator and of demonstrative *that* as the article *the*. But "textual" began to be interpreted in a number of different ways, including text-creating during the subjective act of speaking (see e.g. Breban 2006: 246), the term came to be ambiguous. Therefore I later characterized subjectification as follows: "meanings tend to become increasingly based in the speaker's subjective belief-state/attitude toward the proposition" (Traugott and Dasher 2002: 95), and intersubjectification as: "the development of meanings that encode speaker/writers' attention to the cognitive stances and social identities of addressees" (Traugott 2003: 124). These kinds of perspectives on subjectification and intersubjectification are elaborated in Davidse et al. (2010). It may be noted that in rethinking (inter)subjectification, "textual" was not forgotten, but came to be incorporated within Tendencies toward greater subjectivity on the one hand and toward greater intersubjectivity on the other (see Narrog 2014[2012] for discussion).

Here I mention three different proposals regarding types of (inter)subjectivity and (inter)subjectification, the first synchronic (De Smet and Verstraete 2006), the second both synchronic and diachronic (Brems et al. 2014[2012]a) and the third diachronic (Narrog 2014[2012]). All take up some tripartite version of (2) above. All regard subjectivity and intersubjectivity as a gradient, "less to more" phenomenon. I follow these proposals with my current diachronic constructionalist hypothesis.

De Smet and Verstraete (2006) propose three types of subjectivity, one pragmatic and two semantic (p. 384–388), adopting Halliday and Hasan's (1976) distinction between ideational and interpersonal meanings:

a. pragmatic subjectivity; inherent in language use and "speaker's conceptualization and reasons for choosing" a particular *E* (De Smet and Verstraete 2006: 384, 387). This is what I call "ambient subjectivity" above.
b. ideational subjectivity: evaluative senses of words, e.g. *pleasant, cursed.*
c. interpersonal subjectivity: speaker-positioning and interaction with the interlocutor, e.g. *as, since, for*. As De Smet and Verstraete say, this dual perspective, involving both positioning and interaction is reminiscent of the distinction made in Traugott and Dasher (2002) between subjectivity and intersubjectivity (p. 387).

Drawing on the example of *because* used to express what Sweetser called sociophysical cause, and comparing it with *as, since, for*, De Smet and Verstraete

suggest (p. 383) that criteria for subjectivity include unavailability of the subjective Connectors for focus clefting (*It was because*…, but * *it was as*), *wh*-questioning and negation. These criteria are the ones usually used to identify pragmatic Connectors (see Chapter 4.4.1) and point to the interconnectedness of conventionalized pragmatics and subjectivity of DSMs that I explore further below.

Among important points in De Smet and Verstraete's (2006) proposal is that interpersonal meaning involves the speaking subject's "enactment" of his or her position with respect to the content, in other words, it has an element of illocutionary, speech act force. This is essential to the definition of "procedural" meaning that I have used throughout this book: SP/W cues AD/R about the relationship between D1 and D2 intended. Often this cueing is unconscious, but sometimes it can be highly conscious, as in the case of OBTW2, where "pseudo-enactment" is key (see Chapter 10.4). Another important point is that causals like *because, since* are both intersubjective and subjective in that they cue the AD/R to interpret relationships between events (see also Ghesquière 2010).

Brems et al. (2014[2012]b) investigate intersubjectivity (and to some extent intersubjectification) in terms of "joint attention" (Diessel 2006). They propose distinctions between three types of intersubjectivity (p. 147):

a. attitudinal: meanings "code the speaker's image of (his/her relation to) the hearer, and attention to the face needs and social self of the hearer", cf. hedges,
b. responsive: "elicitation of a certain (speech) act or behavior on the part of the hearer", cf. turn-taking,
c. textual: "meanings that are specifically oriented toward steering the hearer's interpretation, e.g. focus and backgrounding.

In this Chapter I am concerned mainly with (a) and (c). Brems et al. (2014[2012]b) suggest that formal properties of intersubjectivity may be found in:

1. linearization, or position relative to the clause (see Chapter 12),
2. prosody, which is typically linked to linearization (e.g. a rise clause-finally in English typically invites response to a question),
3. directionality (here, despite reservations, they adopt non-subjective > subjective > intersubjective as proposed in Traugott 2003: 134).

While these properties are useful cues, they are not reliable principles for operationalization, as the authors readily admit. Much depends on the theoretical perspective adopted and the data investigated.

Narrog (2014[2012]) draws attention to the importance of what he calls "text/discourse orientation". The term "text/discourse" combines the concept of markers of textual relationships in the narrow sense (e.g. DSMs) with rhetorical moves,

which are "textual" in the wider sense (see Breban 2006 for discussion). With this acknowledgement of the duality of the meaning of "textual", Narrog (2014[2012]: 48) proposes (3) in place of (2):

(3) (subjectification) > intersubjectification > text/discourse orientation

An example that he gives of intersubjectification without prior subjectification is that of an imperative coordinated with a declarative clause, which together function as a conditional, as in (4) (Narrog 2014[2012]: 41, citing Quirk et al. 1985: 832):

(4) Make a move and I'll shoot.

Of this construction Narrog (2014[2012]: 41) says:

> The imperative is clearly intersubjective in being directed at the addressee. If it is used to relate two propositions and imply conditionality between them, corresponding to a connective, it acquires a textual (logical function)

A problem here is that while the imperative is indeed clearly intersubjective, and it is clearly being used "textually" in (4), there is no evidence that it came to be intersubjectified. It has had intersubjective function associated by default with imperative constructions from Indo-European on.

The proposals in (2) and (3) are made on the assumptions that:

a. subjectification, intersubjectification and text-marking constitute a set that can be ordered,
b. it is possible to find directional pathways for the development of this set of changes.

In the case of the hypothesis in (2) above, it was made in the context of work influenced by grammaticalization, an approach to change in which unidirectionality in the strict sense was at the time a major focus of theoretical interest (see Chapter 5.2). The examples I worked on suggested that there was a tendency for meanings that "once subjectified, may be recruited to encode meanings centered on the addressee" (Traugott 2010: 35).

Ghesquière (2010) rightly questioned the ordering hypothesis in (2) on grounds similar to those of De Smet and Verstraete (2006), that causals like *because, since* are both intersubjective and subjective because they cue the AD/R to interpret relationships between events. One does not follow the other.

At this point I propose to keep the rather broad characterizations of subjectification and intersubjectification that I put forward some time ago (5) and (6) respectively, and to add a third process, textualization, as in (7). Like Narrog

(2014[2012]: 33), I consider them to be on "equal standing" for reasons to be explained below.

(5) **Subjectification** is increase in the degree to which SP/Ws overtly base meanings in and orient them toward their own perspective.

(6) **Intersubjectification** is increase in the degree to which SP/Ws overtly pay attention to AD/Rs and orient meanings toward AD/Rs' cognitive stances and social identities.

(7) **Textualization** is increase in the degree to which SP/Ws pay overt attention to text-creation and invite AD/R to interpret textual relationships.

Note that "overt" is key to the three characterizations. The processes lead to externalized linguistic markers/indices of subjectivity, intersubjectivity, and textuality. Note also that the three meaning-shifts proposed here are not mechanisms but processes by which meanings are used and reinterpreted over time. The first two processes, subjectification and intersubjectification, are closely intertwined in language use because SP/W's use language to communicate with AD/Rs. In the process of communicating, SP/Ws create text, as Breban (2006: 246) points out, but in my view that is to be distinguished from using *E*s to specify particular textual meanings. Therefore, although textualization is closely linked to subjectification and intersubjectification, it serves a separate type of communicative function. I currently think that trying to position it in a directional trajectory such as (2) was a mistake, even though it turned out to prompt important thinking about the role of (inter) subjectification in the development of modal functions (Narrog 2014[2012]) and text-creation (Breban 2006).

In Traugott (2010: 35) I proposed that in subjectification "meanings are recruited by the speaker to encode and regulate attitudes and beliefs". This overtly states that a SP (or W) initiates change, a hypothesis that Hansen (2012) challenged, as discussed in Chapter 3.4. Hansen proposed that when a pair of interlocutors negotiate meaning, the hearer is interested in the speaker's point of view and is "predisposed to reinterpret linguistic items as expressing precisely that point of view" (Hansen 2012: 602–603), therefore AD/R is as likely or more likely than SP/W to subjectify certain non-subjective expressions. I agree that AD/Rs are important in change, but AD/Rs' interpretations are in evidence only when AD/Rs produce new uses as SP/Ws, therefore in my view it is SP/Ws who initiate change (see Chapter 3.3.2). However, it is often the case that a SP/W A who in a certain context initially invited the inference of a more subjective meaning for *E1* is not the same SP/W who innovates the more subjective use *E2*. Rather it is SP/W B who innovates and attests *E2*.

11.4 The relationship of textualization and (inter)subjectification in the development of DSMs

Because they cue the relationship between D1 and D2, DSMs are textual markers by definition. This section starts with some generalizations about the relationship of textualization with subjectification and intersubjectification as DSMs come into being (11.4.1). In Section 11.4.2 aspects of the development of Digressives are reprised as a brief example of the issues that are the topic of this Chapter.

11.4.1 Some generalizations

By hypothesis textualization, subjectification and intersubjectification occur simultaneously when a new [[Conjunct] ↔ [DSM] is innovated. This is consistent with Narrog's (2014[2012]: 41) conclusion that "intersubjectification must have accompanied subjectification" when modal *may* was used with concessive meaning as in (8):

(8) Whatever one **may** think of Kenneth Williams, I thought he did a very good rendering. (*London Lund Corpus* [Narrog 2014[2012]: 36)

As innovated uses came to be replicated and conventionalized, the textuality and (inter)subjectivity come to be communicative functional properties of the constructions with which they are associated. Therefore, by hypothesis, when an adverbial comes to be conventionalized and constructionalized as a [[Conjunct] ↔ [DSM]], textualization, subjectification and intersubjectification of the original CircAdv can be said to have occurred. Further subjectification or intersubjectification may occur later, as illustrated by *by the way* below. However, I have not found any evidence of increased textuality in the development of the DSMs I have studied. It is possible that in other domains there might be increase in textuality, given certain assumptions such as that subordination is more grammatical and therefore more textual than coordination (Givón 2018[1979]]: 152), but that is beyond the scope of this study.

11.4.2 Digressive markers, textualization, subjectification and intersubjectification

When used as CircAdvs *by the way* and *by the by* are relatively objective phrases that refer to real world spatial relations. However, as a manner adverbial, *incidentally* involves some subjective assessment of how an event unfolds. When used as [[Conjunct] ↔ [DSM]]s, they are both used with functions that are more textual

than their lexical sources. They are also used in a relatively more subjective way because SP/W signals the relationship to be understood ('I, SP/W, link upcoming D2 to preceding D1 in passing'). They are at the same time used more intersubjectively because SP/W cues that AD/R is to understand the nature of the link between D1 and D2.

As a Conjunct, *incidentally* came to be more subjective pragmatically than it was as a manner adverbial. *Incidentally* was further subjectified in uses that suggest that stance-to-text marker status is emerging, see *And, incidentally, it may be well to mention here one positive benefit* (1870), cited in Chapter 8.3.2 in Example (23).

By the way has a far more interesting history with respect to (inter)subjectification. Recall from Chapter 8.2 that *by the way* was originally used as a spatial CircAdv meaning 'along the way, on the road'. This has no textual or particular subjective or intersubjective meaning. Contextual uses, especially uses in contexts of locutionary verbs like *say, recount, tell, touch on* appear to have led interlocutors to associate *by the way* with talk and text and with the metaphor ARGUMENT IS A JOURNEY. By the early 1600s *by the way* appears in EEBO used to mean 'in passing'. This is a textual use that is subjective ('I writer add this point and present it as not very important'), and intersubjective in the sense that AD/R is invited to understand the text as a kind of footnote. Example (6a) in Chapter 8.2 is repeated here for convenience as (9).

(9) sent him a plaine and cold answer, wash seven times in iordan: **by the way**, I dare boldly say, elisha in himselfe was not proud at other times,
(1640 Fuller, *Ioseph's partie-colored coate* [EEBO])

Soon after we find examples of *by the way* used as a DM that reflects (inter)subjective 'as if casual' stance. Example (6b) in Chapter 8.2 and repeated here as (10) includes a direct address to the reader to pay attention to D2, which is an important generalization but is presented as if it was not so important:

(10) which city is now dwindled to nothing: reader, **by the way**, I observe that cities surnamed the great, come to little at least.
(1662 Thomas, *The history of the worthies of England* [EEBO])

This kind of use came to be conventionalized by the 19thC. It can be assumed to be more subjective than earlier uses such as in (9) because there is an added stance of "as if" unimportantce. *By the way* is no longer used in contexts that overtly mention locution such as *observe* 'note' in the case of (10). However, the alleged thought processes and actions prior to its use may be hinted at or even spelled out in narrative comments on the represented speaker, as in Example (8b) in Chapter 8.2, repeated here in part as (11):

(11) When I paid for my breakfast I half turned away, then turned back casually. "**Oh, by the way**," I said. "Where's this wall TV place?"
(1959 Graham, *The Gallery* [COHA])

Use of *by the way* as a hedge on indirect directives involves increased intersubjectification as in (11) above and in Examples (10) and (11) in Chapter 8.2 (*By the way, you must take lessons in Milford* (10a); *By the way, Tuppy, which is it?* (11a), *By the way, did I hear, or did I not, that…?* (11b)). Such examples show that *by the way* could be used by the end of the 19thC in a considerably more intersubjective way than in the 17th century, when spatial *by the way* was constructionalized as [[Conjunct] ↔ [$DSM_{Digression}$]].

Turning now to *Oh, by the way*, which was outlined in Chapter 10.4, *Oh* in OBTW1 adds an element of both subjectivity and intersubjectivity because it marks change-of-state and is attention-drawing. OBTW2 used as a conceptual unit is highly subjective and evaluative. It trivializes what a third person has been perceived or alleged to do or say and to disaffiliate SP/W from it. It is used to enact a pseudo-representation of what someone else is purported to have said. It has been further subjectified. Unlike the earlier OBTW1, in COHA and COCA OBTW2 is not addressed directly to a second person and is not used to begin a turn. Therefore it does not immediately meet the criterion of intersubjectification as increase in the degree to which SP/W is overtly invited to pay attention to AD/Rs cognitive stances and social identities. However, hearers and readers (as opposed to a direct addressee) of OBTW2 expressions are invited to process D2 and evaluate what might have actually been said in a case like (12), repeated from Chapter 10.4.2, Example (37a):

(12) they had about 90 percent of the deal done, and then China said, **oh, by the way**, this idea of intellectual property theft and forced technology transfer that we've agreed to stop doing, now we're going to stop doing that.
(2019 *Fox – Ingraham* [COCA]; China is alleged to have said they will stop stopping intellectual property theft, etc.)

In sum, from the perspective of constructionalization, the development of *by the way* and other DSMs is a case of textualization, weak subjectification and intersubjectification, all communicative functions (CFs). These occur along with constructionalization and may be followed in individual cases by further subjectification and intersubjectification. Once the processes have occurred, they result in the constructional properties textuality, subjectivity and intersubjectivity. Figure 11.1 specifies for *by the way* those features relevant to this Chapter that are results of the processes: Conjunct, resulting from textualization, subjectivity ('subj') resulting from subjectification, and intersubjectivity ('intersubj') resulting from intersubjectification. "Str.

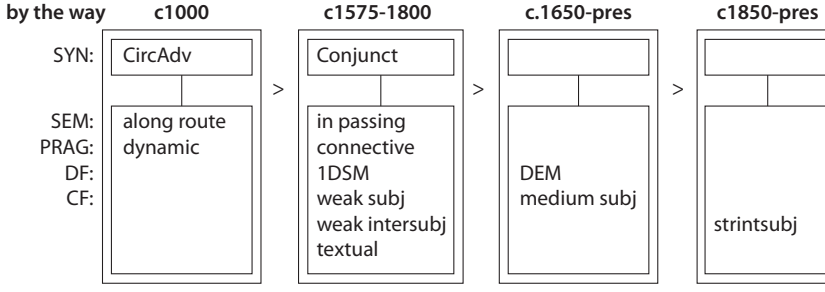

Figure 11.1 Partial model of the development of DM *by the way* showing the outcome of textualization, subjectification and intersubjectification

intsubj" is short for 'stronger intersubjectivity'. Figure 11.1 builds on Figure 8.1 in Chapter 8.2, but for ease of reading does not repeat some details in that Figure that are specific to *by the way*.

Does all this actually need to be specified for *by the way* and every new DSM? No. By hypothesis any new [[Conjunct] ↔ [DSM]] will have undergone textualization, weak subjectification, and weak intersubjectification. This is a default of the functional change and the task to which the new expression is put. What does need to be specified for *by the way* is the construction-specific constructional changes when it became a DM and when it became a hedge.

11.5 Default features of a DSM construction

I propose that when a CircAdv is constructionalized as a [[Conjunct] ↔ [DSM]], it will come to inherit the characteristics in Figure 11.2 by default all at the same time:

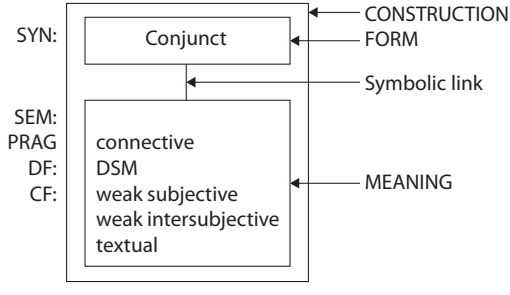

Figure 11.2 Default features inherited at the time of constructionalization as a [[Conjunct] ↔ [DSM]]

Thinking more broadly, since procedurals in general are cues to AD/R to interpret a particular relationship between constructions, whether pragmatic social or epistemic markers, or markers of aspect, modality, case, or comparison, all procedurals may be hypothesized to inherit the communicative functions of weak subjectivity and weak intersubjectivity at the time of constructionalization. Most will not inherit the textual function in the narrow sense of marking text creation used here. The textual function is specific to Connectives, whether coordinate or subordinate.

11.6 Summary

I have proposed the following characterizations which were given in (5)–(7) above in Section 11.3, repeated here as (13)–(15):

(13) **Subjectification** is increase in the degree to which SP/Ws overtly base meanings in and orient them toward their own perspective.

(14) **Intersubjectification** is increase in the degree to which SP/Ws overtly pay attention to AD/Rs and orient meanings toward AD/Rs' cognitive stances and social identities.

(15) **Textualization** is increase in the degree to which SP/W pays overt attention to text-creation and invites AD/R to interpret textual relationships.

All three are grounded in the SP/W – AD/R dyad. The outcomes of all three processes shape language-users' understanding and use of conventionalized constructions. I have analyzed them as properties of communicative function, which I suggested in Section 11.1 should be considered to be among the meaning properties of a construction.

I have argued that all three are involved in the constructionalization of a [[Conjunct] ↔ [DSM]] pairing at the same time, and have suggested that attempts to interpret an order of development among the three is mistaken.

I have also suggested that weak subjectification and weak intersubjectification are characteristic of the constructionalization of all procedurals.

How exactly to operationalize degree of (inter)subjectivity or of (inter)subjectification independently of particular constructions or constructional schemas, has yet to be established.

In the next Chapter I address the question whether the pre-clausal, post-clausal and clause-medial slots in which DSMs can be used are meaningful and to what extent those slots are associated with subjective and intersubjective points of view.

CHAPTER 12

Clausal positions of DMs

12.1 Introduction

Sequential position in DSM (mainly DM) combinations was the topic of Chapter 10. There I suggested that the position of one slot relative to another when DSMs are combined is not meaningful. However, there is a constraint that DSMs with weaker pragmatics precede those with stronger pragmatics. This is probably part of a larger discursive constraint that accounts for preference for modal combinations such as *may and should* over *should and may*, where the deontically weaker modal precedes the stronger one. The focus of the present Chapter is position of the Connector. Cxn relative to the clause D2. Again I will pay most attention to DMs. So far I have largely discussed [[Conjunct] ↔ [DSM]] Connectors in pre-clausal position in the complex subschema [D1 Connector.Cxn D2]. Clause-initial position is well established as a position where discourse framing can be done (see Schiffrin 1987; Aijmer 2002; Sarda et al. 2014; Haselow 2019, among many others). 1DSMs may occur in a variety of positions in the clause, but only multifunctional DMs are likely to have different pragmatic and discourse effects in different positions. In this Chapter I pose the question whether position relative to a clause is a construction, more specifically: "Is the slot in which a DM may be used in English relative to a host clause a construction or form-meaning pairing?"[62] I will answer in the negative. Because the particular pragmatic meaning of a DM may be position-dependent, it does not necessarily follow that position is a construction. Since a construction is an [[F] ↔ [M]] pairing, neither discourse nor communicative function alone qualify position as a construction. To qualify as a construction, position would have to be a slot with a predictable function.

I take "position" to refer to slots in the linear production of text found in the corpora. In the history of English from the first written records on, adverbials have been used in different positions and with different functions. As discussed in some detail in Lenker (2014), especially with respect to the Conjuncts *however* and *therefore*, changes in word order patterns since Old English have affected where and how Conjuncts are used. However, since most of my data concern changes from Middle

62. Many thanks to Noriko Onodera for asking me to think about this question several years ago.

English on when word order had become relatively fixed, variable positions in Old English will not be discussed here.

Since the corpora are not conversational, I will not be considering the question whether position in a set of turns, e.g. first, second, third turn, is meaningful. Important synchronic work on this question is to be found in Heritage (1994, 2002), Deppermann and Günther (2015) and Haselow (2019), among others. It has been suggestive and helpful for discussion of position in this Chapter, although I am not directly concerned with either turn-taking or synchrony.

The question whether position of a DM relative to the clause is a construction assumes that some DMs that can be used as Connectors in pre-clausal position can also be used as Connectors with a linking function in clause-medial and post-clausal positions, e.g. *after all, all the same, by the way*. As mentioned in Chapter 4.3.1, I assume that there is a Connective.Schema [[D1 Connector. Cxn D2] ↔ [signals a largely symmetric link between D1 and D2]]. In this model, pre-clausal position is the default position for Connectors, because it is specified as preceding D2. Some micro-constructions that are DSMs occur in this position only with a clause connecting function (e.g. *and*, and, until recently in most varieties of English, *but*), others in multiple positions in the clause (e.g. *after all, by the way*). As the case studies have suggested, clause connecting functions qualify microconstructions that are DSMs in pre-clausal position to be analyzed as DSMs in other positions as well. This means that they can be used as Connectors in all positions (see also Lenker 2014).

Three hypotheses regarding whether position relative to the clause is a construction can be put forward:

> *Hypothesis 1*: There is one positional construction. It is pre-clausal (form) and it signals the kind of discourse relationships between D1 and D2 to be inferred (meaning). The possibility that a 1DSM can be used in other positions without pragmatic modulation is a specific feature of the 1DSM in question, as is the possibility that a DM may be associated with rather different pragmatics in different positions.
> *Hypothesis 2*: There are three distinct positional constructions, pre-clausal, post-clausal and clause-medial. Each is associated with a different discourse function.
> *Hypothesis 3*: Position relative to the clause is not a construction. Rather, whether a DSM can be used in any position other than pre-clausal, and whether there are functional differences in different positions is construction-specific.

Hypothesis 1, that there is one positional construction, is consistent with Fraser's view of DMs, which is that they occur in patterns of the type Segment 1 – DM – Segment 2 (see Fraser 2009b: 297) and the constructional representation cited

above for the Connective.Schema. But Hypothesis 1 would require a uniquely underspecified positional micro-construction on the meaning side that is redundant with the [D1 Connector.Cxn D2] construction. It would predict that pre-clausal position is the default, therefore most frequently used position for all DSMs and that uses in other positions override this default. However, it is not always the case that use in pre-causal position is the default: *all the same* is preferred in post-clausal position (see Chapter 7.3).

Hypothesis 2, that there are three distinct positional constructions, is not supported by the data I have investigated. For one, some Connectors, e.g. *after all*, can have different functions in the same position. For another, while post-clausal position is associated with concessive meaning for Contrastives, it is not associated with Digressives, e.g. *by the way*. The conclusion to be drawn is that position is not a construction (in English) as positions do not predict the meaning of a [[Conjunct] ↔ [DSM]].

Hypothesis 3, that individual DSMs (and PM microconstructions more generally) may have specific functions in particular positions, is well supported. These variations are part of what we know about a Connector, particularly a DM Connector, when we have learned it. In this Chapter I argue that which positions are available for a particular DSM is Connector-specific. Whether there is a difference in meaning in the different positions is also Connector-specific.[63]

Because textual traditions were in the early periods heavily influenced by Latin and French, it is highly likely that these traditions and the many translations from Latin and French influenced use in particular positions in English. In her study of several causal and adversative Latin "particles", Kroon discusses *autem, vero*, and *at*, which prior to her work were usually taken as equivalents of 'but'-relations (Kroon 1995: 1), but which she analyzes as significantly distinct markers of discourse shifts. She notes the tendency for *autem* to occur in "second position"[64] after the first

63. Jiménez et al. (2018) have proposed that in the Val.Es.Co model of discourse (Briz and Grupo Va.Es.Co. 2003), discourse units (e.g. interventions, acts and subacts) can be "incorporated into position" and that "the function of a discourse marker depends on its structural position, the discourse unit over which it has scope and the meaning of the DM itself" (Jiménez et al. 2018: 122). Jimenéz et al.'s study focuses on hypotheses that have been put forward about the relationship of (inter)subjectivity to position (e.g. Beeching and Detges 2014b), but the findings can be extended to other issues related to position. Whether there is the potential in this work for concluding that position has a function and would therefore qualify as a construction remains to be determined based on more data and more DSM types.

64. This second position phenomenon is associated especially with Indo-European and was formulated as "Wackernagel's law" in 1892 (see Goldstein 2014). The phenomenon is now referred to as V2 syntax, see Section 12.2 immediately below.

constituent that is a "fronted" given or accessible topic and associates this tendency with enhancing the coherence of the text (p. 257):

(1) tu eum orato, ego **autem** orabo vilicum.
You with-him plead, I however will-plead-with bailiff
'you plead with him, while I plead with the bailiff'.
(Pl. Ca. 273 [Kroon: 1995: 275])

Unlike *autem*, the markers *vero* and *at* tend to occur in initial position. *Vero* is primarily modal rather than adversative in function, and is used to indicate "the *actuality* or *reality* of … a communicative event or an event in the represented world" (Kroon 1995: 325, italics original). As for *at*, it signals that "in the upcoming unit certain expectations about the unmarked continuation of the discourse are being frustrated" (p. 363). The differences in word order associated with *autem* (2nd position), *vero* and *at* (pre-clausal position) is likely to have influenced the position in which DSMs and their CircAdv sources were used. While the possible influence of Latin and French word order on translations should not be underestimated, it is beyond the purview of the present study.

The proposal that clause position is not a construction poses a challenge to Goldberg's much cited aphorism *it's constructions all the way down* (e.g. Goldberg 2003: 223). In a squib asking "What would it take for us to abandon Construction Grammar?" Hoffmann (2020) rightly asks us to think what "it's constructions all the way down" means. A theory should be testable and falsifiable. Are there really no parts of a language that are not constructions? Position is a good testing-ground for the hypothesis. I will come back to this point in the conclusion, Section 12.5.

It should be noted that the conclusions in this Chapter pertain specifically to English DMs as defined here in written, largely monologal texts. A different conclusion might be reached with respect to different pragmatic markers in different languages and to turns in conversation. For example, analyzing "discourse particles" in German, Fischer (2017: 337) argues with respect to use of *ja* 'yes' in turn final position to seek confirmation that "it is the structural position (at the end of an assertive utterance produced with rising intonation) that carries this meaning". While it is clear that the meaning does not lie in the particle itself (Fischer 2017: 337), it is possible that it is the intonation that carries the meaning in the case of turn-final *ja* with rising intonation, rather than position.

The structure of the Chapter is as follows. The main positions with respect to the clausal host are discussed in 12.2: pre-clausal position (12.2.1), post-clausal position (12.2.2), and clause-medial position (12.2.3). A hypothesis regarding correlations between position and (inter)subjectification is discussed in Section 2.3. In Section 12.4 *after all* and *by the way* are revisited to exemplify different kinds of discourse function with respect to host clauses. Section 12.5 concludes.

12.2 The main positions with respect to the clausal host

If, as is true of many of the examples in this book, the historical source of a DSM is a CircAdv, that source adverbial will be found in the textual record in the normal position for such adverbials. In Old English this can be initial position. For example, in Old English *þa* 'then' and to a lesser degree *þonne* 'then' were typically used clause-initially, where they triggered what is called "V2 syntax" (see e.g. Los and Kemenade 2012). V2 syntax is characterized by an adverbial such as *þa* 'then', *ne* 'not' or a *wh*-word in initial position followed by a finite verb and then the subject in main clauses. V2 syntax for the most part obsolesced by 1400 except with *wh*-words, a pattern which was later restricted to auxiliaries (*Why did you go*, not *Why went you*?) and some negative adverbs (cf. *Never had she wanted a job so badly*).[65] From Middle English on, in many cases the source adverbial is typically found in the normal position for CircAdvs in general. In PDE locationals, temporals, and manner adverbials are usually found clause-finally in main clauses. But, as has been illustrated in the case studies, a CircAdv can be topicalized and can be used with a framing function in initial position (see Chapter 4.4 and 12.2.1 below). By hypothesis replicated use in this position is a necessary (but not sufficient) context for a shift from CircAdv to Conjunct use.

While most studies of DMs have considered only or mainly use in initial position, there has been increasing interest in the last decade in post-clausal and clause-medial use. Lenker (2010: 200) and Haselow (2019) have suggested that in English post-clausal position may be associated with concessive function and with retrospective contrastive use. Lenker (2014) has investigated medial position as a locus for metatextual markers like *however* and *therefore*.

12.2.1 Pre-clausal position

What I have referred to as pre-clausal position has been called "pre-front field" in Auer (1996). This position is "outside" the clause and has proved problematic for some syntactic theories. Aijmer (2002: 29) comments that Auer finds that "[i]ts loose syntactic attachment makes the pre-front field interactionally and textually

---

65. From the 16thC on spatial adverbial – finite verb – subject came to be used pragmatically to present a referent that is discourse-new, e.g.:

(i) From the director of Azur & Asmar comes another visual stunner.
(2012 *IFC Center* [COCA])

This pattern is a subtype of presentational "preposing constructions" (Birner and Ward 1998).

attractive". *Es* used in pre-clausal position are for the most part PMs that have social or epistemic functions, or 1DSMs and DMs.

Although the data used for this book are typically not interactional, it is nevertheless worthwhile to take note of Haselow's (2015, 2019 *et passim*) proposals regarding correlations between communicative tasks and DMs used in turn-taking. Some of the communicative tasks in turn-taking are as relevant to written documents as they are to natural conversation as talk is represented in drama and fiction. Many texts, especially sermons, are intended to be spoken and to be addressed to an audience, although the text may be a monologue. To what extent represented turn-taking in written texts actually matches what is found in conversational data remains to be investigated.

Haselow (2019: 3) suggests that there are three functional domains in which PMs (he calls them DMs) are used: interaction, discourse structure and cognition. He discusses how PMs are used to fulfill these functions at turn-beginnings and turn-endings. Some of his suggestions regarding use at turn-beginnings are modified in Tables 12.1 and 12.2 to enhance thinking about the communicative functions of DSMs in pre-clausal and post-clausal positions. The assumption is that knowledge of micro-constructions in a usage-based model includes knowledge of how to use them in a usage event to fulfil the usual functions of such events. Written documents are written for readers, and are therefore partially interactive, if only to a limited degree (see Chapter 2.3). Table 12.1 concerns likely communicative tasks to which DSMs are put in pre-clausal position.

Table 12.1 Communicative tasks relevant to use of DSMs in pre-clausal position (based on Haselow's 2019: 5 account of tasks at turn-beginning)

Domain	Communicative tasks
INTERACTION	– getting/claiming the attention of the addressee – dealing with topic-shifting issues
DISCOURSE STRUCTURE	– indicating the kind of relations to prior discourse – initiating a new linguistic action
COGNITION	– providing interpretative cues for an upcoming message

DMs used in pre-clausal position, like DMs used in turn-initial positions, tend to be used to link back to prior discourse and forward to upcoming discourse.

An important additional communicative function is framing, as mentioned in Section 12.1 above. This can occur in either initial or pre-clausal position. In English since the thirteenth century, the clause typically begins with a subject that is given or accessible in the text. However, clause-initial position can also be occupied by a focused element such as *at night* (a CircAdv) in (2):

(2) That is when the dreams began, in which I could fix the mechanisms of life as easily as I could machines. **At night** I saw an elaborate tapestry of iridescent threads. (2012 *The Darth side* [COCA BLOG])

As we have seen, this is a position in which an *E* with appropriate meaning, and in the appropriate context (often report of a locutionary event), can be reinterpreted as a pre-causal, linking [[Conjunct] ↔ [DSM]]. In sum, Conjuncts do not occur in clause-initial position; topicalized (focused) adjunct adverbials may do so. Conjuncts occur in pre-clausal position, sometimes ambiguously with CircAdvs.

In general, DSMs used to mark topic shift occur pre-clausally. This follows from the fact that SP/Ws strategically mark what comes next and its coherence with what precedes (Haselow 2019). Combinations of DMs typically occur pre-clausally as well for the same reason. Neither of these observations leads to the conclusion that the pre-clausal slot is a construction independent of the Connector.Cxn and the microconstructions that instantiate it. The only reasonable generalization about the pre-clausal slot is that it hosts pragmatic *E*s, vocatives, and polarity items such as *Yes* and *No*.

12.2.2 Post-clausal position

Post-clausal position has attracted attention only fairly recently, and mainly in the context of discussion of conversational interaction, especially Haselow (2012, 2013, 2016) and Hancil et al. (2015), but see Kemenade (2021) on a structural analysis of the rise of clause-final *then* in minimally interactional texts. Likely communicative tasks associated with post-clausal position are summarized in Table 12.2:

Table 12.2 Communicative tasks relevant to use of DSMs in post-clausal position (based on Haselow's 2019: 5 account of tasks at turn-ending)

Domain	Communicative tasks
INTERACTION	– legitimizing topic transition
DISCOURSE STRUCTURE	– retrospective indication of the kind of relation to prior discourse
COGNITION	– interpretive fine-tuning of a message just produced (e.g. in terms of epistemic value, illocutionary force, canceling possible implicatures)

Just as initial position needs to be distinguished from pre-clausal position, so final position needs to be distinguished from post-clausal position. In PDE most spatial and temporal CircAdvs may occur in final position, where they may be combined (3).

(3) a. I'll straighten you out **in the city tomorrow.** (1998 *Rounders* [COCA])
b. so that he would be there **in the evening if she came**.
(1990 LeGuin, *Bill Weisler* [COCA])

Whether or not a DSM can be used in post-final position is construction-specific. Some DMs cannot be used in post-clausal position, among them *and* and *also* (they are phrasal coordinators in non-pre-clausal positions). This was also true of *but* until recently. Mulder and Thompson (2008) draw attention to the use of final *but* in represented Australian fiction as a stereotype of Australian English, and investigate its use in both American and Australian dialogue as found in the Santa Barbara Corpus of American English (SBCSAE) and the International Corpus of English – Australia (ICE-AUS) respectively as well as other sources such as films (see also Hancil 2016 on use of final *but* in parts of the UK). Mulder and Thompson found that final *but* ends an intonation unit and concludes a turn, as in (4):[66]

(4) [Talking about whether Abbie is Norwegian or not]
Abbie: It's a complicate[ted story.]
Terry [totally]
Maureen: R(h)eally.
Abbie: And the NA:ME is MY NA:ME is Norwegian **but**,
Maureen: What is his [nationality?]
Terry: [Sola]
Abbie: Hungarian (Mulder and Thompson 2008, Example (8))

In some cases there is what Mulder and Thompson call a "Janus *but*", which appears to serve both as the end of a clause and as the beginning of another, as in (5):

(5) MARCI: I don't know what the real story is,
but,
…(1.1) it sounded kinda neat.
(Mulder and Thompson 2008, Example (15))

Here Marci finishes one intonation unit with *is*, follows it by *but*, which suggests it is potentially final *but*, and then after a 1.1 second pause continues in a way that suggests it is pre-clausal. Mulder and Thompson hypothesize that examples like (5), in which *but* is "left hanging", so to speak, represent an intermediary step toward use of final *but* in Australian English. There it can be used as part of the intonation unit with final, falling contour, and means 'though' as in (6):

66. In this and the following examples from Mulder and Thompson, [ ] indicates overlap in speech, '@' laughter, 'H' pitch-accent, and < > drawn out prosody. Numbers in parentheses, e.g. '(1.1)', indicate length of pause in seconds.

(6) [Diana has just made strange noises]
Kylie: You sounded funny @ @(H)
Diana: I know
Sounded like an alright person **but**.
(3.3)
Kylie: <SING On Saturday, SING>
(Mulder and Thompson 2008, Example (18))

Mulder et al. (2009: 357) suggest that clause-final *but* in Australian English is used to mark contrastive content, to yield a turn and to "index 'Australianness'".

A variety of Connectors can be used in post-clausal position with concessive, therefore connective, meaning, among them *after all*, *anyway, then, though*. Concessive use in this position is consistent with the tendency from the 18thC on for post-clausal position to be strongly associated with concessive 'though, despite what might be expected' meaning for Contrastives (Lenker 2010; Haselow 2012, 2013). However, as was discussed in Chapter 4.5.2, in the case of *after all*, use with concessive meaning in post-clausal position has been on the decline and justificational/elaborative meaning has been generalized to post-clausal position for that DM.

In the case of *by the way*, which is not contrastive, concessive meaning is not associated with post-clausal position. Instead, use in this position is usually used to imply that what precedes (D2) is to be taken as a fairly casual, even dismissive, add-on to D1:

(7) And cutting demand means a long-term program to help treat people who are addicted now and who are committing a lot of the crimes, **by the way**.
(1990 *ABC_Nightline* [COCA])

In post-clausal position *by the way* does not appear to have a backward hedging effect. But it can be used to express epistemic certitude and combativeness in the guise of dismissiveness, as in (8):[67]

(8) Big pharma ran millions of dollars of negative advertisements against me during the campaign, which I won, **by the way**.
(2020 President Trump, White House press briefing, November 20th)

In sum, although there is a tendency for post-clausal use of Contrastives to be concessive, post-clausal position does not determine meaning. *E*s that are DMs (or monofunctional DSMs) can be used in post-clausal position with connective function and can therefore be considered to be Connectors in this position.

67. Thanks to Graeme Trousdale for drawing my attention to this example.

12.2.3 Clause-medial position

Whether there is a difference in meaning in the different positions is Connector-specific. So is whether an *E* can be used with Connector function in clause-medial position. For example, use in medial position in PDE after the finite verb is attested for the DMs *by the way* and *after all*, but is less frequent than pre-clausal or even post-clausal use (see Tables 12.4 and 12.5 below in Sections 12.4.1 and 12.4.2 respectively). By contrast, *also* is not used as a Connector in clause-medial position as a Connector. According to Haselow (2015), the same is true of *anyway*.

"Medial" position is a cover-term for a number of possible positions in the clause. For example, Greenbaum (1969: 78) distinguishes 8 medial positions for linking adverbials. This multiplicity of positions is not limited to Conjuncts. For example, Aijmer (2002: 256) shows that the epistemic PM *actually* can be used in conversation at virtually any constituent break (with slightly different meanings) other than between *as* and *pretty* in *She is not as pretty as she might have been*. She analyzes *actually* as a "discourse particle" (Connector) only in "utterance-initial", "clause-final" and "utterance-final" positions (Aijmer 2002: 257–259). Citing Lenk (1998), she suggests that in other positions it functions primarily as an intensifier.

In her historical study of adverbial Connectors used in medial position, especially *however* and *therefore*, Lenker (2014: 18) shows that "the medial positioning of adverbial connectors" became increasingly frequent from the early Middle English period on, most especially in the 18thC and 19thC, and in academic prose. Discussing adverbial Connectors like *however* and *though*, Lenker (2010: 240–241) notes that in PDE writing, 40% of this set of Connectors occur in medial position, whereas only 2.5% occur in spoken registers. Citing Biber et al. (1999: 891) she interprets the large percentage in writing as a major stylistic development in English writing. In Lenker (2014: 30) she hypothesizes that the development of clause-medial use of Connectors in writing is related to the loss of case morphology and of V2 syntax and the fixing of word order. Like syntactic clefting and focus particles such as *only, particularly*, Connectors used in medial position provide options for marking topic and focus.

Lenker argues that *E*'s with adverbial linking function (called Conjuncts here) are used with a dual function in medial position. They serve as Connectors and in addition have information-structuring function that in spoken language is often marked by prosody. She argues that from late Middle English on two main medial positions can be identified:

1. "Post-initial", i.e. after subject or topicalized CircAdv (Lenker 2014: 17). Here Connectors draw attention to and focus the subject (Lenker 2014: 30),

2. Post-Auxiliary Connectors are "discourse partitioners" that separate topic from comment (Lenker 2014: 31).

Lenker's evidence for the focus function of post-initial use is that subjects that precede the Connector are typically contrastive or specific, e.g. Possessive NP, *this* NP. Typically they are not anaphoric pronouns, which tend to be topics (Lenker 2014: 31–32).[68] In Section 12.4, we will see that the uses of *however, therefore*, and *though* that Lenker identifies as focusing the subject and partitioning the clause are supported by medial uses of *after all* with main verb BE and are supported by medial uses of *by the way*. Lenker's findings are therefore not limited to contrastives. They are construction-specific.

12.3 A hypothesis about the relationship between subjectivity, intersubjectivity and position

In this section I discuss a widely cited hypothesis put forward by Beeching and Detges (2014b) regarding properties of various kinds of linguistic items at "left periphery" and "right periphery" (pre-clausal and post-clausal position). I will explain why the approach taken in this book is not consistent with Beeching and Detges's hypothesis. The authors originally proposed that the correlations posited in Table 12.3 below are universal. However, the correlations were reconceptualized as tendencies because there are various counterexamples, for example use of stance adverbials like *frankly* that Beeching and Detges cite themselves (p. 11) and of the epistemic adverbials *surely* and *no doubt* discussed in Traugott (2014). Beeching and Detges were concerned primarily with characterizing conversational interaction, but, as in the case of Haselow's (2019) set of communicative tasks in Table 12.1 and 12.2 above, the properties could potentially also be useful for thinking about monologal texts as well.

Table 12.3 is a partial listing of the hypothesized uses that Beeching and Detges draw attention to, omitting properties specific to turn-taking such as response-marking and response-inviting. "LP" is short for 'left periphery', "RP" for 'right periphery', "forthc" for 'forthcoming'.

---

68. Interestingly, from the 18thC on, the distribution of *however* identified by Lenker is the opposite of what Kroon (1995) found for Latin *autem* 'but, however'. *Autem* was used primarily after an accessible, given topic, *however* after a focused subject (see Section 12.1 above). Reasons for the difference presumably lie in rhetorical and stylistic traditions that are beyond the scope of this book.

Table 12.3 Some hypothesized usages of linguistic items on the left and right periphery (based on Beeching and Detges 2014b: 11)

LP	RP
links to previous discourse	anticipation of forthc discourse
focalizing, topicalizing, framing	modalizing
subjective	intersubjective

A comparison of Table 12.3 with Tables 12.1 and 12.2 above reveals some differences in formulation of properties at clause endings. "Anticipation of forthcoming discourse" is considered to be a property of RP in Table 12.3, but a somewhat similar task "providing interpretative cues for an upcoming message" is cited in Table 12.1 as relevant at clause-beginning. A task at clause-ending that Haselow cites is "retrospective indication of the kind of relation to prior discourse", not prospective anticipation of upcoming discourse. My concept of DSMs and procedurals being used by SP/W to cue AD/R's interpretation is consistent with Haselow's formulation. So is my concept of post-clausal uses: they tend to provide a "look-back" to D1, especially when they implicate concession, as shown below in Section 12.4. However, Haselow does not mention "modalizing", an important property identified by Beeching and Detges (2014b: 11) at RP. It accounts for the correlation between concessive uses of Contrastives in post-clausal position.

Of particular interest in Table 12.3 is the proposed correlation between subjective and intersubjective uses and position. Focalizing, topicalizing and framing are all subjective acts, so correlation of subjectivity (resulting from subjectification) with LP is well attested. However, as discussed in Chapter 11 and outlined in 11.5, in the case of DSMs, (inter)subjectivity and (inter)subjectification are simultaneously present, so division of labor between LP and RP is not generally attested with respect to these factors, at least not in monologal discourse.

In sum, I here confirm the proposal in Chapter 11.5 and modeled in Figure 11.2 that the communicative function of DSM microconstructions is subjective, intersubjective and textual at the same time.

12.4 Two case studies revisited with position in focus

In this section I outline the positional uses of two DMs that have already been discussed at some length: *after all* (12.4.1) and *by the way* (12.4.2), this time with focus on potential functional differences in the different positions. For both, I have provided examples from COHA in the decade 1850, to maximize comparison. While *after all* exemplifies correlations between position and pragmatic and functional differences up to about 1900, *by the way* does not, so they are interesting

to compare. The studies are coarse-grained. More differences would be revealed by more extensive analysis of more data and especially by detailed multifactorial analysis.

12.4.1 Positions in which elaborative and contrastive *after all* is used

In Chapter 4.5.2 I showed that *after all* originated in a temporal CircAdv that was reinterpreted as an inferential Conjunct meaning 'in the end' in contexts of belief and argumentation and that it was used in rather different ways depending on position. In PDE *after all* used in pre-clausal position tends to be understood to signal that D2 is an elaboration on and justification of D1. In medial position it tends to be understood as justificational or (in the context of generic statements with *BE*), epistemic and confirmatory. In post-clausal position it is often used to mark D2 as justification of D1; it can also be used to mark D2 as a concessive countering an expectation set up in D1, i.e. as a Contrastive. All three positional variants exemplify further subjectification after constructionalization.

Historically, the following very broad stages of development were identified:

1. Late 15thC 'at the end' (temporal CircAdv used in any position suitable for a such an adverbial; use in topicalized position presumably enabled change to Conjunct use)
2. Early 17thC 'in the end' (inferential adverbial 'on reflection, therefore', a Conjunct used in pre-clausal position)
3. Late 17thC justificational DM use (pre-clausal)
4. Early 18thC 'despite expectations' (contrastive, concessive DM use in post-clausal position)
5. Early 19thC 'of course' (epistemic DM use mainly in clause-medial position collocating with main verb *BE* in generic statements)
6. During the 20thC justificational use, which was largely associated with pre-clausal position, came to be extended to post-clausal position.

Here I take a closer look at positional distributions of *after all* in COHA. Given the complexities of use in medial position, I discuss that separately from pre- and post-clausal uses.

Examples of DM use of *after all* in pre- and post-clausal function selected from the COHA 1850s data are:

(9) a. "I talked pretty severely to him, and he got frightened. **After all**, the best way is to use very pointed language to these fellows."
(1850 Yonge, *Poor and proud* [COHA]); pre-clausal, elaborative/justificational ('D2 is my reason for talking pretty severely to him')

b. he is too much like the sour Gage, as we call his mother, to be good for much. But, **after all**, he is not so bad as Dick Rodd, who has never been confirmed, [1855 Yonge, *Scenes and characters* COHA]; pre-clausal, concessive ('although he is not good for much', D2).

c. he gives a record of his sensations. Sensations are the great things **after all.** (1855 Poe, *Works*, Vol. 4 [COHA]; post-clausal; D2 is justification for D1).

d. "Perhaps I shall have to give it up **after all**," said she. "But I will not give up till I am beaten." (1850 Abbott, *Mary Erskine* [COHA]); post-clausal, concessive (D2 'despite wishing not to').

These examples, all from the 1850s, show that neither pre- nor post-clausal position determines the meaning of *after all.*

However, certain readings are preferred throughout the 19thC data in COHA in particular positions:

1. Justification is the preferred use in pre-clausal position (9a); this continues to be the preferred use in PDE.
2. Concessive is the preferred use in post-clausal position up to the 1950s; it always involves look-back to D1 or invites a reconstructed situation in D1 despite which D2 is said to apply.
3. A concessive reading is strongly favored in the 1850 data in pre-clausal position when *after all* is combined with a Contrastive such as *but, yet* or *though* as in (9b). Such combinations almost completely disappear by the 1950s, but if they occur *after all* has a justificational meaning and is no longer affected by the Contrastive, as shown by (10):

   (10) "I thought they wasn't treatin' you quite respectful." "Respectful?… Nonsense – why should they? I'm no older than any of them, except Lily." "Well, yes, that's true, but **after all** you was a stranger".
   (1953 Shellabarger, *Lord Vanity* [COHA]; justification)

4. In the 1850s *after all* appears in post-clausal position with justificational meaning as in (9c). At this time post-clausal justfication use it was very infrequent, but, as shown in Table 12.4 below, it came to be used with greater frequency post-clausally by the 2000s.
5. A concessive use is strongly favored until the 1950s in post-clausal position as in (9d).

Use of *after all* with connective function in medial position is exemplified in (11):

(11) a. Highly as riches are esteemed – the one great good in life as they are regarded – they never have given and never will give this best of all blessings. How little, how very little of the world's happiness, **after all**, flows from the possession of money.
(1852 Arthur, *True riches* [COHA]; justification)

b. This account cleared up the otherwise unaccountable mystery, and showed that the landlord, **after all**, had had no idea of fooling me.
(1851 Melville, Moby Dick [COHA]; concessive 'showed that the landlord had no plan to fool me although I suspected he had')

c. His thought, too, is no less smart than his style…; and this **after all** is the great objection to that manner of writing.
(1851 Barrow, *New Englander and Yale Review* [COHA]; justification combined with epistemic certitude)

d. In most books, the I, or first person, is omitted; in this it will be retained; that, in respect to egotism, is the main difference. We commonly do not remember that it is, **after all**, always the first person that is speaking.
(1854 Thoreau, *Walden* [COHA]; justification and reminder of epistemic truth)

As (11a) and (11b) show, *after all* may be used with justificational or concessive meaning in clause-medial position. If the predicate is the main verb BE, *after all* is usually understood to be used in its justificational sense. In addition, used in clause-medial position before BE it may implicate SP's epistemic certitude (11c), or used after BE it may be intended to remind AD/R of some well-known truth (11d). As mentioned in Chapter 4.5.2, this use with the copula in clause-medial position developed in the 19thC.

In examples where *after all* precedes main verb BE, if an inquit or verb of cognition is interpolated, the break occurs after [subject, after all] as in (12a). There are no examples of a verb of locution (an "inquit") or verb of cognition after [BE, after all] in COHA. However, there is one in COCA (12b):

(12) a. "She scarcely deserved this fidelity on his part," said the monarch, with a dark frown, … "The difficulty, **after all**," he said to him himself, "is not so much to die for one we love, as to find one worthy of dying for".
(1851 Ballou, *The Circassian slave* [COHA]; justification with epistemic certitude)

b. they have no love left for something as quaint as America, and America was **after all**, I said, "always a state of mind to begin with."
(1995 Erickson, *A nation of nomads* [COCA]; justification with epistemic generalization)

The distinction here is reminiscent of Lenker's observation mentioned at the end of Section 12.2.3, that post-subject medial *however* and *though* are used to focus the subject, but post-verb use focuses the complement. However, the constraint she identifies on subjects with these Contrastives, that they be non-anaphoric, does not apply in the case of *after all*. In (13) it is used following the anaphoric pronoun *he*:

(13) It would be a fearful thing for you, my child, should he, **after all**, turn a charmed ear to the voice of that Syren, the end of whose song is destruction.
(1859 Cary, *The adopted daughter* [COHA]; concessive)

Table 12.4 below gives the numbers of pre-clausal, clause-medial and post-clausal uses in the first 200 hits of COHA release 2009 every fifty years starting with 1850.[69] I do not distinguish post-subject and post-finite verb uses, but I do distinguish epistemic (e) and concessive (c) uses. By way of reminder, as pointed out in Chapter 1.4, I take DM position before a subordinator to be clause-medial, because the subordinate clause is an argument of the clause, as in:

(14) But it may be said **after all**, that we have done the Bishop's doctrine injustice.
(1844 *Review of the errors of the times* [COHA])

As in any corpus set, numbers of hits are skewed by authorial preferences (see Schmid and Mantlik 2015 for evidence of distinct authorial preferences for use of [N BE that] constructions, e.g. *the idea was that X*). In the 1850s data from COHA release 2009, of the 20 DM hits from Melville's *Moby Dick* (1851), only 5 are pre-clausal, but of the 17 DM hits from Stowe's *Uncle Tom's Cabin* (1852), 9 are. In Table 12.4, the total number of DM uses in a particular position is followed by a breakdown into justificational (j), and concessive (c) uses. For medial position, I further distinguish epistemic (e) uses. In general what I counted as epistemic 'as is well known' coincides with presence of predicate *BE*, as in (11d). "Non-DM" uses are mainly literal Notable in Table 12.4 is the proportional decline in use of concessive *after all* especially in pre-clausal position from the 1950s on. This is directly correlated with decline in use of combinations with Contrastive DMs *but* and *yet*, which triggered concessive implicatures (see Example (9b)). Notable also is the decline in the last period of post-clausal concessive use, balanced by a proportional increase in justificational use. adjectival modifier uses (*after all these years*), or pronominal uses where *all* means 'everything' (*after all is over*). "Non-DM" also includes 3 examples of monofunctional 1DSM 'in the end' use in the 1850 data.

---

69. Data in COHA for Tables 12.4 and 12.5 were retrieved in March and September 2019 using the then-available corpus and checked in August 2021. COHA has since been expanded; data from TV and Movies now appear at the beginning of the 1950s and 2000, and therefore there are discrepancies between the first 200 hits and the set I extracted. However, all the examples I cite are still in the corpus.

Table 12.4 Number of examples of DM *after all* attested in COHA release 2009 in different positions (the first 200 hits; data retrieved Sept. 2019)

	Pre-clausal	Medial	Post-clausal	Non-DM
1850s	67 (39j/28c)	58 (12j/23e/23c)	45 (3j/42c)	30
1900s	68 (41j/27c)	25 (7j/12e/ 6c)	77 (5j/72c)	30
1950s	109 (107j/2c)	22 (5j/ 6e/11c)	46 (9j/37c)	23
2000s	76 (75j/1c)	23 (9j/10e/ 4c)	58 (28j/30c)	43

The main tendencies attested in COHA from the early 19thC on with respect to correlations between meaning and position of DM uses of *after all* are outlined in Figure 12.1. To clarify the uses in specific positions, features are repeated across periods, even if they are the same, unlike in prior Figures.

In Figure 12.1 the numerals 1, 2, 3 are shorthand for pre-clausal position, medial position and post-clausal position, respectively. The main changes associated with *after all* are in medial and post-clausal position. The fact that concession declines in post-clausal position from the 2000s on is shown by the use of grey font:

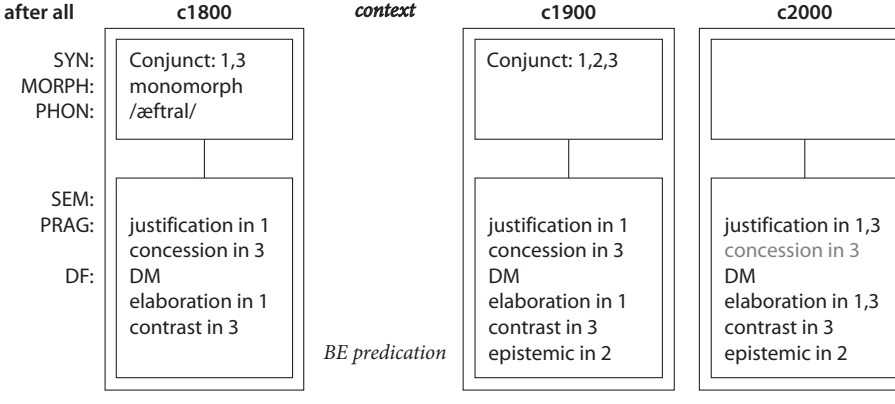

Figure 12.1 DM *after all*: main correlated tendencies between position and discourse function in COHA release 2009 and changes to them

12.4.2 Positions in which digressive *by the way* is used

In Chapter 8.2 *by the way* was shown to originate in a dynamic spatial CircAdv that was reinterpreted as a Conjunct in contexts of verbs of locution, topicalization and relativization. The following very broad steps in the development of *by the way* can be identified.

1. 11thC 'en route' (CircAdv used in any position suitable for a such an adverbial; use in topicalized position presumably enabled change to Conjunct use)
2. Late 16thC 'in passing' (digressive Conjunct used as a 1DSM in pre-clausal position and elsewhere, mainly in religious and philosophical texts)
3. Late 17thC 'incidentally' (digressive DM used in pre-clausal position and elsewhere; I call this a "dismissive digression")
4. Mid 19thC use as hedge, mainly in pre-clausal position
5. 20thC use as marker of aggression, especially in the context of *Oh*; pre-clausal position

Unlike in the case of *after all*, there is relatively little correlation of position with distinctive variation in pragmatics or discourse function.

To allow for comparison with (9), examples of *by the way* in pre- and post-clausal positions and uses from the 1850s data in COHA are presented in (15):

(15) a. Give my compliments to your sister Julia. **By the way**, do you know that I always admired her very much?
(1854 Holmes, *Tempest and sunshine* [COHA]; pre-clausal; evaluation presented as unimportant; dismissive digression)

b. little Lewie is very ill, I fear. **By the way**, Harry, run and tell Matthew that just as soon as he is warm, he must drive as fast as possible to the village,
(1853 Bradford, *Lewie* [COHA]; pre-clausal; hedge introducing an urgent directive)

c. for what he called "mature deliberation" – often one of the greatest enemies to upright, generous, and disinterested action – to hope, faith, and charity, that I know of, **by the way.** (1855 Southworth, *The missing bride* [COHA]; post-clausal, dismissive digression, probably a hedge on what precedes)

d. one [reminder] was that of the marriage of Margaret, her handmaid, to the man with the four children (who had lost his wife, **by the way**); and the other was… (1853 Cozzens, *Prismatics* [COHA]; implicating a dismissive attitude to the explanatory information provided ('the man with four children could marry Margaret because his wife had died')

Use of *by the way* in post-clausal position can in instances where the content of D2 is serious be interpreted as more dismissive of the content of D2 than use in pre-clausal position (see e.g. (15c) and (15d)). In the data for the 1850s, *by the way* occurs in post-clausal position only infrequently (see Table 12.5 below). It occurs more frequently in the later data. There it may also be used in a short formula of the kind *I'm Jane, by the way* (2006), *My name is Jefferson, by the way* (2005). The

context for these formulaic uses is a perceived awkward time-delay between meeting and introduction (see Chapter 8.2).

In the case of some medial uses, especially copula clauses with main verb BE, *by the way* may be used to focus the subject rather than to imply dismissiveness or secondary importance. This is exemplified in (16a), where *by the way* is used clause-medially as an indirect way to call attention to what the speaker actually considers to be an important point about how a sick person's mind works. Post-verbally, *by the way* appears to focus on the complement. This distinction is reminiscent of Lenker's (2014) hypothesis that post-subject medial contrastive Conjuncts focus the preceding subject, and elsewhere in the clause they enhance the default focus.

(16) a. Her mind was unfavorable to her cure; and this, **by the way**, is a very important particular in the fortunes of the sick. (1856 Simms, *Charlemont* [COHA]]; medial, post-stressed pronoun; drawing attention to the link between D1 and D2)

b. she had been wandering off into the fields at the foot of the garden, where it was safe and still. There is, **by the way**, a peculiar awe in the utter hush of the earliest morning hours. (1869 Philips, *Men, women, and ghosts* [COHA]; medial, drawing attention to the 'peculiar awe')

DM uses of *by the way* signal that D2 is a digression and express a degree of (often feigned) incoherence with the main topic, hence (feigned) dismissiveness of the content of D2. This feigned dismissiveness enabled use as a hedge on face-threatening content in D2, as in (17):

(17) And **by the way**, let me beg you not to call a TROTTING MATCH a RACE. (1859 Holmes, *Autocrat of the breakfast table* [COHA])

In (17) there is a double hedge as both *by the way* and the indirect directive *let me beg you* are used.

Numbers for the first 200 hits of DM uses of *by the way* in COHA (release 2009) for the years 1850, 1900, 1950 and 2000 are shown in Table 12.5, for comparison with *after all* in Table 12.4. It is not always possible to tell with certainty whether a hedge (h) is intended. What I have counted as hedges are those uses of *by the way* that precede a D2 with probably face-threatening content. They are typically questions or directives followed by a grudging response or silence, as in (18) (Example (10a) in Chapter 8.2, repeated here for convenience):

(18) Verry still played. "Her talent is wonderful," said father, taking the cigar from his mouth. "**By the way**, you must take lessons in Milford; I wish you would learn to sing." I acquiesced, but I had no wish to learn to play. (1862 Stoddard, *The Morgesons* [COHA)

'Non-DM hits' refer to *by the way* used in a phrase, not as a Connector that connects clauses (e.g. *a new one by the way*, 1856), as a spatial adverbial meaning 'en route', a manner expression ('the manner in which'), or monofunctional 1DSM 'in passing'.

Notable in Table 12.5 is that pre-clausal position has been prototypical throughout the period. But there has been a steady increase of use in post-clausal position relative to the first period and modest decrease of use in clause-medial position relative to the first period.

Table 12.5 Number of examples of DM *by the way* attested in COHA release 2009 in different positions (the first 200 hits; data retrieved March 2019)

	Pre-clausal	Clause-medial	Post-clausal	Non-DM
1850s	58 (19h)	45	5	92
1900s	120 (79h)	20	9	51
1950s	110 (27h)	24	36	30
2000s	72 (28h)	32	43	43

Tendencies attested in COHA from the early 19thC with respect to correlations between meaning and position of *by the way* are outlined in Figure 12.2 ("pref" is short for 'preferred'). As in Figure 12.1 in Section 12.3.1, the numerals 1, 2, 3 are shorthand for pre-clausal position, clause-medial position and post-clausal position, respectively. Note that, to highlight the new use as a hedge, in Figure 12.2 features are not repeated across periods. Conjunct use is specified as being preferred in position 1; that does not exclude use in positions 2 and 3. The pragmatic component is specified as applying to all three positions. Unlike in the case of *after all*, there is only one significant change and it is associated with pre-clausal position.

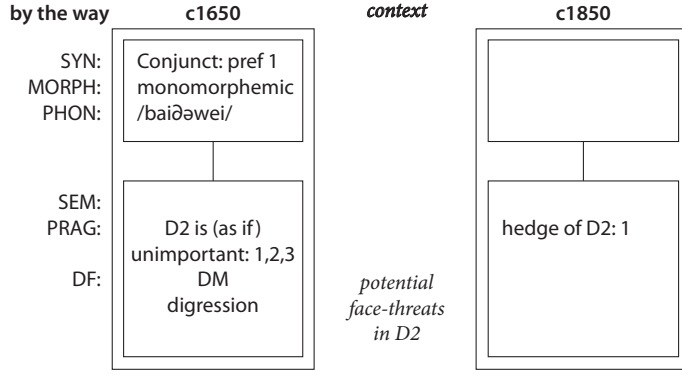

Figure 12.2 DM *by the way*: main change in correlated tendencies between position and discourse function in COHA

12.5 Conclusion

Pre-clausal position is associated by default with Conjuncts. Pre-clausal position is the default position in English for DSMs that link D1 with D2, whether mono- or multi-functional, but other positions are possible for some DSMs. As the history of *by the way* shows, there may be changes in the discourse function of a connective in pre-clausal position. As the history of *after all* shows, there may also be changes in the discourse function of a DM in post-clausal position. Therefore no clear generalization about function in position can be made.

Positionally, *but* has been preferred pre-clausally, *all the same* post-clausally. Post-clausal use of *but* is probably not a case of analogization to *all the same*, as the latter is not a multifunctional DM and is not preferred in spoken varieties of English, but the preference in 18thC English for contrastive *E*s to be used concessively in that position was probably a factor.

Whether or not a Connector can be used with a linking meaning in several positions and what its discourse function is in those positions depends on the individual Connector. As *after all* and *by the way*, among others, show, some Connectors can be used in all three positions, some (e.g. *after all*) with significantly different discourse functions, others with less distinct functions (e.g. *by the way*). Some, like elaborative *and* and inferential *so*, can be used as Connectors only in pre-causal position. If *and* is used clause-internally, it is to coordinate Ns, Vs or As. Clause-internally, *so* signals degree (*X was so frustrated*), not inferential connection.

To the extent that it is possible to say anything specific about the meaning of a clause position in English, there has been a tendency since the 18thC to associate concessive meaning with contrastive markers in post-clausal position. However, use in post-clausal position does not determine this meaning, as evidenced by the increasing tendency over time for *after all* to be used with justificational meaning in post-clausal position (see Table 12.4). Furthermore, concessive meaning is not associated with non-contrastive markers such as *by the way* (or with PMs such as tag questions like *isn't it*, or general extenders such as *and stuff*). Concessive meaning is therefore not a predictable feature of post-clausal use, except partially for contrastive Connectors.

The data discussed here lead to the conclusion that position is not a construction. Rather, positional function needs to be specified in terms of tendencies for each Connector micro-construction individually (Hypothesis 3). In the case of Contrastives, the tendency for post-clausal DMs to be used concessively can be generalized for the Contrastive subschema, but not for others.

As mentioned in the Introduction, the argument that position is not a construction is a testing ground for Goldberg's claim that "it's constructions all the way down" (e.g. Goldberg 2003: 223). Some researchers have said that a weakness of

construction grammar is that it is not testable (see e.g. Dunn 2017). However, Perek (2012) tested a hypothesis about ditransitives and their hypothesized relationship with caused motion constructions experimentally. Testing the "constructions all the way down" hypothesis in this Chapter was not conducted experimentally. However, I believe that logic shows that slot positions are not constructions. One counter-example, and a partial one at that, given the correlation between Contrastives, concessive pragmatics and post-clausal position, does not undermine the theory of construction grammar, but rather strengthens it by providing further evidence that it is testable.

Several questions remain to be explored in future work on position in the development of Connectors. Among them are:

a. Some Connectors are used in post-clausal position, e.g. *by the way, after all, all the same*, others not, e.g. *and*, *now*, and for most SP/Ws, *but*. Are there any constraints on meanings/functions in post-clausal position beyond the likelihood that contrastive DSMs will have concessive meaning when used in this position?
b. To what extent can the communicative tasks outlined by Haselow (2019) for turn-taking, and recast in Tables 12.1 and 12.2 in terms more suited to the kinds of monologal discourses often found in historical texts, be drawn on to enrich understanding of positional uses in both monologal and interactional texts?
c. Lenker (2014) has suggested that partitioning of a clause into topic and focus is a task that can be performed with Contrastives in medial position. What kinds of communicative tasks should be identified for other kinds of metatextual *E*s used in medial position, especially in writing (see Section 12.2.3)?

CHAPTER 13

Changes in networks and nodes

13.1 Introduction

Constructions clearly do not exist in a vacuum. Conceptualized as units that are stored in the constructicon, they are by hypothesis linked to each other to various degrees of strength in networks. A key concept in Goldberg's model of construction grammar is captured in her statement that "language is a network of pairings of form and function", that is, constructions.[70] She envisages "construction-based theory" as an account of "the network of constructions [that] captures our grammatical knowledge *in toto*" (Goldberg 2006: 18, italics original). These networks are specified in the constructicon.

How to conceptualize networks, the nodes where links join up, and relations that constitute the network is a question to which answers have evolved substantially in the last decade and are still evolving. Diessel (2019) is devoted to the question of how to treat networks and nodes in construction grammar especially with respect to acquisition. Sommerer and Smirnova (2020) addresses the question from a diachronic construction grammar perspective. Here I seek to answer Smirnova and Sommerer's (2020: 3) question "How can node creation and node loss be implemented in the network model?" I start by outlining some largely synchronic proposals that have been made about networks and go on to discuss issues that arise in historical approaches to networks.

Discussion is limited here, as in the rest of the book, to language-internal properties. However, a proposal for multidimensional networks can be found in Fried and Östman (2005), one of the first papers to account for pragmatic markers from a constructional perspective. The proposal is synchronic, and network links are shown for, among other markers, the Czech PM *jestli* ('I wonder', in certain contexts). This PM serves as a deference marker and hedge. It is restricted to interactional conversation and is shown to involve links not only to linguistic functions such as questions, requests and assertions, but also to cultural coherence and speaker-based attitudes.

---

70. https://www.einsteinfoundation.de/en/people-projects/einstein-visiting-fellows/adele-goldberg/; accessed August 12th 2020.

In what follows some discussion is provided of the network metaphor and what it encompasses. Two types of network that have received considerable attention in the constructional literature are discussed: a "vertical" inheritance network and a "horizontal" network (Section 13.2). Section 13.3 is devoted to the role of networks in change. In Section 13.4 suggestions are made about how to represent changing networks. Section 13.5 tentatively provides a way to model context in a changing network. Section 13.6 is a brief conclusion.

13.2 The network metaphor

The concept of a "network" is based in computational research on neural networks (see e.g. Rumelhart and McClelland 1986). In usage-based views of grammar, networks are dynamic and can be modified. From this perspective, replicated innovations "could be viewed as the formation of associative connections" (Hilpert 2018: 33). It is hard, if not impossible, to show network changes adequately in two-dimensional space on a page, so most networks proposed for linguistic change, whether in acquisition or in historical records are modelled in terms of successive stages of synchronic networks.

Networks have "nodes" where two or more links join, as in Figure 13.1:

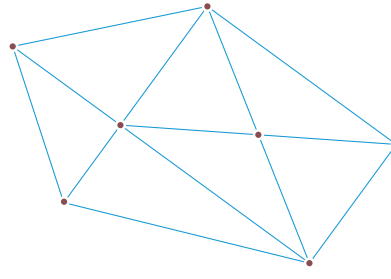

Figure 13.1 A simple network with nodes

To date, the nodes have usually been conceptualized in construction grammar as constructions, i.e. rich bundles of features with values. For purposes of describing particular changes, the hypothesis that nodes are constructions is a helpful shorthand and is maintained here. Mindful of the problem of the "cognitive commitment" for work in historical linguistics (Chapter 3.2.2), I do not seek to suggest that the nodes that I posit as a linguist are anything more than very crudely analogous to neural networks. More psychologically based alternatives with focus on the way the mind works have been suggested in which the structure lies mainly in the links that are activated in language use, not in the nodes (e.g. Hudson 2007, 2015; Schmid 2017).

Goldberg (1995) suggested that there are various kinds of links between constructions. Budts and Petré (2020) provide a contemporary view on types of links. Two main types of network relationships have received a lot of attention in the construction grammar literature: a "vertical" taxonomic inheritance relationship between constructions at different levels of abstraction and a "horizontal" relationship between constructions that are similar. A third type of link is "paradigmatic", the relationship between sets of alternative constructions, for example in morphology, tight sets of alternatives in a tense or case frame (Diewald and Smirnova 2012), or more loosely among modal auxiliaries in English (e.g. Budts and Petré 2020). Other links include metaphorical links, such as, in Goldberg's view (1995: 89), between ditransitive *Kim gave Bill a pear* and caused motion *Kim gave a pear to Bill*, where reception of the pear is understood as the result of moving the pear toward Bill. Among the examples of DSMs discussed here, a metaphorical link was inferred in the Middle English period between the CircAdvs *by the way*, *further* and a "journey" of argumentation in the construction of texts.

In this section I discuss "vertical" networks (13.2.1) and "horizontal" networks (13.2.2) in more detail. The vertical and horizontal networks by hypothesis capture part of what we know about Conjuncts and their DSM subclasses at a particular "synchronic" moment in time. This knowledge is only a small part of our knowledge of the conceptual space in which DSMs arise and are used. How vertical and horizontal networks can be conceptualized in work on change is discussed in Section 13.3.

13.2.1 "Vertical" inheritance networks

In early work on construction grammar networks were conceptualized as taxonomic, hierarchical relationships (Goldberg 1995, 2006; see also Boas 2013), which characterize abstractions over utterance types. These are sometimes referred to as "vertical inheritance networks". To illustrate, the Connector.Cxn within the Connective.Schema (see Figure 4.3 in Chapter 4.3.1) is the abstract "mother" of several "daughters" as in Figure 13.2. The daughters include the Elaborative.Cxn (EC), Inferential.Cxn (IC), Contrastive.Cxn (CC) and Digressive. Cxn (DC) as well as others, represented by a box containing '…', such as the Return to Prior Topic.Cxn (see Chapter 9). All of the subschemas can be used to signal topic-shift. *E*s in the bottom line, e.g. [also], represent microconstructions, abstractions over constructs with the same pairing of form and meaning. The abstract Connector.Cxn can be specified in terms of the combined default characteristics of Connectors outlined in Chapter 4.4.2 (Figure 4.4) and Chapter 11.5 (Figure 11.2), with some additional features.

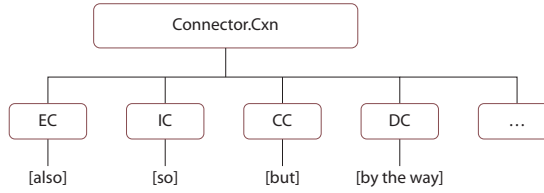

Figure 13.2 Partial taxonomic network of the Connector.Cxn in present day English

Figure 13.3 specifies in addition that the construction occurs in a [D1 __ D2] frame. This predicts that the default position for Connectors is pre-clausal. It also specifies non-integration with the host clause and [+conventional] pragmatic status, and also procedural status. The notation […] in Figure 13.3 means that the feature in question is unspecified and will be specified by particular microconstructions. Degree of contentfulness and truth-conditionality as well as degree of topic-shifting function are construction-specific. So is multifunctional DM status and overrides the monofunctional 1DSM status specified in Figure 13.3.

All [[Conjunct] ↔ [DSM]] micro-constructions "below" the level of the Connector.Cxn in a taxonomic hierarchy "inherit" these characteristics, so the characteristics do not have to be specified for every subschema and every micro-construction (this is called "default inheritance", Goldberg 2013: 2–23). Such default inheritance may be overridden or elaborated in particular cases. With respect to position, for

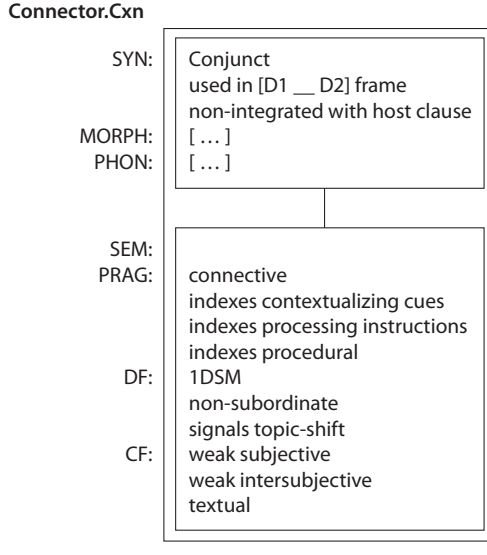

Figure 13.3 Model of the characteristics of a Connector.Cxn

example, because of default inheritance, pre-clausal position does not need to be specified for subschemas and individual microconstructions. However, other specific positional factors, such as that *after all* can be used in clause-medial and post-clausal positions as well as pre-clausal position (see Chapter 12.3.1) and that *all the same* can occur in post-clausal as well as pre-clausal position (see Chapter 7.3), need to be specified for those particular micro-constructions. Whether a particular micro-construction is discourse functionally a multifunctional DM also needs to be specified individually.

By hypothesis the default features in Figure 13.3 can be reconstructed as far back as Old English at a minimum for *ac* 'and', and have remained constant throughout the history of English. Changes involve the development of overrides and construction-specific additional features when a particular 1DSM has been constructionalized.

13.2.2 "Horizontal" networks

Because vertical taxonomies do not account for variation, in some work they have been expanded, particularly with respect to polysemy (e.g. Cappelle 2006; Van de Velde 2014 and Perek 2015). Cappelle suggests that just as in morphology there are allomorphs such as /-s, -z, əz/, so there can be "allostructions", variants ("alternations") of constructions under certain conditions, e.g. *pick the book up ~ pick up the book*. Noting that "structurally different elements can fulfil the same function", Van de Velde (2014: 141) suggests that variants that are polysemous can be thought of as linked "horizontally" within a hierarchy at the same level of abstraction. A syntactic example he gives is the position of the finite verb (Vfin) in Dutch: in main clause declaratives it is usually in second position (Vfin–2), in polarity questions, conditionals and imperatives it is usually in first position (Vfin–1) and in subordinate clauses it is usually in final position (Vfin-n). These predictable variants can be displayed in a combined vertical and horizontal network such as Figure 13.4. Horizontal relationships are shown with a double-headed arrow.

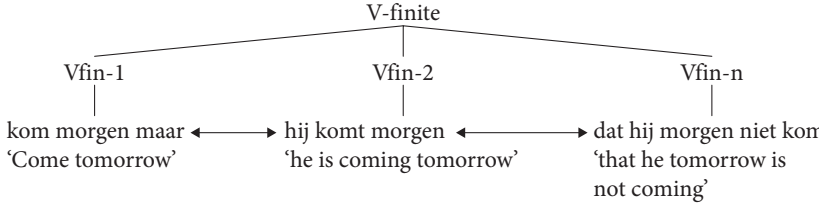

Figure 13.4 Dutch finite clauses as a constructional network (based on Van de Velde 2014: 150)

13.3 Networks and change

Network hierarchies are constantly being internally modified. For example, productivity, the openness to new collocates of a construction, may change. So may membership of the (sub)schema. As discussed in Chapter 6, in the Elaborative. Cxn subschema, Old English *eac* 'also' obsolesced in Standard English by the end of the 19thC, but continued to be used as *eke* in some other varieties of English. *In addition* was added to the construction in the 20thC. As discussed in Chapter 8, in the Digressive.Cxn subschema *by the by* obsolesced during the 20thC. Relationships between schemas and external subschemas can also change, see e.g. Torrent (2015) on reorganization of the *para* 'for to' infinitive construction between Peninsula and Brazilian Portuguese and Gyselinck (2020) on the "intensifying fake resultative" in Dutch (the equivalent of *drinks himself drunk*).

There has not been much discussion of how vertical inheritance and horizontal models can be used to enhance thinking about constructional change. In sections 13.3.1 and 13.3.2 I briefly point to ways in which the concepts of vertical and horizontal networks respectively can be used in thinking about changes in nodes and networks. In 13.3.3 the importance of including context is addressed and in 13.3.4 a tentative diachronic network model is proposed for *by the way* to exemplify the prior discussion.

13.3.1 Vertical inheritance networks and change

Here I point to a few factors that should be considered with regard to vertical inheritance networks: strengthening and weakening of links, and reorganization of parts of a taxonomy.

The taxonomic relationships in vertical inheritance networks can be built and strengthened as schemas are expanded with new members (type productivity), or they can be weakened as members cease to be used (obsolescence). Strengthening of inheritance links occurs when new members of a subschema are adopted (e.g. *also, all the same, by the way*). Any link from a mother node (e.g. Connector.Cxn) to a micro-construction may be strengthened, resulting from the expansion of a particular subdomain such as Contrastives.

Adoption of a new member of a taxonomy via the process of constructionalization typically entails breaking a symbolic link in one hierarchy and establishing a new one.[71] We have seen how various spatial adverbials like *by the way, further*, which inherited features from the Spatial.Schema, were replicated in topicalized

71. An exception is borrowing, unless perhaps when SP/W is totally bilingual.

position within a clause and then reinterpreted in [D1 ___ D2] frames in discourses referring metaphorically to textual space. These were eventually constructionalized as Conjuncts in the context of the [D1 __ D2] frame and incorporated into one of the subschemas of the Connector.Cxn, depending on their lexical meaning (Elaborative.Cxn in the case of *further*, Digressive.Cxn in the case of *by the way*). As this was happening, their inheritance relationship to the Spatial.Cxn was slowly stretched and destabilized. However, they were not stretched in this way in other contexts; therefore *by the way* and *further* continued to be used as instantiations of the Spatial.Cxn in other contexts. As links from the metaphorical use to the Connector.Cxn were strengthened by replication and expansion of contextual collocates, both sets of taxonomic relationship were destabilized. Eventually, as the two spatial adverbials came to be used more and more in [D1 __ D2] contexts, the vertical inheritance link with the Spatial.Schema came to be broken and a new Conjunct use developed within the Elaborative.Cxn and the Digressive.Cxn subschemas of the Connector.Cxn respectively. In the case of *by the way* the result was a polysemy/horizontal relationship between the source use 'along the route' and the new metaphorical 'in the course of the argument' use in the late 1500s. In other words there is "layering" (Hopper 1991) of the earlier and later uses.

13.3.2 Horizontal networks and change

With respect to horizontal networks, Zehentner (2019), drawing on Perek (2015), used the idea of horizontal network relationships to support the concept of "dative alternation", arguing that ditransitives and prepositional variants are in a horizontal polysemy relationship (e.g. *Kim gave Bill a pear ~ Kim gave a pear to Bill*, which both have the basic meaning 'X cause Y to receive Z', but have different forms: subject – verb – object 1 object 2 [SUBJ V OBJ1 OBJ2] and subject-verb- object2 *to* object 1 [SUBJ V OBJ2 to OBJ1]). This horizontal network relationship was established by about 1500. It is a highly productive construction, sanctioning many new verbs, e.g. *fax someone something/fax something to someone*. Zehentner and Traugott (2020) argue that the "benefactive" alternation (e.g. *Kim bought Bill a pear ~ Kim bought a pear for Bill*) also involves a horizontal network relationship with the shared meaning 'X intend Y to receive Z' and the forms [SUBJ V OBJ1 OBJ2] and [SUBJ V OBJ2 *for* OBJ1]. This relationship was established around 1600, about a century later than the dative alternation. Unlike the dative alternation, the benefactive alternation is low in productivity, with few new collocates and few new members. Nevertheless, by hypothesis, "paradigmatic analogy" (Perek 2015) of the benefactive alternative to the already extant dative alternation led to the conceptual combining of the two sets in a higher level, more abstract ditransitive schema (Zehentner and Traugott 2020: 197). This schema vertically licenses two

constructions each of which has a horizontally linked relationship with a prepositional phrase. Figure 13.5 provides a sketch of the network relationships as they are attested in corpora from the 17thC on:

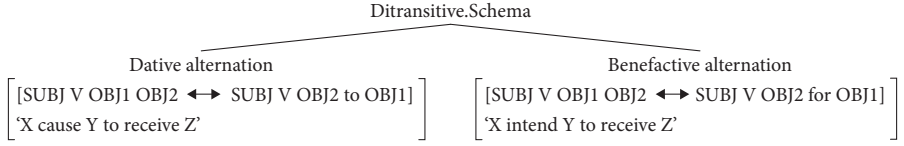

Figure 13.5 Modern English ditransitive.schema

It is usually assumed that constructions linked as alternative expressions in a horizontal relationship are daughters of the same mother node, as illustrated by Figures 13.4 and 13.5 above.

Just as vertical relationships can be broken, so can horizontal ones, for example by analogical leveling (Van de Velde 2014). For example, in Old English finite verbs were largely used in V2 position in main clauses and in final position in subordinate clauses (as in Dutch as shown in Figure 13.4 above), but during Middle English this relationship was largely broken and syntactic V2 obsolesced. However, typically "the semantic differences that are formally expressed by the horizontally related nodes, [sic] survive" (Van de Velde 2014: 159). This is possible because "form-function changes involve strengthening of already available resources with extension to new domains when a subsystem comes under pressure" (Van de Velde 2014: 173). In the case of V2, its discourse functional property of focusing the expression that precedes survives in "preposing constructions" such as "From the director of X comes Y" (see footnote 65 in Chapter 12).

13.4 Representing changing networks

In Chapter 12.3.1 the meaning correlations between position and use of *after all* were discussed. Such factors should be represented in a network model that seeks to account for a particular change in detail. How best to do so in two dimensions remains to be determined. Figure 13.6 below attempts to sketch the network links and the break in the inheritance link relevant to the history of *after all*.

In modeling Figure 13.6 I assume two abstract high-level schemas: temporal and connective terms of taxonomic hierarchy, which coexist in constructional space in the constructicon. The network links associated with the development of *after all* include the following schemas: Temporal.Schema (the CircAdv use); Inferential/reasoning.Schema (first the 'in the end', then the justification meanings), Contrastive.Schema (the concessive meaning) and the Epistemic.Schema (the 'of

course' meaning in clause-medial uses after BE). Horizontal networks are often conceptualized as alternative expressions of the same mother node, as mentioned above in Section 13.3.2. Because concessive is a type of reasoning, it can be argued that the inferential/reasoning schema is its mother node, in which case a horizontal relationship can easily be established, especially as in its early uses it is highly context-dependent on negation and conditional reasoning. However, to capture the fact that at least intuitively *after all* can be replaced by such contrastive *E*s as *all the same*, and that for much of its earlier history *after all* functioned as a marker of contrast, especially in final position, it seems appropriate to suggest that horizontal relationships can exist between polysemies of a DM across domains. Specifically, I posit that justificational use is directly linked to the inferential/reasoning schema, concessive use is linked horizontally to the contrastive schema. At a later stage in its history, a horizontal link with the epistemic schema was established.

In Figure 13.6 the following notation is used: downward arrows show the historical trajectory, e.g. temporal uses precede justificational and concessive uses. Constructionalization of *after all* as an [[Inferential] ↔ [DSM]] Conjunct resulted in a break in the inheritance link and a new link with the Connector.Cxn. This is shown by an angled upward arrow. The link to Epistemic.Schema attested by use of *after all* in an 'of course' meaning when used clause-medially is a link to an "external" schema outside the domain of the Connector.Cxn; this relationship is shown by double curvy arrows. Horizontal links between justification and concessive readings and between justification and epistemic readings are shown by double-headed horizontal arrows. Lines without arrow heads between the Connector.Cxn, the Inferential.Cxn and the Contrastive.Cxn show that Connector.Cxn is the mother node for these two subschemas.

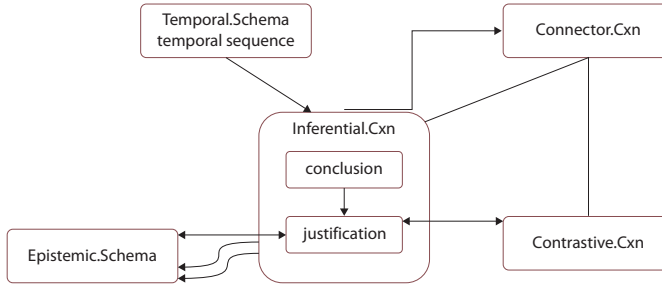

Figure 13.6 Partial sketch of network links in the development of *after all*

Figure 13.6 models relationships between construction types. Figure 13.7, by contrast, is a preliminary sketch of some of the factors that need to be accounted for in a network model at the higher level of schemas, illustrated by the constructionalization of the DSM *by the way*. As summarized above, when used sufficiently

frequently in topicalized initial position and especially in [D1 __ D2] contexts, it came to be associated with the Connector.Cxn.

Figure 13.7 distinguishes subschemas within the two high-level abstract schemas, spatial and connectivity. Each of the high-level schemas has subschemas. The Spatial.Schema has at least dynamic and static subschemas (see Chapter 8.2). As proposed in Chapter 4.3.1, Figure 4.3, the Connectivity.Schema has at least two subschemas: the non-subordinate Connective.Schema and the Subordinate.Schema. The Dynamic.SpatialCxn licenses various CircAdv microconstructions, one of which is *by the way*. As in the case of *after all* represented in Figure 13.6, constructionalization of *by the way* involved a break in the Dynamic Spatial inheritance link and a new link with the Connector.Cxn; this is represented by an upward angled arrow. "subord" is short for 'subordinate'.

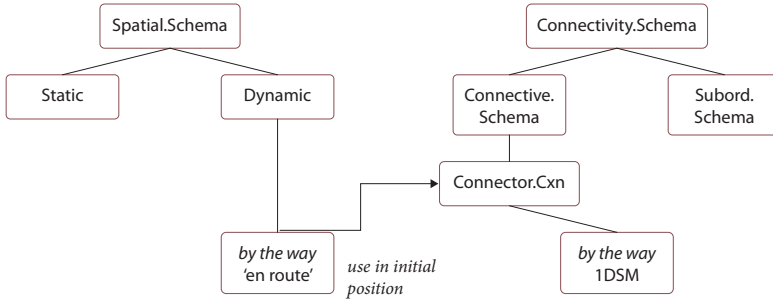

Figure 13.7 Partial sketch of the initial network change undergone by *by the way* c1500

Any link to a mother node may be weakened and eventually lost, e.g. the link with the Spatial.Schema has been lost in the case of *but* (Chapter 7.2). In the case of the Connectors discussed in this book, there is extensive assignment of particular adverbials to existing Connector type subschemas. Often the original use persists, e.g. *now* can still be used as a temporal adverbial, *further* as a spatial one. Sometimes, however, the original micro-construction will be used less and less and replaced by another one, for example, dynamic spatial adverbial use of *by the way* was largely replaced with *on/along the way*. This kind of loss, called "constructional death" in Smirnova and Sommerer (2020: 3) is not often discussed in any great depth. An exception is Kuo (2020), who conceptualizes loss in a Chinese contrastive schema in terms of constructional competition of micro-constructions within a schema and adjustment of fit to prototypes. In a usage-based grammar, the notion of "competition" is problematic as constructions are not agents that can compete. However, fit to prototypes, understood as SP/Ws' tendency to analogize to existing patterns, is a very likely motivation for change.

13.5 Incorporating context into network models

Figure 13.7 includes the place-holder 'use in initial position' for contextual factors in network change. It is slightly expanded in Figure 13.8 below. A challenge that I have not seen addressed except in passing in constructional work on networks is how context interacts with them. Assuming that the network links represent potentials for activation in contexts, contexts can be conceptualized as guiding interpretations that lead to gradual modifications of constructional components, especially those that are related to meaning especially, in the case of DSMs, to discourse function.

As discussed in Chapter 3.7, contexts that researchers on change find relevant are those that are, or appear to be, enabling factors in change. They are co-texts, the discursive activities which are replicated and invite various inferences that may eventually lead SP/Ws to create new symbolic links between form and meaning. In the data discussed in this book, two contextual uses have been repeatedly identified.

1. Topicalized use of a CircAdv in clause initial position, where it frames the up-coming content of the clause. This appears to be a necessary context for the constructionalization of a CircAdv or other *E* as a Conjunct.
2. As mentioned in Chapter 4.4.3, Lenker (2010: 38) notes that many *E*s with CircAdv origins were reinterpreted as Connectors in contexts where the CircAdv was used to modify a verbal phrase containing a *verbum dicendi* or 'verb of speaking'.

It is well established that context can be "imported into discourse" (see e.g. Levinson 2003). The first type of context allowed for the kind of profile shift from atypical use of the CircAdv to default use as a Conjunct outlined in Chapter 4.4. The second type of context resulted in the kind of context importation has been theorized as "context-absorption" (Bybee et al. 1994: 296). As (Kuteva 2001: 151) describes it, "what before the context-absorption had to be made explicit in the immediate context no longer needs to be stated", although it can be. The 'say' context was absorbed into DSMs which conventionally implicate illocutionary functions such as Elaboration and Digression. Other examples of context-absorption include the way in which *after all* was used in the 17thC in the context of *for* and *yet*, the first of which cues justification, the second concession. The meanings of these Connectors was absorbed into *after all*, one of several distributional changes that enabled it to be used independently as a DM (see Chapter 4.5.2).

My aim here is to put these ideas together and tentatively propose a way of representing context in a historical account of constructional network changes. I assume that the network links represent potentials for activation in contexts, as suggested by Hudson (2010: 89). Replicated contexts can be seen as guiding

interpretations that lead to gradual modifications of constructional components, especially those that are meaning related. The changes discussed in this book are cases of gradual shift in meaning and function from CircAdvs and other *Es* toward Conjunct status resulting from strengthening and weakening of links in replicated contexts, so context ideally needs to be included in a network model.

Figure 13.8 is a highly simplified and truncated preliminary attempt to represent the role of discursive context in changes involved in use of *by the way* over time. The kinds of factors that a diachronic network account of this DM should represent include the approximate date when the changes occurred, the break in the inheritance link as in Figure 13.7, contextual links involving use of the CircAdv in topicalized initial position and in references to 'talk' and links in its later history to mitigators. In the earliest period,[72] the *by (the) way* micro-construction inherits syntactic and semantic properties as a CircAdv from the Spatial.Schema in a "vertical network" relationship as shown by the simple downward link. Topicalized use in complex clauses, combined with contexts featuring replicated reference to talk by hypothesis enabled reinterpretation and constructionalization as a Conjunct member of the Connector.Cxn in [D1 __ D2] contexts. The result of constructionalization is a break in the vertical network relationship and is represented as in Figures 13.6 and 13.7 by an angled upward arrow. In Figure 13.8 there is also an angled downward arrow from CircADv to DSM use. This is meant to reflect the kind of stretching of CircAdv use in context and gradual reassignment to Connector status hypothesized to occur in the process of conventionalization of replicated innovations. The replicated contextual factors are represented in curvy boxes. The new [[Conjunct] ↔ [DSM]] *by the way* 'in passing' was later reinterpreted as a DM, a constructional change as the shift concerns discourse function, not form. Later yet, a link to mitigation, a subschema of social markers (the SocMark.Cxn) was made. This is also a constructional change, the creation in this case of an associative connection to an external schema. None of these later changes involves a break in the symbolic links, and therefore no angled brackets link the changes.

The representation in Figure 1.8 uses the same notational diacritics as were used in Figures 13.6 and 13.7:

72. The 650 date reflects the period of the earliest documentation in English, but the Spatial. Schema pre-existed that date in earlier Germanic and Indo-European.

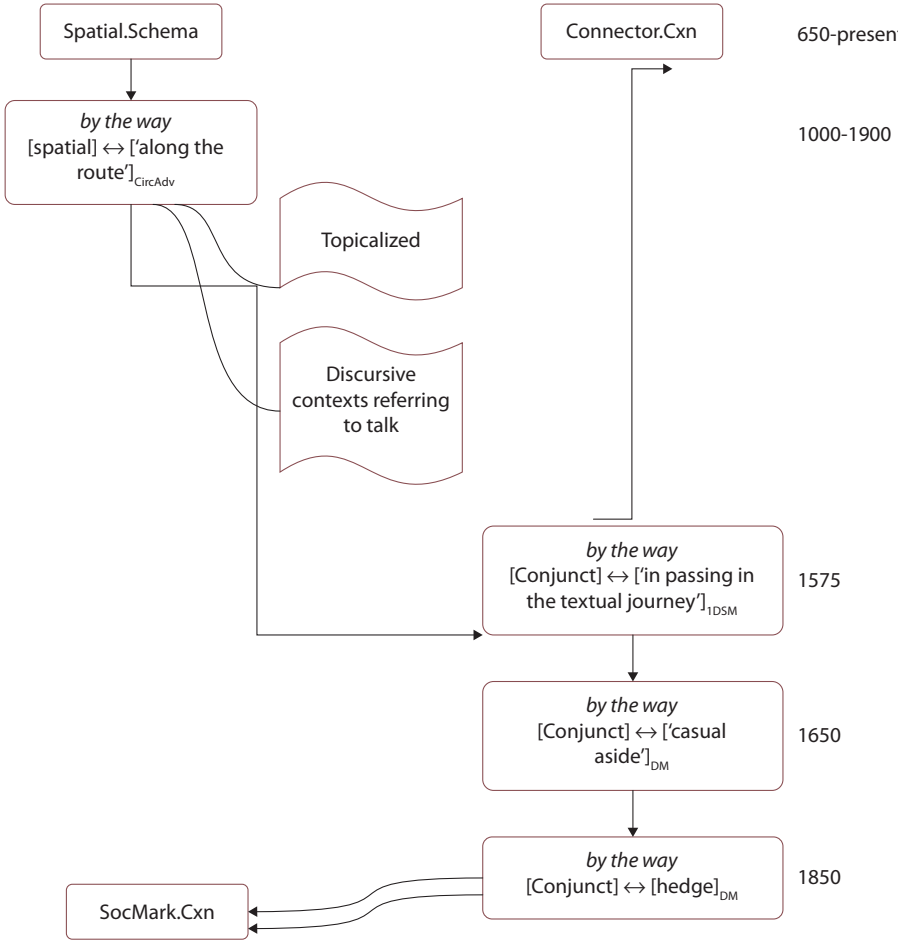

Figure 13.8 Partial sketch of changes in the development of *by the way* in terms of network connections and context

13.6 Conclusion

In this Chapter I have attempted to address the question posed in Smirnova and Sommerer (2020: 3) "How can node creation and node loss be implemented in the network model?" The network representations capture the hypothesis that specific micro-constructions are developed within the context of knowledge of related conceptual categories.

In the representations proposed, nodes and links are equally important. The nodes reflect observed patterns of use of both substantive and abstract *Es*, while the links reflect association with the conceptual domains in which these *E*s are used. Challenges for the future include the question: How can the role of networks in the constructicon and changes to them best be conceptualized?

CHAPTER 14

Conclusion and prospects

14.1 Introduction

In this final Chapter, I briefly summarize the main points of the book (Section 14.2). I then suggest some areas of further research that arise directly out of the work presented here (Section 14.3).

14.2 Summary of main points

The major objective has been to emphasize the importance of infusing more pragmatics into constructional thinking and suggesting ways to do so in terms of Diachronic Construction Grammar. The ubiquity in language use of primarily pragmatic expressions, such as illocutionary uses of speech act verbs like *promise, accuse* and discourse structuring Connectors like *after all* is part of a language user's knowledge of language and should be paid more attention in construction grammar accounts than has been customary in the past. How such expressions come into being and how they change over time also deserve more attention as a window into the dynamic processes of language production and processing.

As an example of the importance throughout the history of English of pragmatic expressions, I have developed a historical constructionalist perspective on the rise of Connectors that are pragmatic Discourse Structuring Markers in English and have offered a uniform account of the history of several of these types of expressions. This class of Connectors is often referred to as Discourse Markers (DMs). They signal a relationship between discourse segment 1 (D1) and discourse segment 2 (D2) (e.g. Fraser 1996, 2006). However, a key point developed in this book is that a single term is inadequate to account for such Connectors. They are on two gradients:

1. from primarily contentful to primarily pragmatic in meaning,
2. from monofunctional to multifunctional.

Connectors like *in addition, instead*, are on the relatively [+truth-conditional] pole and are monofunctional. I have called them 1DSMs. Connectors like *after all, by the way* in their contemporary uses are on the [–truth-conditional] pole. They are multifunctional. I have called these DMs.

A Discourse Structuring Trajectory Hypothesis was proposed in Chapter 4.4.2 and tested in later chapters. A refined version appears in Chapter 7.5 as Figure 7.4 and is repeated here as Figure 14.1:

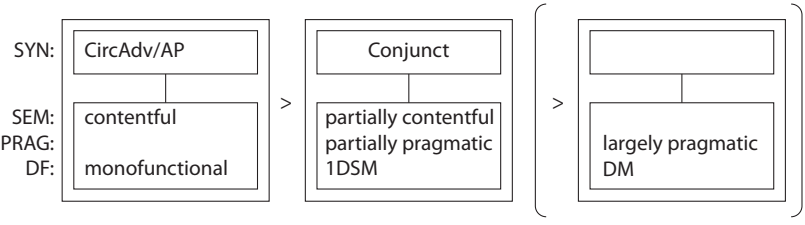

Figure 14.1 Discourse Structuring Marker Trajectory Hypothesis

The case studies showed that this hypothesis is robustly supported by the histories of a variety of DSMs in English. In all cases, an internal change was a necessary step in the development, typically use of a CircAdv in clause initial position, where it was ambiguous between a topicalized adverbial argument of the clause and a pre-clausal Connector. In many cases SP/Ws have not (yet) used [[Conjunct] ↔ [DSM]] constructions as DMs, but rather still use them in relatively contentful, monofunctional 1DSM ways. As Finkbeiner (2019b: 185) notes, "what kinds of pragmatic aspects play a role in what kinds of constructions is an empirical question and has to be specified individually for every construction in the analysis".

I have suggested how such specification can be achieved by using constructional features such as are proposed in Croft (2001). The distinction made there between three types of meaning: semantic, pragmatic and discourse functional has proved useful in specifying differences between Connectors at a very general level of analysis as well as stages of development. More fine-grained analysis can be built on these features. I have added communicative function (CF) in addition to discourse function (DF). To repeat Figure 13.3 in Chapter 13.2.1 as Figure 14.2, I have suggested that throughout the history of English, when a CircAdv is constructionalized as a 1DSM in PDE it will inherit the features in the Figure by default. Other features are construction-specific and dependent in part of the original content of the expression and in part on the DSM function to which a micro-construction is put.

A second objective, in addition to advocating for greater infusion of pragmatics in construction grammar, has been to suggest some issues that are of importance to understanding change from a constructional usage-based perspective. Following Schmid (2017) I have distinguished between innovations, shifts in use by individual speakers and writers, which come to be replicated and entrenched, and changes that are conventionalized within a group or community of speakers

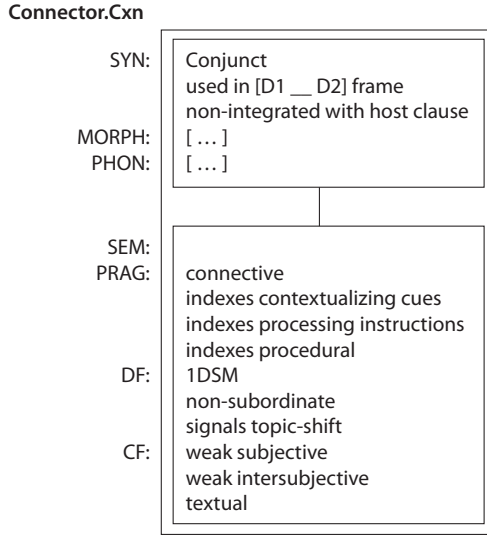

Figure 14.2 Model of characteristics of a connector.Cxn

and writers. There can be no change without innovation, but not all innovation leads to change. I have modified the characterization of the distinction made in Traugott and Trousdale (2013) between constructionalization, the creation of a new form-meaning pairing, and constructional changes, the modulation of either the form or the meaning of a construction. I have also tested the distinction and found that it accounts well for the difference between the development of [[Conjunct] ↔ [DSM]] pairings (constructionalization) and the later development of DM uses (constructional change).

A recurrent theme in the book is the importance of replicated assemblies of discourse contexts in enabling change. In work by Goldberg and others, utterances are conceptualized as exemplifying unification of a number of constructions, e.g. *Do you understand?* is conceptualized as unifying the proposition *you understand* with interrogative, subject-auxiliary inversion and present tense constructions, among others. This kind of thinking has led practitioners of construction grammar to conceptualize and recognize contexts that may enable change as sets of multiple discursive types. Drawing on Petré's (2019) hypothesis that assemblies of contextual uses precede constructionalization, I suggested that several types of replicated external discursive contexts may have enabled the development of Connector [[Conjunct] ↔ [DSM]] use. For example, assembled contexts that appear to have enabled the constructionalization of the dynamic spatial CircAdv *by the way* as a 1DSM include:

a. use in reports of talk along a route, in which the route is the background to the informationally more important talk,
b. use of the metaphor of argumentation as journey to signal the structure of the argumentation,
c. use in appositive relative clauses, in which the adverbial is focused.

Such contextual uses led to gradual modification of the meaning of the CircAdv *by the way* with which they were assembled. By contrast, post-constructionalization changes had to do with slight modulations in the degree of dismissiveness associated with 1DSM *by the way* when used in different positions, especially post-clausal vs. pre-clausal uses, and in the degree to which dismissiveness was conceptualized as fake hedging use. The distinction between assemblies of external constructional contexts prior to constructionalization and more local modulations post-constructionalization was not adequately developed in Traugott and Trousdale (2013). Evidence for it underscores the validity of the proposed contrast between constructionalization and constructional changes.

Obsolescence is one kind of post-constructionalization change that has not received much attention from a constructional perspective (but see Kuo 2020 in which obsolescence is discussed in late-stage grammatical change in Chinese). A usage-based way of thinking about constructional loss is that speakers cease to use a construction because they have come to prefer a different one instead. SP/Ws may have ceased to use *by the by* because the N *by* 'place' had become uncommon and did not support the construction with semantic transparency in the way that *way* 'route' supports *by the way* or *after* and *all* support DM *after all*. In these examples, although univerbated *E*s are (relatively) non-compositional, the phonology allows some access to the contentful *E*s from which the DMs arose. A hypothesis worth exploring is that when two or more constructions have approximately the same function, *E*s that are more compositional and allow some potential access to the contentful source are more robust and "prototypical" than those that are unanalyzable. The "dying" construction is marginalized and because it is (i) less accessible compositionally on the meaning dimension, (ii) less analyzable on the form dimension and (iii) less frequently used, it may eventually cease to be used. This kind of obsolescence usually occurs at the micro-constructional level. Sociocultural factors may be among other reasons for obsolescence. EModE (c1500–1700) is a period in which there is ample evidence of the rise of various types of Conjuncts with PM including DM functions (see Brinton 2008; Lutzky 2012). But it is also a period in which several epistemic *E*s such as *in truth* obsolesced, probably in tandem with socio-political changes brought about by the Protestant Reformation and the Renaissance (see e.g. Wierzbicka 2006; Bromhead 2009). Several individual

micro-constructions ceased to be used, but the Epistemic.Schema remained strong, and new adverbials like *certainly* came to be preferred. To what extent obsolescence is linked with sociocultural change deserves further study.

14.3 Some suggestions for further work

I have suggested questions for further work on particular issues in the relevant Chapters, e.g. questions regarding position are posed at the end of Chapter 12, in Section 12.5. Here I mention some general areas for further work that arise out of the present study.

First, how useful are the constructional models I have used for languages other than English? How can they be integrated with other models of resources in the constructicon that specify form-meaning pairings that are primarily syntactic and referential (see e.g. Perek and Patten 2019)?

Second, can the Discourse Structuring Marker Trajectory Hypothesis in Figure 14.1 be generalized to PMs in general, and to languages other than English? Similarly, is the set of default features in Figure 14.2 that arises when a Connector is constructionalized specific to English, or can it be generalized to other varieties of English and most especially to other languages?

Third, are any generalizations to be made about whether particular kinds of assemblies lead to the development of certain kinds of constructions, and under what circumstances? The importance of contexts referring to locution has been mentioned several times here, as well as in earlier work by Lenker (2010).

Fourth, regarding data. The origins and development of the [[Conjunct] ↔ [DSM]] pairings have been studied in terms of changes evidenced by three main corpora: EEBO, COHA, and COCA. These resources, large as they are, present only a part of the historical record of English, and only that part that has survived in manuscript and printed or transcribed form and has been selected for the corpus. As is well known, these corpora therefore provide insight into the linguistic knowledge of a privileged group of speakers who were literate. Other, mostly smaller, corpora that represent spoken language more closely should be investigated to enrich the findings, especially to test whether, as appears from EEBO, the historical development of a DM like *by the way* is largely grounded in translations from Latin, French, and Greek, and in "high style" philosophical and religious genres.

Lastly, while it may not be possible to reliably reconstruct sign language patterns or prosodic patterns from earlier times, some insights into what they may have been like, how to study them and how to model them in constructional terms will hopefully emerge from research that currently seeks on the one hand to understand

sign language from a cognitive perspective (e.g. Wilcox and Occhino 2016; Wilcox and Martinez 2020) and on the other to integrate prosody into construction grammar (e.g. Alm and Fischer 2021; Gras and Elvira-García 2021).

References

Data resources and corpora

CED *A Corpus of English Dialogues 1560–1760*. (2006). Compiled by M. Kytö, & J. Culpeper, in collaboration with T. Walker, & D. Archer. Uppsala University. 1.4 million words. http://www.helsinki.fi/varieng/CoRD/corpora/CED/ http://www.engelska.uu.se/corpus.html.

CEECS *Corpus of Early English Correspondence Sampler, 1414–1680*. (1998). Compiled by T. Nevalainen, H. Raumolin-Brunberg, J. Keränen, M. Nevala, A. Nurmi, & M. Palander-Collin. University of Helsinki. 450,000 words. http://www.helsinki.fi/varieng/domains/CEEC.html.

CLMET_3_0. *The Corpus of Late Modern English Texts*, version 3.0 compiled by H. de Smet, H.-J. Diller, & J. Tyrkkö. Leuven University. c. 34 million words 1710–1920. https://perswww.kuleuven.be/~u0044428/clmet3_0.htm.

COCA *The Corpus of Contemporary American English. 1990–2019*. Compiled by M. Davies. Brigham Young University. Release March 2020, 1. billion words. https://www.english-corpora.org/coca/.

COHA *Corpus of Historical American English. 1810–2009*. Compiled by M. Davies. Brigham Young University. 400 million words. (Superseded by *Corpus of Historical American English*. 1820s–2010s. Compiled by M. Davies. Brigham Young University. Release 2021, 475 million words. https://www.english-corpora.org/coha/.)

DOE *Dictionary of Old English A-I*. diPaolo Healey, A. et al. (Eds.). University of Toronto. https://www.doe.utoronto.ca/pages/index.html.

DOEC *Dictionary of Old English Corpus*. (2009). Original release 1981 compiled by A. Cameron, A. Crandell Amos, S. Butler & A. diPaolo Healey. Release 2009 compiled by A. diPaolo Healey, J.Holland, I. McDougall, & D. McDougall, with Xin Xiang. University of Toronto. c3 million running words of Old English, c1 million running words of Latin. http://www.helsinki.fi/varieng/CoRD/corpora/DOEC/index.html.

EEBO Davies, M. (2017). *Early English Books Online*. Part of the SAMUELS Project, available online at https://www.english corpora.org/eebo/.

Fisher Fisher Corpus of American Telephone Calls. See C. Cieri, D. Graff, O. Kimball, D. Miller, & K. Walker (2004, 2005). *Fisher English Training Speech, Part 1, Part 2. Transcripts*. Philadelphia, PA: Linguistic Data Consortium.

HC *Helsinki Corpus of English Texts. 730–1710*. (1991). Compiled by M. Rissanen (Project leader), M. Kytö (Project secretary); L. Kahlas-Tarkka, M. Kilpiö (Old English); S. Nevanlinna, I. Taavitsainen (Middle English); T. Nevalainen, H. Raumolin-Brunberg (Early Modern English). Department of English, University of Helsinki, 1.5 million words. http://www.helsinki.fi/varieng/CoRD/corpora/HelsinkiCorpus/

HGCW *Harvard's Geoffrey Chaucer Website*. https://chaucer.fas.harvard.edu/pages/literary-works

ICE-AUS *International Corpus of English – Australia*. (1991–1995). Compiled by P. Peters. Macquarie University, c1 million words. https://www.ausnc.org.au/corpora/ice.

ICE-GB *International Corpus of English-Great Britain*. International http://ice-corpora.net/ice/index.htm.

MED *The Middle English Dictionary*. (1956–2001). Ann Arbor: University of Michigan Press. https://quod.lib.umich.edu/m/middle-english-dictionary/dictionary.

NOW corpus *News on the Web*. (2010-present). https://www.english-corpora.org/now/.

OED *Oxford English Dictionary*. (2018). Oxford University Press. http://www.oed.com/.

OSS *OpenSource Shakespeare, An Experiment in Literary Technology, 2003–2005*, compiled by E. M. Johnson, George Mason University, http://www.opensourceshakespeare.org/.

SBCSAE *Santa Barbara Corpus of American Spoken English, Parts 1–4*. (2000–2005). Compiled by J. W. Du Bois, W. L. Chafe, C. Meyer, S. A. Thompson, R. Englebretson & N. Martey. Philadelphia: Linguistic Data Consortium.

List of references

Acton, E. K. (2019). Pragmatics and the social life of the English definite article. *Language* 95(1), 37–65. https://doi.org/10.1353/lan.2019.0010

Aijmer, K. (1986). Why is *actually* so popular in spoken English? In G. Tottie, & I. Bäcklund (Eds.), *English in speech and writing: A symposium* (pp. 119–127). Almqvist and Wiksell.

Aijmer, K. (1997). *I think*– an English modal particle. In T. Swan, & O. J. Westvik (Eds.), *Modality in Germanic languages: Historical and comparative perspectives* (pp. 1–47). Mouton de Gruyter. https://doi.org/10.1515/9783110889932.1

Aijmer, K. (2002). *English Discourse Particles. Evidence from a Corpus*. John Benjamins. https://doi.org/10.1075/scl.10

Alm, M., & Fischer, K. (2021). The problem of prosodic variability in the definition of constructions: The list construction. Paper presented at ICCG11, Antwerp, August.

Andersen, H. (2001). Actualization and the (uni)directionality. In H. Andersen (Ed.,) *Actualization: Linguistic change in progress* (pp. 225–248). John Benjamins. https://doi.org/10.1075/cilt.219.11and

Anthonissen, L. (2020). Cognition in construction grammar: Connecting individual and community grammars. In P. Petré, & L. Anthonissen (Eds.), *Constructionalist Approaches to Individuality in Language*. Special issue, Cognitive Linguistics 31(2), 185–212. https://doi.org/10.1515/cog-2019-0023

Anttila, R. (2003). Analogy: The warp and woof of cognition. In B. D. Joseph & R. D. Janda (Eds.), *The handbook of historical linguistics* (pp. 425–440). Blackwell. https://doi.org/10.1002/9780470756393.ch10

Ariel, M. (2019). *Or* constructions: Code, inference and cue too. In R. Finkbeiner (Ed.), *On the role of pragmatics in Construction Grammar*. Special issue, Constructions and Frames 11(2), 193–219. https://doi.org/10.1075/cf.00028.ari

Ariel, M., & Mauri, C. (2019). An 'alternative' core for *or*. *Journal of Pragmatics* 149, 40–59. https://doi.org/10.1016/j.pragma.2019.06.004

Athanasiadou, A., Canakis, C., & Cornillie, B. (Eds.) (2006). *Subjectification: Various paths to subjectivity*. Mouton de Gruyter. https://doi.org/10.1515/9783110892970

Auer, P. (1996). The pre-front field in spoken German and its relevance as a grammaticalization position. *Pragmatics* 6, 295–322. https://doi.org/10.1075/prag.6.3.03aue

Auer, P. (2005). Projection in interaction and projection in grammar. *Text* 25, 7–36.

Auer, P. & Maschler, Y. (eds). (2016). *NU/NÅ: A Family of Discourse Markers across the Language of Europe and Beyond*. De Gruyter. https://doi.org/10.1515/9783110348989

Auer, P., & Pfänder, S. (2011). Constructions: Emergent or emerging? In P. Auer, & S., Pfänder (Eds.), *Constructions: Emerging and Emergent* (pp. 1–21). De Gruyter Mouton.

Barðdal, J. (2008). *Productivity: Evidence from case and argument structure in Icelandic*. John Benjamins. https://doi.org/10.1075/cal.8

Barðdal, J., Smirnova, E., Sommerer, L., & Gildea, S. (Eds). (2015). *Diachronic Construction Grammar*. John Benjamins. https://doi.org/10.1075/cal.18

Beeching, K. (2005). Politeness-induced semantic change: The case of *quand-même*. *Language Variation and Change* 17(2), 155–180. https://doi.org/10.1017/S0954394505050076

Beeching, K., & Detges, U. (Eds.) (2014a). *Discourse functions at the left and right periphery: Crosslinguistic investigations of language use and language change*. Brill. https://doi.org/10.1163/9789004274822

Beeching, K., & Detges, U. (2014b). Introduction. In K. Beeching, & U. Detges (Eds.), *Discourse functions at the left and right periphery: Crosslinguistic investigations of language use and language change* (pp. 1–23). Brill.

Beeching, K., & Murphy, J. (2019). Doing (mock) im/politeness: Norms and variations in the use of politeness formulae. Special issue, *Journal of Pragmatics* 142, 201–206. https://doi.org/10.1016/j.pragma.2019.01.027

Benveniste, É. (1971[1958]). Subjectivity in language. In É. Benveniste, *Problems in general linguistics* (pp. 223–230), trans. by M. E. Meek. Coral Gables: FL: University of Miami Press.

Bergs, A., & Diewald, G. (2008). *Constructions and language change*. de Gruyter Mouton. https://doi.org/10.1515/9783110211757

Bergs, A., & Diewald, G. (Eds.) (2009a). *Contexts and constructions*. John Benjamins. https://doi.org/10.1075/cal.9

Bergs, A., & Diewald, G. (2009b). Contexts and constructions. In A. Bergs, & G. Diewald (Eds.), *Contexts and constructions* (pp. 1–14). John Benjamins. https://doi.org/10.1075/cal.9.01ber

Biber, D., & Finegan, E. (1988). Adverbial stance types in English. *Discourse Processes* 11, 1–34. https://doi.org/10.1080/01638538809544689

Biber, D., & Finegan, E. (1989). Styles of stance in English: Lexical and grammatical marking of evidentiality and effect. *Text* 9(1), 93–124.

Biber, D., & Gray, B. (2011). Grammatical change in the noun phrase: The influence of written language use. *English Language and Linguistics* 15(2), 223-250. https://doi.org/10.1017/S1360674311000025

Biber, D., & Gray, B. (2012). The competing demands of popularization vs. economy: Written language in the age of mass literacy. In T. Nevalainen, & E. C. Traugott (Eds.), *The Oxford handbook of the history of English* (pp. 314–328). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199922765.013.0028

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Pearson Education.

Biberauer, T., & Roberts, I. (2017). Parameter setting. In A. Ledgeway, & I. Roberts (Eds.), *The Cambridge handbook of historical syntax* (pp. 134–162). Cambridge University Press. https://doi.org/10.1017/9781107279070.008

Birner, B. J., & Ward, G. (1998). *Information status and non-canonical word order in English*. John Benjamins. https://doi.org/10.1075/slcs.40

Blakemore, D. (1987). *Semantic constraints on Relevance*. Blackwell.

Blakemore, D. (2002). *Relevance and linguistic meaning: The semantics and pragmatics of Discourse Markers*. Cambridge University Press. https://doi.org/10.1017/CBO9780511486456

Boas, H. C. (2013). Cognitive construction grammar. In T. Hoffmann, & G. Trousdale (Eds.), *The Oxford Handbook of Construction Grammar* (pp. 233–252). Oxford University Press.

Boas, H. C., & Sag, I. A. (Eds.) (2012). *Sign-based construction grammar*. CSLI Publications.

Bolinger, D. (1989). *Intonation and its Uses: Melody in grammar and discourse*. Edward Arnold.

Börjars, K., Vincent, N., & Walkden, G. (2015). On constructing a theory of grammatical change. *Transactions of the Philological Society* 113, 363–382. https://doi.org/10.1111/1467-968X.12068

Boye, K, & Harder, P. (2007). Complement-taking predicates: Usage and linguistic structure. *Studies in Language* 31(3), 596–606. https://doi.org/10.1075/sl.31.3.03boy

Boye, K., & Harder, P. (2012). A usage-based theory of grammatical status and grammaticalization. *Language* 88, 1–44. https://doi.org/10.1353/lan.2012.0020

Bréal, M. (1964[1900]). *Semantics: Studies in the Science of Meaning*, trans. by Mrs. H. Cust. Dover.

Breban, T. (2006). Grammaticalization and subjectification of the English adjectives of general comparison. In A. Athanasiadou, C. Canakis, & B. Cornillie (Eds.), *Subjectification: Various paths to subjectivity* (pp. 241–278). Mouton de Gruyter. https://doi.org/10.1515/9783110892970.241

Brems, L. (2011). *Layering of size and type noun constructions in English*. de Gruyter Mouton. https://doi.org/10.1515/9783110252927

Brems, L., Ghesquière, L., & Van de Velde, F. (2014[2012]a). *Intersubjectivity and intersubjectification in grammar and discourse*. John Benjamins. https://doi.org/10.1075/bct.65

Brems, L., Ghesquière, L., & Van de Velde, F. (2014[2012]b). Intersubjectivity and intersubjectification: Typology and operationalization. In L. Brems, L. Ghesquière, & F. Van de Velde (Eds.), *Intersubjectivity and intersubjectification in grammar and discourse* (pp. 129–153). John Benjamins. https://doi.org/10.1075/bct.65.07ghe

Brinton, L. J. (1996). *Pragmatic Markers in English: Grammaticalization and discourse functions*. Mouton de Gruyter. https://doi.org/10.1515/9783110907582

Brinton, L. (2006). Pathways in the development of Pragmatic Markers in English. In A. van Kemenade, & B. Los (Eds.), *The handbook of the history of English* (pp. 307–334). Blackwell. https://doi.org/10.1002/9780470757048.ch13

Brinton, L. J. (2008). *The comment clause in English: Syntactic origins and pragmatic development*. Cambridge University Press. https://doi.org/10.1017/CBO9780511551789

Brinton, L. J. (2017). *Evolution of Pragmatic Markers in English: Pathways of change*. Cambridge University Press. https://doi.org/10.1017/9781316416013

Briz, A., & Grupo Val.Es.Co. (2003). Un sistema de unidades para el estudio del lenguaje coloquial [A system of units for the study of colloquial language]. *Oralia* 6, 7–61.

Bromhead, H. (2009). *The reign of truth and faith: Epistemic expressions in 16th and 17th century English*. Mouton de Gruyter. https://doi.org/10.1515/9783110216028

Buchstaller, I., & Traugott, E. C. (2006). *The lady was al demonyak*: Historical aspects of adverb *all*. *English Language and Linguistics* 10, 345–370. https://doi.org/10.1017/S136067430600195X

Budts, S., & Petré, P. (2020). Putting connections center stage in Diachronic Construction Grammar. In L. Sommerer, & E. Smirnova (Eds.,) *Nodes and networks in Diachronic Construction Grammar* (pp. 317–351). John Benjamins. https://doi.org/10.1075/cal.27.09bud

Bybee, J. (2003). Mechanisms of change in grammaticization: The role of frequency. In B. D. Joseph, & R. D. Janda (Eds.), *The handbook of historical linguistics* (pp. 602–623). Blackwell. https://doi.org/10.1002/9780470756393.ch19

Bybee, J. L. (2010). *Language, usage and cognition*. Cambridge University Press. https://doi.org/10.1017/CBO9780511750526

Bybee, J. L., & Pagliuca, W. (1987). The evolution of future meaning. In A. Giacalone Ramat, O. Carruba, & G. Bernini (Eds.), *Papers from the 7th International Conference on Historical Linguistics* (pp. 109–122). John Benjamins. https://doi.org/10.1075/cilt.48.09byb

Bybee, J., Perkins, R., & Pagliuca, W. (1994). *The evolution of grammar: tense, aspect, and modality in the languages of the world*. University of Chicago Press.

Caffi, C. (2013). On mitigation. In M. Sbisà, & K. Turner (Eds.), *Pragmatics of speech actions* (pp. 258–286). Walter de Gruyter.

Campbell, L. (Ed.) (2001). Grammaticalization: A critical assessment. *Language Sciences* 23, Nos. 2–3.

Cappelle, B. (2006). Particle placement and the case for 'allostructions'. In D. Schönefeld (Ed.), *Constructions all over: Case studies and theoretical implications*. Special issue, *Constructions*. http://www.blogs.uni-osnabrueck.de/constructions/files/2014/06/2006-SI-Cappelle22-80-1-PB.pdf

Cappelle, B. (2017). What's pragmatics doing outside constructions? In I. Depraetere, & R. Salkie (Eds.), *Semantics and pragmatics: Drawing a Line* (pp. 115–151). Springer.

Carston, R. (2002). *Thoughts and utterances: The pragmatics of explicit communication*. Blackwell. https://doi.org/10.1002/9780470754603

Catford, J. C. (1965). *A linguistic theory of translation*. Oxford University Press.

Chaves, R. P. (2007). Coordinate structures: Constraint-based syntactic-semantic processing. Unpublished dissertation, University of Lisbon.

Chomsky N. (2005). Three factors in language design. *Linguistic Inquiry* 36(1), 1–22. https://doi.org/10.1162/0024389052993655

Clark, E. V., & Clark, H. H. (1979). When nouns surface as verbs. *Language* 55, 767–811. https://doi.org/10.2307/412745

Comrie, B. 1989[1981]. Language universals and linguistic typology: Syntax and morphology. *Chicago University Press*.

Conrad, S., & Biber, D. (2000). Adverbial marking of stance in speech and writing. In S. Hunston, & G. Thompson (Eds.), *Evaluation in text: Authorial stance and the construction of discourse* (pp. 56–73). Oxford University Press.

Coussé, E., Andersson, P. & Olofsson, J. (Eds.) (2018a). *Grammaticalization meets Construction Grammar*. John Benjamins. https://doi.org/10.1075/cal.21

Coussé, E., Andersson, P. & Olofsson, J. (2018b). Grammaticalization meets Construction Grammar: Opportunities, challenges and potential incompatibilities. In E. Coussé, P. Andersson, & J. Olofsson (Eds.), *Grammaticalization meets Construction Grammar* (pp. 3–19). John Benjamins. https://doi.org/10.1075/cal.21.c1

Croft, W. (2000). *Explaining language change*. Longman, Pearson Education.

Croft, W. (2001). *Radical Construction Grammar: Syntactic theory in typological perspective*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198299554.001.0001

Croft, W. (2005). Logical and typological arguments for Radical Construction Grammar. In J.-O. Östman & M. Fried (Eds.), *Construction grammars: Cognitive grounding and theoretical extensions* (pp. 273–314). John Benjamins. https://doi.org/10.1075/cal.3.11cro

Cuenca, M. J. (2015). Lexical connectives as grounding devices in political discourse. Paper presented at IPrA 14, Antwerp, Belgium.

Cuenca, M. J., & Bach, C. (2007). Contrasting the form and use of reformulation markers. *Discourse Studies* 9(2), 149–275. https://doi.org/10.1177/1461445607075347

Culpeper, J., & Kytö. M. (2010). *Early Modern English dialogues: Spoken interaction as writing*. Cambridge University Press.

Davidse, K., Vandelanotte, L., & Cuyckens, H. (Eds.) 2010. *Subjectification, intersubjectification and grammaticalization*. De Gruyter Mouton. https://doi.org/10.1515/9783110226102

Defour, T. (2007). A diachronic study of the pragmatic markers *well* and *now*. Fundamental research into semantic development and grammaticalisation by means of a corpus study. Doctoral dissertation, Ghent University.

Degand, L., & Evers-Vermeul, J. (2015). Grammaticalization or pragmaticalization of discourse markers? *Journal of Historical Pragmatics* 16(1), 59–85. https://doi.org/10.1075/jhp.16.1.03deg

Degand, L., & Fagard, B. (2011). *Alors* between discourse and grammar: The role of syntactic position. *Functions of Language* 18, 29–56. https://doi.org/10.1075/fol.18.1.02deg

Degand, L., & Simon-Vandenbergen, A.-M. (Eds.) (2011a.) *Grammaticalization, pragmaticalization and/or (inter)subjectification: Methodological issues for the study of Discourse Markers*. Special issue, *Linguistics* 49 (2).

Degand, L., & Simon-Vandenbergen, A. M. (2011b). Introduction: Grammaticalization and (inter)subjectification of Discourse Markers. In L. Degand, & A.-M. Simon-Vandenbergen (Eds.), *Grammaticalization, pragmaticalization and/or (inter)subjectification: Methodological issues for the study of Discourse Markers*. Special issue, *Linguistics* 49(2), 287–294. https://doi.org/10.1515/ling.2011.008

Dehé, N., & Wichmann, A. (2010). Sentence-initial *I think (that)* and *I believe (that)*: Prosodic evidence for use as main clause, comment clause and discourse marker. *Studies in Language* 34, 36–74. https://doi.org/10.1075/sl.34.1.02deh

Deppermann, A., & Günthner, S., (Eds.) (2015). *Temporality in interaction*. John Benjamins. https://doi.org/10.1075/slsi.27

Depraetere, I., & Salkie, R. (Eds.) (2017). *Semantics and pragmatics: Drawing a line*. Springer. https://doi.org/10.1007/978-3-319-32247-6

De Smet, H. (2009). Analysing reanalysis. *Lingua* 119, 1728–1755. https://doi.org/10.1016/j.lingua.2009.03.001

De Smet, H. (2012). The course of actualization. *Language* 88(4), 601–633. https://doi.org/10.1353/lan.2012.0056

De Smet, H., Ghesquière, L., & van de Velde, F. (Eds.) (2013). *On multiple source constructions in language change*. John Benjamins. https://doi.org/10.1075/bct.79

De Smet, H., & Verstraete., J.-C. (2006). Coming to terms with subjectivity. *Cognitive Linguistics* 17(3), 365–392. https://doi.org/10.1515/COG.2006.011

Detges, U. (2016). Does reanalysis need ambiguity? In M. Bauer, & N. Potysch (Eds.), *Ambiguity. An interdisciplinary approach*. Peter Lang.

Detges, U., & Waltereit, R. (2002). Grammaticalization vs. reanalysis: A semantic-pragmatic account of functional change in language. *Zeitschrift für Spachwissenschaft* 21, 151–195. https://doi.org/10.1515/zfsw.2002.21.2.151

Diessel, H. (2006). Demonstratives, joint attention, and the emergence of grammar. *Cognitive Linguistics* 17(4), 463–489. https://doi.org/10.1515/COG.2006.015

Diessel, H. (2015). Usage-based construction grammar. In E. Dąbrowska, & D. Divjak (Eds.), *Handbook of cognitive linguistics* (pp. 295–321). Mouton de Gruyter. https://doi.org/10.1515/9783110292022-015.

Diessel, H. (2017). Usage-based linguistics. In M. Aronoff (Ed.), *Oxford research encyclopedia of linguistics*. Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.363.

Diessel, H. (2019). *The grammar network: How linguistic structure is shaped by language use*. Cambridge University Press. https://doi.org/10.1017/9781108671040

Diewald, G. (2002). A model for relevant types of contexts in grammaticalization. In I. Wischer, & G. Diewald (Eds.), *New reflections on grammaticalization* (pp. 103–120). John Benjamins. https://doi.org/10.1075/tsl.49.09die

Diewald, G. (2011). Pragmaticalization (defined) as grammaticalization of discourse functions. *Linguistics* 49, 365–390. https://doi.org/10.1515/ling.2011.011

Diewald, G., & Smirnova, E. (2012). Paradigmatic integration. In K. Davidse, T. Breban, L. Brems and T. Mortelmans (Eds.), *Grammaticalization and language change: New reflections*, (pp. 111–134). John Benjamins. https://doi.org/10.1075/slcs.130.05die

Dostie, G. (2009). Discourse markers and regional variation in French: A lexico-semantic approach. In K. Beeching, N. Armstrong, & F. Gadet (Eds.), *Sociolinguistic variation in Contemporary French*, (pp. 201–214). John Benjamins. https://doi.org/10.1075/impact.26.15dos

Drew, P., & Heritage, J. (1992). Analyzing talk at work: An introduction. In P. Drew, & J. Heritage (Eds.), *Talk at work: Interaction in institutional settings* (pp. 3–65). Cambridge University Press.

Dunn. J. (2017). Learnability and falsifiability of construction grammars. *Proceedings of the Linguistic Society of America* 6(1). https://doi.org/10.3765/plsa.v2io.4009

Durkin, P. (2014) *Borrowed words: A history of loanwords in English*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199574995.001.0001

Eckardt, R. (2006). *Meaning change in grammaticalization: An enquiry into semantic reanalysis*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199262601.001.0001

Eckardt, R. (2009). APO: Avoid Pragmatic Overload. In M.-B. Mosegaard Hansen, & J. Visconti (Eds.), *Current trends in diachronic semantics and pragmatics* (pp. 21–41). Brill. https://doi.org/10.1163/9789004253216_003

Ehmer, O., & Rosemeyer, M. (2018). Inferences in interaction and language change. In O. Ehmer, & M. Rosemeyer (Eds.), *Inferences in interaction and language change*, Special issue, *Open Linguistics*, 4(1), 1–20. https://doi.org/10.1515/opli-2018-0026

Enfield, N. J. (2006). Heterosemy and the grammar-lexicon trade-off. In F. K. Ameka, A. Dench, & N. Evans (Eds.), *Catching language* (pp. 297–320). De Gruyter Mouton.

Erman, B. & Kotsinas, U.-B. (1993). Pragmaticalization: The case of *ba'* and *you know*. *Studier i Modernspråkvetenskap* 10, 76–93. Almqvist and Wiksell.

Evans, N., & Wilkins, N. (2000). In the mind's ear: The semantic extensions of perception verbs in Australian languages. *Language* 76(3), 546–592. https://doi.org/10.2307/417135

Evers-Vermeul, J., Degand, L., Fagard, B., & Mortier, L. (2011). Historical and comparative perspectives on subjectification: A corpus-based analysis of Dutch and French causal connectives. *Linguistics* 49(2), 445–478. https://doi.org/10.1515/ling.2011.014

Fagard, B., & Sarda, L. (2014). From local adverbials to discourse markers: Three case studies in the diachrony of French. *Pragmatic Approaches to Text Structuring*, halshs-01242141.

Fauconnier, G. (2008). Pragmatics and cognitive linguistics. In L. R. Horn, & G. Ward (Eds.), *The handbook of pragmatics* (pp. 657–674). Blackwell.

Fetzer, A. (2012). Context in interaction: Relating pragmatic wastebaskets. In R. Finkbeiner, J. Meibauer, & P. B. Schumacher (Eds.), *What is a context? Linguistic approaches and challenges* (pp. 105–127). John Benjamins. https://doi.org/10.1075/la.196.08fet

Fillmore, C. J. (1976). Frame semantics and the nature of language. In S. R. Harnad, H. D. Steklis, & J. Lancaster (Eds.), *Origins and evolution of language and speech* (pp. 20–32). New York Academy of Sciences. https://doi.org/10.1111/j.1749-6632.1976.tb25467.x

Fillmore, C. J., Kay, P., & O'Connor, M. C. (1988). Regularity and idiomaticity in grammatical constructions. *Language* 64, 501–538. https://doi.org/10.2307/414531

Fillmore, C. J., & Baker, C. F. (2001). Frame semantics for text understanding. *Proceedings of WordNet and other lexical resources workshop* (pp. 59–63). NAACL.

Finkbeiner, R. (Ed.). (2019a). *On the role of pragmatics in Construction Grammar*. Special issue, *Constructions and Frames* 11(2). https://doi.org/10.1075/cf.00027.fin

Finkbeiner, R. (2019b). Reflections on the role of pragmatics in Construction Grammar. In R. Finkbeiner (Ed.,) *On the role of pragmatics in Construction Grammar*. Special issue, *Constructions and Frames* 11(2), 171–192. https://doi.org/10.1075/cf.00027.fin

Fischer, K. (Ed.) (2006). *Approaches to Discourse Particles*. Elsevier. https://doi.org/10.1163/9780080461588

Fischer, K. (2017). Cognitive linguistics and pragmatics. In B. Dancygier (Ed.), *The Cambridge handbook of cognitive linguistics* (pp. 330–346). Cambridge University Press. https://doi.org/10.1017/9781316339732.021

Fischer, O. (2007). *Morphosyntactic change: Functional and formal perspectives*. Oxford University Press.

Fischer, O. (2011). Grammaticalization as analogically driven change? In H. Narrog, & B. Heine (Eds.), *The Oxford handbook of grammaticalization* (pp. 31–42). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199586783.013.0003

Fitzmaurice, S. (2010). Coalitions, networks, and discourse communities in Augustan England: The Spectator and the early eighteenth-century essay. In R. Hickey (Ed.), *Eighteenth-century English: Ideology and change* (pp. 106–132). Cambridge University Press. https://doi.org/10.1017/CBO9780511781643.008

Flach, S. (2020). Constructionalization and the Sorites paradox. In L. Sommerer, & E. Smirnova (Eds.), *Nodes and networks in Diachronic Construction Grammar* (pp. 45–67). John Benjamins. https://doi.org/10.1075/cal.27.01fla

Foolen, A. (1996). Pragmatic particles. In J. Verschueren, J.-O. Östman, J. Blommaert, & C. Bulcan (Eds.), *Handbook of pragmatics* 1996 (pp. 1–24). John Benjamins. https://doi.org/10.1075/hop.2.pra3

Frank-Job, B. (2006). A dynamic-interactional approach to Discourse Markers. In K. Fischer (Ed.), *Approaches to Discourse Particles* (pp. 395–415). Elsevier.

Fraser, B. (1988). Types of English discourse markers. *Acta Linguistica Hungarica* 38, 19–33.

Fraser, B. (1996). Pragmatic Markers. *Pragmatics* 6, 167–190. https://doi.org/10.1075/prag.6.2.03fra

Fraser, B. (2006). Towards a theory of Discourse Markers. In K. Fischer (Ed.,) *Approaches to Discourse Particles* (pp. 189–204). Elsevier.

Fraser, B. (2009a). Topic orientation markers. *Journal of Pragmatics* 41(5), 892–898. https://doi.org/10.1016/j.pragma.2008.08.006

Fraser, B. (2009b). An account of discourse markers. *International Review of Pragmatics* 1, 293–320. https://doi.org/10.1163/187730909X12538045489818

Fraser, B. (2010). Pragmatic competence: The case of hedging. In G. Kaltenböck, W. Mihatsch, & S. Schneider (Eds.), *New approaches to hedging*. (pp. 15–34). Brill.

Fraser, B. (2015). The combining of Discourse Markers: A beginning. *Journal of Pragmatics* 86, 48–53. https://doi.org/10.1016/j.pragma.2015.06.007

Fraser, B., & Traugott, E. C. (2017). *But yet, see now this is another kind of Catch22*. A study of metatextual marker sequences. Paper presented at IPrA 15, July 16–21, Belfast.

Frawley, W. (2013[1992]). *Linguistic semantics*. Routledge.

Fried, M., & Nikiforidou, K. (Eds). (2015). *On the interaction of constructions with register and genre*. Special issue, *Constructions and Frames* 7(2).

Fried, M., & Östman, J.-O. (2005). Construction grammar and spoken language: The case of pragmatic particles. *Journal of Pragmatics* 37, 1752–1778. https://doi.org/10.1016/j.pragma.2005.03.013

Geeraerts, D. (1997). *Diachronic prototype semantics: A contribution to historical lexicology*. Clarendon Press.

Geeraerts, D. (2003). Decontextualizing and recontextualizing tendencies in 20th-century linguistics and literary theory. In E. Mengel, H.-J. Schmid, & M. Steppat (Eds.), *Anglistentag 2002 Bayreuth* (pp. 369–379). Wissenschaftlicher Verlag.

Geeraerts, D., & Cuyckens, H. (2007). *The Oxford handbook of cognitive linguistics*. Oxford University Press.

Ghesquière, L. (2010). On the subjectification and intersubjectification paths followed by the adjectives of completeness. In K. Davidse, L. Vandelanotte, & H. Cuyckens (Eds.), *Subjectification, intersubjectification and grammaticalization* (pp. 277–313). De Gruyter Mouton. https://doi.org/10.1515/9783110226102.3.277

Givón, T. (2018[1979]). *On understanding grammar*. John Benjamins, 2nd, rev. edn. https://doi.org/10.1075/z.213

Goldberg, A. E. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Goldberg, A. E. (2003). Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences* 7, 219–224. https://doi.org/10.1016/S1364-6613(03)00080-9

Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford University Press.

Goldberg, A. E. (2013). Constructionist approaches. In T. Hoffmann, & G. Trousdale (Eds.), *The Oxford handbook of construction grammar* (pp. 15–31). Oxford University Press.

Goldberg, A. E. (2019). *Explain me this*. Princeton University Press.

Goldberg, J. A. (1980). Discourse particles: An analysis of the role of *y'know, I mean, well* and *actually* in conversations. Unpublished dissertation, Cambridge University.

Goldstein, D. (2014). Wackernagel's Law I. In G. Giannakis et al. (Eds). *Encyclopedia of ancient Greek language and linguistics* (pp. 508–513). Brill.

Gras, P., & Elvira-García, W. (2021). The role of intonation in Construction Grammar: On prosodic constructions. *Journal of Pragmatics* 180, 234–247. https://doi.org/10.1016/j.pragma.2021.05.010

Greenbaum, S. (1969). *Studies in English adverbial usage*. Arnold.

Gregory, M. L., & Michaelis, L. A. (2001). Topicalization and left-dislocation: A functional opposition revisited. *Journal of Pragmatics* 33, 1665–1706. https://doi.org/10.1016/S0378-2166(00)00063-1

Grice, H. P. (1989[1967]). Logic and conversation. In H. P. Grice, *Studies in the way of words* (pp. 22–40). Harvard University Press.

Gries, S. Th., & Hilpert, M. (2012). Variability-based neighbor clustering: A bottom-up approach to periodization in historical linguistics. In T. Nevalainen, & E. C. Traugott (Eds.), *The Oxford handbook of the history of English* (pp. 134–144). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199922765.013.0014

Grondelaers, S., & Geeraerts, D. (2003). Towards a pragmatic model of cognitive onomasiology. In H. Cuyckens, R. Dirven, & J. R. Taylor (Eds.), *Cognitive approaches to lexical semantics* (pp. 67–92). De Gruyter Mouton. https://doi.org/10.1515/9783110219074.67

Grondelaers, S., Speelman, D., & Geeraerts, D. (2007). Lexical variation and change. In D. Geeraerts, & H. Cuyckens (Eds.), *The Oxford handbook of cognitive linguistics* (pp. 988–1011). Oxford University Press.

Gyselinck, E. (2020). (Re)shaping the constructional network: Modeling shifts and reorganizations in the network hierarchy. In L. Sommerer, & E. Smirnova (Eds.), *Nodes and networks in Diachronic Construction Grammar* (pp. 107–140). John Benjamins. https://doi.org/10.1075/cal.27.03gys

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Longman.

Hancil, S. (2016). Final *but*, theticality and subjectification. *Anglophonia* 22. https://doi.org/10.4000/anglophonia.1043

Hancil, S., Haselow. A., & Post, M. (Eds.) (2015). *Final particles*. De Gruyter Mouton. https://doi.org/10.1515/9783110375572

Hansen, M.-B. M. (1998). *The function of Discourse Particles. A study with special reference to Spoken Standard French*. John Benjamins. https://doi.org/10.1075/pbns.53

Hansen, M.-B. M. (2008). *Particles at the semantics/pragmatics interface: Synchronic and diachronic issues: A Study with special reference to the French phasal adverbs*. Elsevier.

Hansen, M.-B. M. (2012). The semantics of pragmatic expressions. In H.-J. Schmid (Ed.), *Cognitive pragmatics* (pp. 589–613). De Gruyter Mouton. https://doi.org/10.1515/9783110214215.587

Hansen, M.-B. M., & Visconti, J. (2009). Current trends in diachronic semantics and pragmatics. In M.-B. M. Hansen, & J. Visconti (Eds.), *Current trends in diachronic semantics and pragmatics* (pp. 1–19.) Brill.

Hansen, M.-B. M., & Waltereit, R. (2006). GCI theory and language change. *Acta lingvistica hafniensia* 38, 235–268. https://doi.org/10.1080/03740463.2006.10412210

Harder, P. (2012). Emergent and usage-based models of grammar. In H.-J. Schmid (Ed.), *Cognitive pragmatics* (pp. 507–532). De Gruyter. https://doi.org/10.1515/9783110214215.507

Harris, A., & Campbell, L. (1995). *Historical syntax in cross-linguistic perspective*. Cambridge University Press. https://doi.org/10.1017/CBO9780511620553

Haselow, A. (2012). Discourse organization and the rise of final *then* in the history of English. In I. Hegedüs, & A. Fodor (Eds.), *English historical linguistics 2010. selected papers from the sixteenth international conference on English historical linguistics (ICEHL 16), Pécs, 23–27 August* 2010, (pp. 153–175). John Benjamins. https://doi.org/10.1075/cilt.325.07has

Haselow, A. (2013). Arguing for a wide conception of grammar: The case of final particles in spoken discourse. *Folia Linguistica* 47,375–424. https://doi.org/10.1515/flin.2013.015

Haselow, A. (2015). Left vs. right periphery in grammaticalization: The case of *anyway*. In A. D. M. Smith, G. Trousdale, & R. Waltereit (Eds.), *New Directions in Grammaticalization Research*, 157–186. John Benjamins. https://doi.org/10.1075/slcs.166.08has

Haselow, A. (2016). A processual view on grammar: Macrogrammar and the final field in spoken syntax. *Language Sciences* 54, 77–101. https://doi.org/10.1016/j.langsci.2015.12.001

Haselow, A. (2019). Discourse marker sequences: Insights into the serial order of communicative tasks in real-time turn production. *Journal of Pragmatics* 146, 1–18. https://doi.org/10.1016/j.pragma.2019.04.003

Haspelmath, M. (1998). Does grammaticalization need reanalysis? *Studies in Language* 22, 315–351. https://doi.org/10.1075/sl.22.2.03has

Haspelmath, M. (1999). Why is grammaticalization irrecersible? *Linguistics* 37, 1043-1068. https://doi.org/10.1515/ling.37.6.1043

Haspelmath, M. (2004). On directionality in language change with particular reference to grammaticalization. In O. Fischer, M. Norde, & H. Perridon (Eds.), *Up and Down the Cline-The Nature of Grammaticalization* (pp. 17–44). John Benjamins. https://doi.org/10.1075/tsl.59.03has

Hasselgård, H. (2010). *Adjunct adverbials in English*. Cambridge University Press. https://doi.org/10.1017/CBO9780511676253

Hata, K. (2016). On the importance of the multimodal approach to discourse markers: A pragmatic view. *International Review of Pragmatics* 8(1), 36–54. https://doi.org/10.1163/18773109-00801002

Hawkins, J. A. (1983). *Word order universals*. Academic Press.

Hawkins, J. A. (2004). *Efficiency and complexity in grammars*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199252695.001.0001

Heine, B. (2002). On the role of context in grammaticalization. In I. Wischer, & G. Diewald (Eds.), *New reflections on grammaticalization* (pp. 83–101). John Benjamins. https://doi.org/10.1075/tsl.49.08hei

Heine, B. (2013). On discourse markers: Grammaticalization, pragmaticalization, or something else? *Linguistics* 51(6), 1205–1247. https://doi.org/10.1515/ling-2013-0048

Heine, B. (2019). Some observations on the dualistic nature of discourse processing. *Folia Linguistica* 52(2), 411–442. https://doi.org/10.1515/flin-2019-2016

Heine, B., Claudi, U., & Hünnemeyer, F. (1991). *Grammaticalization: A conceptual framework*. University of Chicago Press.

Heine, B., Kaltenböck, G., Kuteva, T., & Long, H. (2017). Cooptation as a discourse strategy. *Linguistics* 55(4), 813–855. https://doi.org/10.1515/ling-2017-0012

Heine, B., Kaltenböck, G., Kuteva, T., & Long, H. (In press). *The rise of Discourse Markers*. Cambridge University Press.

Heine, B., & Miyashita, H. (2008). Accounting for a functional category: German *drohen* 'to threaten'. *Language Sciences* 30, 53–101. https://doi.org/10.1016/j.langsci.2007.05.003

Heine, B., & Reh, M. (1984). *Grammaticalization and reanalysis in African Languages*. Buske.

Heritage, J. (1994). A change-of-state token and aspects of its sequential placement. In J. M. Atkinson, & J. Heritage (Eds.), *Structures of social action* (pp. 299–345). Cambridge University Press.

Heritage, J. (2002). *Oh*-prefaced responses to assessments: A method of modifying agreement/disagreement. In C. Ford, B. Fox, & S. Thompson (Eds.), *The language of turn and sequence* (pp. 196–224). Oxford University Press.

Hernández-Campoy, J. M., & Conde-Silvestre, J. C. (Eds.) (2012). *The handbook of historical sociolinguistics*. Wiley-Blackwell. https://doi.org/9781118257227

Hilpert, M. (2008). *Germanic future constructions: A usage-based approach to language change*. John Benjamins. https://doi.org/10.1075/cal.7

Hilpert, M. (2013). *Constructional change in English: Developments in allomorphy, word-formation and syntax*. Cambridge University Press. https://doi.org/10.1017/CBO9781139004206

Hilpert, M. (2018). Three open questions in Diachronic Construction Grammar. In E. Coussé, J. Olofsson, & P. Andersson (Eds.), *Grammaticalization meets construction grammar*, 22–39. John Benjamins. https://doi.org/10.1075/cal.21.c2

Himmelmann, N. P. (2004). Lexicalization and grammaticization: Opposite or orthogonal? In W. Bisang, N. P. Himmelmann, & B. Wiemer (Eds.), *What makes grammaticalization – A look from its fringes and its components* (pp. 21–42). Mouton de Gruyter.

Hoffmann, S. (2004). Using the OED quotations database as a corpus – a linguistic appraisal. *ICAME* 28(4), 17–30.

Hoffmann, T. (2020). What would it take for us to abandon Construction Grammar? Falsifiability, confirmation bias and the future of the constructionist enterprise. *Belgian Journal of Linguistics* 34, 149–161. https://doi.org/10.1075/bjl.00042.hof

Hoffmann, T., & Trousdale, G. (Eds.) (2013). *The Oxford handbook of construction grammar*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195396683.001.0001

Hopper, P. J. (1987). Emergent grammar. In J. Aske, N. Berry, L. Michaelis, & H. Filip (Eds.), *Berkeley Linguistics Society 13: General session and parasession on grammar and cognition* (pp. 139–157). Berkeley Linguistics Society. https://doi.org/10.3765/bls.v13i0.1834

Hopper, P. J. (1991). On some principles of grammaticization. In E. C. Traugott, & B. Heine (Eds.), *Approaches to grammaticalization*, Vol. 1, 17–35. John Benjamins. https://doi.org/10.1075/tsl.19.1.04hop

Hopper, P. J. (2008). Emergent serialization in English: Pragmatics and typology. In J. Good (Ed.), *Linguistic universals and language change* (pp. 252–284). Oxford University Press.

Hopper, P. J. (2011). Emergent grammar and temporality in interactional linguistics. In P. Auer, & S. Pfänder (Eds.), *Constructions: Emerging and emergent* (pp. 22–44). De Gruyter Mouton. https://doi.org/10.1515/9783110229080.22

Hopper, P. J., & Traugott, E. C. (2003[1993]). *Grammaticalization*. Cambridge University Press, 2nd, rev. edn.

Huddleston, R., & Pullum, G. K. (2002). *The Cambridge grammar of the English language*. Cambridge University Press. https://doi.org/10.1017/9781316423530

Huddleston, R., Payne, J., & Peterson, P. (2002). Coordination and supplementation. In R. Huddleston, & G. K. Pullum, *The Cambridge grammar of the English language*, Chapter 15. Cambridge University Press. https://doi.org/10.1017/9781316423530.016

Hudson, R. (2007). *Language networks: The new Word Grammar*. Oxford University Press.

Hudson, R. (2010). *An introduction to Word Grammar*. Cambridge University Press. https://doi.org/10.1017/CBO9780511781964

Hudson, R. (2015). Word Grammar. In B. Heine, & H. Narrog, *The Oxford handbook of linguistic analysis*, 2nd edn. Published online. https://doi.org/10.1093/oxfordhb/9780199677078.013.0033.

Hüning, M., Booij, G. (2014). From compounding to derivation: The rise of derivational affixes through constructionalization. In F. von Mengden, & H. Simon (Eds.), *Refining grammaticalization*. Special issue, *Folia Linguistica* 48(2), 579–604. https://doi.org/10.1515/flin.2014.019

Ingham, R. (2015). Spoken and written register differentiation in pragmatic and semantic functions in two Anglo-Norman corpora. In J. Gippert, & R. Gehrke (Eds.), *Historical corpora: Challenges and perspectives* (pp. 269–280). Narr.

Jakobson, R. (1956). Two aspects of language and two types of aphasic disturbance. In R. Jakobson, & M. Halle (Eds.), *Fundamentals of language* (pp. 53–87). Mouton.

Jaszczcolt, K. M. (2019). Rethinking being Gricean: New challenges for metapragmatics. *Journal of Pragmatics* 145: 15–24. https://doi.org/10.1016/j.pragma.2019.01.024

Jiménez, S. S., Estellés Arguedas, M., & Pons Bordería, S. (2018). Beyond the notion of *periphery*: An account of polyfunctional discourse markers within the Val.Es.Co model of discourse. In K. Beeching, C. Ghezzi, & P. Molinelli (Eds.), *Positioning the self and others: Linguistic perspectives* (pp. 105–125). John Benjamins. https://doi.org/10.1075/pbns.292.05sal

Joseph, B. D. (1997). On the linguistics of marginality: The centrality of the periphery. *Chicago Linguistic Society* 33, 197–213.

Joseph, B. D., & Janda, R. D. (2003). On language, change, and language change – or, of history, linguistics, and historical linguistics. In B. D. Joseph, & R. D. Janda (Eds.), *The handbook of historical linguistics* (pp. 3–180). Blackwell.

Kaltenböck, G., & Heine, B. (2014). Sentence grammar vs. thetical grammar: Two competing domains. In B. MacWhinney, A. Malchukov, & E. Moravcsik (Eds.), *Competing motivations in grammar and usage* (pp. 348–363). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198709848.003.0021

Kaltenböck, G., Heine, B., & Kuteva, T. (2011). On thetical grammar. *Studies in Language* 35(4), 852–897. https://doi.org/10.1075/sl.35.4.03kal

Kay, P., & Fillmore, C. J. (1999). Grammatical constructions and linguistic generalizations: The *What's X doing Y* construction. *Language* 75, 1–33. https://doi.org/10.2307/417472

Kemenade, A. van. (2012). Rethinking the loss of verb second. In T. Nevalainen, & E. C. Traugott (Eds.), *The Oxford handbook of the history of English* (pp. 822–834). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199922765.013.0067

Kemenade, A. van. (2021). Syntactic change and pragmatic maintenance: The discourse particle then over the history of English. *Paper presented at the 22nd Diachronic Syntax Conference*, Konstanz, May 24th.

Kemenade, A. van, & Links, M. (2020). Discourse particles in early English: Clause structure, pragmatics and discourse management. *Glossa a Journal of General Linguistics* 5(1). https://doi.org/10.5334/gjgl.1020

Kemmer, S., & Barlow, M. (1999). Introduction: A usage-based conception of language. In M. Barlow, & S. Kemmer (Eds.), *Usage-based models of language* (pp. vii–xxviii). CSLI publications.

Kempson, R. (1975). *Presupposition and the delimitation of semantics*. Cambridge University Press.

Kim, J.-B., & Davies, M. (2020). English *what with* absolute constructions: A Construction Grammar perspective. *English Language and Linguistics* 24(4), 637–666. https://doi.org/10.1017/S1360674319000169

Kim, J.-B., & Sag, I. A. (2005). English object extraposition: A constraint-based approach. In S. Müller (Ed.), *Proceedings of the HPSG05 Conference* (pp. 192–212). CSLI Publications. https://doi.org/10.21248/hpsg.2005.11

Kiparsky, P. (1968). Linguistic universals and linguistic change. In E. Bach, & R. T. Harms (Eds.), *Universals in linguistic theory* (pp. 171–202). Holt, Rinehart and Winston.

Kiparsky, P. (2012). Grammaticalization as optimization. In D. Jonas, J. Whitman, & A. Garrett (Eds.), *Grammatical change: Origins, nature, outcomes* (pp. 15–51). Oxford University Press.

König, E. (1985). Where do concessives come from? On the development of concessive connectives. In J. Fisiak (Ed.), *Historical semantics and historical word formation* (pp. 263–282). Mouton de Gruyter. https://doi.org/10.1515/9783110850178.263

König, E. (2020). Beyond exophoric and endophoric uses: Additional discourse-functions of demonstratives. In Å. Ness, A. Margetts, & Y. Treis (Eds.), *Demonstratives in Discourse* (pp. 21–42). Language Science Press.

Koops, C., & Lohmann, A. (2015). A quantitative approach to the grammaticalization of discourse markers: Evidence from their sequencing behavior. *International Journal of Corpus Linguistics* 20, 232–259. https://doi.org/10.1075/ijcl.20.2.04koo

Kroon, C. (1995). *Discourse particles in Latin: A study of NAM, ENIM, AUTEM, VERO and AT*. J.C. Gieben. https://doi.org/10.1163/9789004408999

Kuo, Y. H. (2020). Late-stage grammatical change in Chinese: A constructional account. PhD thesis, Edinburgh University.

Kuteva, T. (2001). *Auxiliation: An enquiry into the nature of grammaticalization*. Oxford University Press.

Kuteva, T., Heine, B, Hong, B., Long, H., Narrog, H., & Rhee, S. (2019). *World lexicon of grammaticalization*. Cambridge University Press, 2nd edn. https://doi.org/10.1017/9781316479704

Lakoff, G., & Johnson, M. (2003[1980]). *Metaphors we live by*. Chicago University Press, 2nd edn.

Lakoff, R. (1971). If's, and's, and but's about conjunction. In C. J. Fillmore, & D. T. Langendoen (Eds.), *Studies in linguistic semantics* (pp. 115–150). Holt, Rinehart and Winston.

Langacker, R. W. (1977). Syntactic reanalysis. In C. N. Li (Ed.), *Mechanisms of syntactic change* (pp. 57–139). University of Texas Press. https://doi.org/10.7560/750357-005

Langacker, R. W. (1987). *Foundations of cognitive grammar, Vol. I: Theoretical prerequisites*. Stanford University Press.

Langacker, R. W. (1988). A usage-based model. In B. Rudzka-Ostyn (Ed.), *Topics in cognitive linguistics* (pp. 127–161). John Benjamins. https://doi.org/10.1075/cilt.50.06lan

Langacker, R. W. (1990). Subjectification. *Cognitive Linguistics* 1, 5–38. https://doi.org/10.1515/cogl.1990.1.1.5

Langacker, R. W. (1991). *Foundations of cognitive grammar, Vol. II: Descriptive application*. Stanford University Press.

Langacker, R. W. (1999). Losing control: Grammaticalization, subjectification, and transparency. In A. Blank, & P. Koch (Eds.), *Historical semantics and cognition* (pp. 47–175). Mouton de Gruyter.

Langacker, R. W. (2006). Subjectification, grammaticization, and conceptual archetypes. In A. Athanasiadou, C. Canakis, & B. Cornillie (Eds.), *Subjectification: Various paths to subjectivity* (pp. 17–40). Mouton de Gruyter. https://doi.org/10.1515/9783110892970.17

Lass, R. (2000). Language periodization and the concept "Middle". In I. Taavitsainen, T. Nevalainen, P. Pahta, & M. Rissanen (Eds.), *Placing Middle English in context* (pp. 7–42). Mouton de Gruyter. https://doi.org/10.1515/9783110869514.7

Lehmann, C. (2015[1995]). *Thoughts on Grammaticalization*. Language Science Press, 3rd rev. edn.

Lenk, U. (1998). *Marking discourse coherence: Functions of Discourse Markers in spoken English*. Tübingen: Narr.

Lenker, U. (2010). *Argument and rhetoric. Adverbial connectors in the history of English*. De Gruyter Mouton. https://doi.org/10.1515/9783110216066

Lenker, U. (2014). Knitting and splitting information: Medial placement of linking adverbials in the history of English. In S. E. Pfenninger, O. Timofeeva, A.-C. Gardner, A. Honkapohja, M. Hundt, & D. Schreier (Eds.), *Contact, variation and change in the history of English* (pp. 11–38). John Benjamins. https://doi.org/10.1075/slcs.159.02len

Levinson, S. C. (2000). *Presumptive meanings: The theory of Generalized Conversational Implicature*. MIT Press, Bradford. https://doi.org/10.7551/mitpress/5526.001.0001

Levinson, S. C. (2003). Contextualizing "contextualization cues". In S. L. Eerdmans, C. L. Prevignano, & P. J. Thibault (Eds.), *Language and interaction: Discussions with John J. Gumperz* (pp. 31–39). John Benjamins. https://doi.org/10.1075/z.117.04lev

Lewis, D. M. (2000). Some emergent discourse connectives in English: Grammaticalization via rhetorical patterns. Unpublished Ph.D. dissertation, University of Oxford.

Lewis, D. M. (2007). From temporal to contrastive and causal: The emergence of connective *after all*. In A. Celle, & R. Huart (Eds.), *Connectives as discourse landmarks* (pp. 88–99). John Benjamins. https://doi.org/10.1075/pbns.161.09lew

Lewis, D. M. (2011). A discourse-constructional approach to the emergence of discourse markers in English. In L. Degand, & A.-M. Simon-Vandenbergen (Eds.), *Grammaticalization, pragmaticalization and/or (inter)subjectification: Methodological issues for the study of Discourse Markers*. Special issue, *Linguistics* 49(2), 415–443. https://doi.org/10.1515/ling.2011.013

Lightfoot, D. (1979). *Principles of diachronic syntax*. Cambridge University Press.

Lightfoot, D. (1991). *How to set parameters: Arguments from language change*. MIT Press.

Lohmann, A., & Koops, C. (2016). Aspects of discourse marker sequencing – Empirical challenges and theoretical implications. In G. Kaltenböck, E. Keizer, & A. Lohmann (Eds.), *Outside the clause: Form and function of extra-clausal constituents* (pp. 417–446). John Benjamins. https://doi.org/10.1075/slcs.178.14loh

López-Couso, M. J. (2010). Subjectification and intersubjectification. In A. H. Jucker, & I. Taavitsainen, (Eds.), *Historical pragmatics* (pp. 127–163). de Gruyter Mouton.

Los, B., & Kemenade, A. van. (2012). Information structure and syntax in the history of English. In A. Bergs, & L. J. Brinton (Eds.), *English historical linguistics: An international handbook*, Vol. 2, 1475–1490. De Gruyter Mouton.

Lutzky, U. (2012). *Discourse Markers in Early Modern English*. John Benjamins. https://doi.org/10.1075/pbns.227

Lyngfelt, B. (2018). Introduction: Constructicons and constructicography. In B. Lyngfelt, L. Borin, K. Ohara, & T. T. Torrent (Eds.), *Constructicography. Constructicon development across languages* (pp. 1–18). John Benjamins. https://doi.org/10.1075/cal.22.01lyn

Lyons, J. (1982). Deixis and subjectivity: Loquor, ergo sum? In R. J. Jarvella & W. Klein (Eds.), *Speech, place, and action: Studies in deixis and related topics* (pp. 101–124). Wiley.

MacWhinney, B., & O'Grady, W. (Eds.). (2015). *The handbook of language emergence*. Wiley Blackwell. https://doi.org/10.1002/9781118346136

Mair, C. (1997). Parallel corpora: A real-time approach to the study of language change in progress. In M. Ljung (Ed.), *Corpus-based studies in English* (pp. 95–209). Rodopi.

Mauri, C., & Auwera, J. van der. (2012). Connectives. In K. Allan, & K. M. Jaszczolt (Eds.), *The Cambridge handbook of pragmatics* (pp. 377–401). Cambridge University Press. https://doi.org/10.1017/CBO9781139022453.021

Meillet, A. (1958[1912]). L'évolution des formes grammaticales [The evolution of grammatical forms]. In A. Meillet, *Linguistique historique et linguistique générale* [Historical linguistics and general linguistics] (pp. 130–148). Champion.

Michaelis, L. A. (2013). Sign-Based construction grammar. In T. Hoffmann, & G. Trousdale (Eds.), *The Oxford handbook of construction grammar* (pp. 133–152). Oxford University Press.

Milroy J., & Milroy, L. (1985). Linguistic change, social network and speaker innovation. *Journal of Linguistics* 21(2), 339–383. https://doi.org/10.1017/S0022226700010306

Miskoviċ-Lukoviċ, M., & Dediaċ, M. N. (2012). The discourse marker *odnoso* at the ICTY: A case of disputed translation in war crimes trials. *Journal of Pragmatics* 44, 1355–1377. https://doi.org/10.1016/j.pragma.2012.06.013

Mittwoch, A., Huddleston, R., & Collins, P. (2002). The clause adjuncts. In R. Huddleston, & G. K. Pullum, *The Cambridge grammar of the English language*, Chapter 8. Cambridge University Press. https://doi.org/10.1017/9781316423530.009

Mulder, J., & Thompson, S. A. (2008). The grammaticalization of *but* as a final particle in conversation. In R. Laury (Ed.), *Crosslinguistic studies of clause combining: The multifunctionality of conjunctions* (pp. 179–204). John Benjamins. https://doi.org/10.1075/tsl.80.09mul

Mulder, J., Thompson, S. A., & Williams, C. P. (2009). Final *but* in Australian English conversation. In P. Peters, P. Collins, & A. Smith (Eds.), *Comparative studies in Australian and New Zealand English: Grammar and beyond* (pp. 337–358). John Benjamins. https://doi.org/10.1075/veaw.g39.19mul

Murray, D. (1979). 'Well'. *Linguistic Inquiry* 10(4), 727–734.

Narrog, H. (2014[2012]). Beyond intersubjectification: Textual uses of modality and mood in subordinate clauses as part of *speech-act orientation*. In L. Brems, L. Ghesquière, & F. Van de Velde (Eds.), *Intersubjectivity and intersubjectification in grammar and discourse* (pp. 29–52). John Benjamins. https://doi.org/10.1075/bct.65.03nar

Nesselhauf, N. (2010). The development of future time expressions in Late Modern English: Redistribution of forms or change in discourse? *English Language and Linguistics* 14(2), 163–186. https://doi.org/10.1017/S1360674310000043

Nevalainen, T. (1990). Modeling functional differentiation and function loss: The case of "but". In S. Adamson, V. Law, N. Vincent, & S. Wright (Eds.), *Papers from the 5th International Conference on English Historical Linguistics* (pp. 337–355). John Benjamins. https://doi.org/10.1075/cilt.65.20nev

Nevalainen, T. (1991). *BUT, ONLY, JUST: Focusing on adverbial change in Modern English* 1500–1900. Société Néophilologique.

Nicolle, S. (2011). Pragmatic aspects of grammaticalization. In H. Narrog, & B. Heine (Eds.), *The Oxford handbook of grammaticalization* (pp. 401–412). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199586783.013.0032

Noël, D. (2007). Diachronic construction grammar and grammaticalization theory. *Functions of Language* 14, 177–202. https://doi.org/10.1075/fol.14.2.04noe

Noël, D. (2017). The development of non-deontic BE BOUND TO in a radically usage-based diachronic construction grammar perspective. *Lingua* 199, 72–93. https://doi.org/10.1016/j.lingua.2017.07.012.

Norde, M. (2009). *Degrammaticalization*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199207923.001.0001

Oates, S. L. (2000). Multiple discourse marker occurrence: Creating hierarchies for natural language generation. *Proceedings of ANLP-NAACL* (pp. 41–45).

Onodera, N. O. (2004). *Japanese Discourse Markers: Synchronic and diachronic discourse analysis*. John Benjamins. https://doi.org/10.1075/pbns.132

Östman, J.-O. (1981). *You know: A Discourse-Functional Approach*. John Benjamins. https://doi.org/10.1075/pb.ii.7

Parkes, M. B. (1992). *Pause and effect: An introduction to the history of punctuation in the West*. University of California Press.

Percillier, M. (2020). Allostructions, homostructions or a constructional family? Changes in the network of secondary predicate constructions in Middle English. In L. Sommerer, & E. Smirnova (Eds.), *Nodes and networks in Diachronic Construction Grammar* (pp. 213–242). John Benjamins. https://doi.org/10.1075/cal.27.06per

Perek, F. (2012). Alternation-based generalizations are stored in the mental grammar: Evidence from a sorting task experiment. *Cognitive Linguistics* 23(3), 601–635. https://doi.org/10.1515/cog-2012-0018

Perek, F. (2015). *Argument structure in usage-based construction grammar*. John Benjamins.

Perek, F. (2020). Productivity and schematicity in constructional change. In L. Sommerer, & E. Smirnova, (Eds.), *Nodes and networks in Diachronic Construction Grammar* (pp. 141–166). John Benjamins. https://doi.org/10.1075/cal.27.04per

Perek, F., & Patten, A. L. (2019). Towards an English constructicon using patterns and frames. In S. Hunston & F. Perek (Eds.), *Constructions in Applied Linguistics*. Special issue, *International Journal of Corpus Linguistics*. 24(3), 354-384.

Petré, P. (2016). Grammaticalization by changing co-text frequencies, or why [BE Ving] became the 'progressive'. *English Language and Linguistics* 20(1), 31–54. https://doi.org/10.1017/S1360674315000210

Petré, P. (2019). How constructions are born. The role of patterns in the constructionalization of *be going to INF*. In B. Busse, & R. Möhlig-Falke (Eds.), *Patterns in language and linguistics: New perspectives on a ubiquitous concept* (pp. 157–192). De Gruyter Mouton. https://doi.org/10.1515/9783110596656-007

Petré, P., & Anthonissen, L. (2020). Individuality in complex systems: A constructionalist approach. In P. Petré, & L. Anthonissen (Eds.), *Constructionalist approaches to individuality in language*. Special issue, *Cognitive Linguistics* 31(2), 185–212. https://doi.org/10.1515/cog-2019-0033

Petré, P., & Van de Velde, F. (2018). The real-time dynamics of the individual and the community in grammaticalization. *Language* 94(4), 867–901. https://doi.org/10.1353/lan.2018.0056

Petruck, M. R. (2011). Advances in frame semantics. *Constructions and Frames* 3(1), 1–8. https://doi.org/10.1075/cf.3.1.00pet

Pons Bordería, S. (2008). Do discourse markers exist? On the treatment of discourse markers in Relevance Theory. *Journal of Pragmatics* 40, 1411–1434. https://doi.org/10.1016/j.pragma.2008.03.013

Pons Bordería, S., & Estellés Arguedas, M. (2009). Expressing digression linguistically: Do digressive markers exist? *Journal of Pragmatics* 41, 921–936. https://doi.org/10.1016/j.pragma.2008.08.011

Pratt, L., & Denison, D. (2000). The language of the Southey-Coleridge circle. *Language Sciences* 22, 401–422. https://doi.org/10.1016/S0388-0001(00)00013-9

Prévost, S. (2011). *A propos* from verbal complement to discourse marker: A case of grammaticalization? *Linguistics* 49(2), 391–413. https://doi.org/10.1515/ling.2011.012

Pullum, G., & Huddleston, R. (2002). Adjectives and adverbs. In R. Huddleston, & G. Pullum (Eds.), *The Cambridge grammar of the English language* (pp. 526–595). Cambridge University Press. https://doi.org/10.1017/9781316423530.007

Pulvermüller, F. (2002). *The neuroscience of language: On brain circuits of words and serial order*. Cambridge University Press.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. Longman.

Ramat, P., & Ricca, D. (1994). Prototypical adverbs: On the scalarity/radiality of the notion ADVERB. *Rivista di Linguistica* 6, 289–326.

Recanati, F. (2010). *Truth-conditional pragmatics*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199226993.001.0001

Rissanen, M. (2004). Grammaticalisation from side to side: On the development of *beside(s)*. In H. Lindquist, & C. Mair (Eds.), *Corpus approaches to grammaticalization in English* (pp. 151–170). John Benjamins. https://doi.org/10.1075/scl.13.08ris

Rosén, H. (2009). Coherence, sentence modification, and sentence-part modification – the contribution of particles. In P. Baldi, & P. Cuzzolin (Eds.), *New perspectives on historical Latin syntax: Vol. 1. Syntax of the sentence* (pp. 317–441). Mouton de Gruyter.

Rostila, J. (2004). Lexicalization as a way to grammaticalization. In F. Karlsson (Ed.), *Proceedings of the 20th Scandinavian Conference of Linguistics*. http://www.ling.helsinki.fi/kielitiede/20scl/Rostila.pdf.

Rouchota, Villy. (1998). Procedural meaning and parenthetical Discourse Markers. In A. H. Jucker, & Y. Ziv (Eds.), *Discourse Markers: Descriptions and theory* (pp. 97–126). John Benjamins. https://doi.org/10.1075/pbns.57.07rou

Rumelhart, D. E., & McClelland, J. L. (Eds.) (1986). *Parallel distributed processing: Explorations in the microstructures of cognition*, 2 Vols. MIT Press. https://doi.org/10.7551/mitpress/5236.001.0001

Sankoff, G. (2019). Language change across the life-span: Three trajectory types. *Language* 95(2), 197–229. https://doi.org/10.1353/lan.2019.0029

Sarda, L., Carter-Thomas, S., Fagard, B., & Charolles, M. (Eds.) (2014). *Adverbials in use: From predicative to discourse functions*. Presses Universitaires de Louvain.

Saussure, F. de. (1983[1916]). *Course in general linguistics*. Trans. by R. Harris. Open Court.

Schegloff, E. A., & Sacks, H. (1973). Opening up closings. *Semiotica* 4, 289–327. https://doi.org/10.1515/semi.1973.8.4.289

Schiffrin, D. (1987). *Discourse Markers*. Cambridge University Press. https://doi.org/10.1017/CBO9780511611841

Schmid, H.-J. (Ed.) (2012). *Cognitive pragmatics*. De Gruyter. https://doi.org/10.1515/9783110214215

Schmid, H.-J. (2016). Why Cognitive Linguistics must embrace the social and pragmatic dimensions of language and how it could do so more seriously. *Cognitive Linguistics* 27(4), 543–557. https://doi.org/10.1515/cog-2016-0048

Schmid, H.-J. (2017). A framework for understanding linguistic entrenchment and its psychological foundations. In H.-J. Schmid (Ed.), *Entrenchment and the psychology of language learning: How we reorganize and adapt linguistic knowledge* (pp. 9–35). De Gruyter Mouton and American Psychological Association. https://doi.org/10.1037/15969-002

Schmid, H.-J. (2020). *The dynamics of the linguistic system: Usage, conventionalization, and entrenchment*. Oxford University Press. https://doi.org/10.1093/oso/9780198814771.001.0001

Schmid, H.-J., & Mantlik, A. (2015). Entrenchment in historical corpora? Reconstructing dead authors' minds from their usage profile. *Anglia* 133(4), 583–623. https://doi.org/10.1515/ang-2015-0056

Schönefeld, D. (2011). *Converging evidence: Methodological and theoretical issues for linguistic research*. John Benjamins. https://doi.org/10.1075/hcp.33

Schourup, L. (2016[1985]). *Common discourse particles in English conversation*. Routledge.

Schwenter, S. A., & Traugott, E. C. (1995). The semantic and pragmatic development of substitutive complex prepositions in English. In A. H. Jucker (Ed.), *Historical pragmatics* (pp. 243–273). John Benjamins. https://doi.org/10.1075/pbns.35.16sch

Schwenter, S. A., & Waltereit, R. (2010). Presupposition accommodation and language change: From additivity to speech-act marking. In K. Davidse, L. Vandelanotte, & H. Cuyckens (Eds.), *Subjectification, intersubjectification and grammaticalization* (pp. 75–102). De Gruyter Mouton. https://doi.org/10.1515/9783110226102.2.75

Shinzato, R. (2014). Subjectivity, intersubjectivity and Japanese grammar. In K. Kabata, & T. Ono (Eds.), *Usage-based approaches to Japanese grammar: Towards the understanding of human language* (pp. 85–108). John Benjamins. https://doi.org/10.1075/slcs.156.08shi

Simon-Vandenbergen, A.-M., & Willems, D. (2011). Cross-linguistic data as evidence in the grammaticalization debate: The case of discourse markers. In L. Degand, & A.-M. Simon-Vandenbergen (Eds.), *Grammaticalization, pragmaticalization and/or (inter)subjectification: methodological issues for the study of discourse markers*. Special issue, *Linguistics* 49(2), 333–364.

Smirnova, E. (2015). Constructionalization and constructional change: The role of context in the development of constructions. In J. Barðdal, E. Smirnova, L. Sommerer, & S. Gildea (Eds.), *Diachronic Construction Grammar* (pp. 81–106). John Benjamins. https://doi.org/10.1075/cal.18.03smi

Smirnova, E., & Sommerer, L. (2020). The nature of the node and the network – Open questions in Diachronic Construction Grammar. In L. Sommerer, & E. Smirnova (Eds.), *Nodes and networks in Diachronic Construction Grammar* (pp. 1–42). John Benjamins. https://doi.org/10.1075/cal.27.int

Sommerer, L. (2018). *Article emergence in Old English: A constructionalist perspective*. De Gruyter Mouton. https://doi.org/10.1515/9783110541052

Sommerer, L., & Smirnova, E. (Eds.) (2020). *Nodes and networks in Diachronic Construction Grammar*. John Benjamins. https://doi.org/10.1075/cal.27

Sorva, E. (2007). Grammaticalization and syntactic polyfunctionality: The case of *albeit*. In U. Lenker, & A. Meurman-Solin (Eds.), *Connectives in the history of English* (pp. 115–143). John Benjamins. https://doi.org/10.1075/cilt.283.08sor

Sperber, D., & Wilson, D. (1995[1986]). *Relevance: Communication and cognition*. Blackwell, 2nd, rev. edn.

Swan, T. (1994). A note on Old English and Old Norse initial adverbials and word order with special reference to sentence adverbials. In T. Swan, E. Mørck, & O. J. Westvik (Eds.), *Language change and language structure* (pp. 233–270). Mouton de Gruyter. https://doi.org/10.1515/9783110886573.233

Sweetser, E. E. (1988). Grammaticalization and semantic bleaching. In S. Axmaker, A. Jaisser, & H. Singmaster (Eds.), *Berkeley Linguistics society 14: General session and parasession on grammaticalization* (pp. 389–405). Berkeley Linguistics Society. https://doi.org/10.3765/bls.v14i0.1774

Sweetser, E. E. (1990). *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure*. Cambridge University Press. https://doi.org/10.1017/CBO9780511620904

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and syntactic description, Vol. III: Grammatical categories and the lexicon* (pp. 57–149). Cambridge University Press, 2nd edn.

Talmy, L. (2000). *Toward a cognitive linguistics, Vol. I. Concept structuring systems*. The MIT Press. https://doi.org/10.7551/mitpress/6847.001.0001

Terkourafi, M. (2009). On de-limiting context. In A. Bergs, & G. Diewald (Eds.), *Contexts and constructions* (pp. 17–42). John Benjamins. https://doi.org/10.1075/cal.9.02ter

Thompson, S. A., & Mulac, A. J. (1991). A quantitative perspective on the grammaticalization of epistemic parentheticals in English. In E. C. Traugott, & B. Heine (Eds.), *Approaches to grammaticalization*, Vol. 2, 213–329. John Benjamins. https://doi.org/10.1075/tsl.19.2.16tho.

Timberlake, A. (1977). Reanalysis and actualization in syntactic change. In C. N. Li (Ed.), *Mechanisms of syntactic change* (pp. 141–177). University of Texas Press. https://doi.org/10.7560/750357-006

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Harvard University Press.

Torrent, T. T. (2015). On the relation between inheritance and change: The Constructional Convergence and Construction Network Reconfiguration hypotheses. In J. Barðdal, E. Smirnova, L. Sommerer, & S. Gildea (Eds.), *Diachronic Construction Grammar* (pp. 173–211). John Benjamins. https://doi.org/10.1075/cal.18.06tor

Tottie, G. (2014). On the use of *uh* and *um* in American English. *Functions of Language* 21(1), 6–29. https://doi.org/10.1075/fol.21.1.02tot

Traugott, E. C. (1982). From propositional to textual and expressive meanings: Some semantic-pragmatic aspects of grammaticalization. In W. P. Lehmann, & Y. Malkiel (Eds.), *Perspectives on historical linguistics* (pp. 245–271). John Benjamins. https://doi.org/10.1075/cilt.24.09clo

Traugott, E. C. (1997[1995]). The role of the development of discourse markers in a theory of grammaticalization. Paper presented at ICHL XII, Manchester, 1995. http://www.stanford.edu/~traugott/papers/discourse.pdf

Traugott, E. C. (1999). The role of pragmatics in a theory of semantic change. In J. Verschueren (Ed.), *Pragmatics in 1998: Selected papers from the 6th International Pragmatics Conference*, Vol. 2, 93–102. International Pragmatics Association.

Traugott, E. C. (2003). From subjectification to intersubjectification. In R. Hickey (Ed.), *Motives for language change* (pp. 124–139). Cambridge University Press. https://doi.org/10.1017/CBO9780511486937.009

Traugott, E. C. (2004). Historical pragmatics. In L. R. Horn, & G. Ward (Eds.), *Handbook of pragmatics* (pp. 538–561). Blackwell.

Traugott, E. C. (2010). (Inter)subjectivity and (inter)subjectification: A reassessment. In K. Davidse, L. Vandelanotte, & H. Cuyckens (Eds.), *Subjectification, intersubjectification and grammaticalization* (pp. 29–71). de Gruyter Mouton. https://doi.org/10.1515/9783110226102.1.29

Traugott, E. C. (2014). On the function of the epistemic adverbs *surely* and *no doubt* at the left and right peripheries of the clause. In K. Beeching, & U. Detges (Eds.), *Discourse functions at the left and right periphery: Crosslinguistic investigations of language use and language change* (pp. 72–91). Brill.

Traugott, E. C. (2018a). Modeling language change with constructional networks. In S. Pons Bordería, & Ó. Loureda Lemos (Eds.), *Beyond grammaticalization and Discourse Markers: New issues in the study of language change* (pp. 17–50). Brill.

Traugott, E. C. (2018b). Rethinking the role of invited inferencing in change from the perspective of interactional texts. In O. Ehmer, & M. Rosemeyer (Eds.), Inferences in interaction and language change. Special issue, *Open Linguistics* 4(1), 19–34. https://www.degruyter.com/view/j/opli.2018.4.issue-1/issue-files/opli.2018.4.issue-1.xml. https://doi.org/10.1515/opli-2018-0002

Traugott, E. C. (2020a). The development of "digressive" discourse–topic shift markers in English. In B. Fagard, & M. Charolles (Eds.), *Topic shifters in contrastive perspective*. Special issue, *Journal of Pragmatics* 156, 121–135. https://doi.org/10.1016/j.pragma.2019.02.002

Traugott, E. C. (2020b). Expressions of stance-to-text: Discourse management markers as stance markers. In G. Kaltenböck, M. J. López-Couso, & B. Méndez-Naya (Eds.), *Investigating stance in English: Synchrony and diachrony*. Special issue, *Language Sciences* 82. Online at https://doi.org/10.1016/j.langsci.2020.101329.

Traugott, E. C. (2020c). Is *back to my point* a pragmatic marker? An inquiry into the historical development of some metatextual discourse management markers. In J. Martines, S. Rodríguez, & J. Antolí (Eds.), *Context and linguistic change*. Special issue, *Catalan Journal of Linguistics*, 13–29. https://doi.org/10.5565/rev/catjl.307

Traugott, E. C. In press. *Ten lectures on a diachronic constructionalist approach to Discourse Structuring Markers*. Brill. Special issue, https://doi.org/10.1163/9789004507050

Traugott, E. C., & Dasher, R. B. (2002). *Regularity in semantic change*. Cambridge University Press.

Traugott, E. C., & Trousdale, G. (2013). *Constructionalization and constructional changes*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199679898.001.0001

Traugott, E.C., & Heine, B. (Eds). 1991. *Approaches to Grammaticalization*, 2 Vols. John Benjamins.

Trousdale, G., & Traugott, E. C. (2021). Rethinking constructionalization: The history of *by the way*. Paper presented at ISLE 6, Joensuu, June.

van Bogaert, J. (2011). *I think* and other complement-taking mental predicates: A case of and for constructional grammaticalization. *Linguistics* 49(2), 295–332. https://doi.org/10.1515/ling.2011.009

Vandelanotte, L. (2009). *Speech and thought representation in English: A cognitive-functional approach*. Mouton de Gruyter. https://doi.org/10.1515/9783110215373

van der Auwera, J. (2009). The Jespersen cycles. In E. van Gelderen (Ed.), *Cyclical Change* (pp. 35–71). John Benjamins. https://doi.org/10.1075/la.146.05auw

Van de Velde, F. (2014). Degeneracy: The maintenance of constructional networks. In R. Boogaart, T. Colleman, & G. Rutten (Eds.), *Extending the scope of construction grammar*, 141–180. De Gruyter. https://doi.org/10.1515/9783110366273.141

Verhagen, A. (1995). Subjectification, syntax, and communication. In D. Stein, & S. Wright, (Eds.), *Subjectivity and subjectivisation in language* (pp. 103–28). Cambridge University Press. https://doi.org/10.1017/CBO9780511554469.006

Verhagen, A. (2005). *Constructions of intersubjectivity. Discourse, syntax, and cognition*. Oxford University Press.

Verhagen, A. (2007). Construal and perspectivization. In D. Geeraerts, & H. Cuyckens (Eds.), *The oxford handbook of cognitive linguistics* (pp. 48–81). Oxford University Press.

Vincent, N. (2015). Compositionality and change. In C. Bowern, & B. Evans (Eds.), *The Routledge handbook of historical linguistics*. Routledge.

Waltereit, R. (2006). The rise of discourse markers in Italian: A specific type of language change. In K. Fischer (Ed.), *Approaches to discourse particles* (pp. 61–76). Elsevier.

Waltereit, R. (2012). On the origins of grammaticalization and other types of language change in discourse strategies. In K. Davidse, T. Breban, L. Brems, & T. Mortelmans (Eds.), *Grammaticalization and language change: New reflections* (pp. 51–72). John Benjamins. https://doi.org/10.1075/slcs.130.03wal

Weinreich, U., Labov, W., & Herzog, M. [2017(1968)]. Empirical foundations for a theory of language change. In W. P. Lehmann, & Y. Malkiel (Eds.), *Directions for historical linguistics* (pp. 95–189). University of Texas Press, repr.

Wierzbicka, A. (2006). *English: Meaning and culture*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195174748.001.0001

Wilcox, S., & Occhino, C. (2016). Historical change in signed languages. *Oxford Handbooks Online*. https://doi.org/10.1093/oxfordhb/9780199935345.013.24

Wilcox, S., & Martinez, R. (2020). The conceptualization of space: *Places* in signed discourse. *Frontiers in Psychology* 11: 1406. https://doi.org/10.3389/fpsyg.2020.01406

Winograd, T. (1976). Towards a procedural understanding of semantics. *Revue Internationale de Philosophie* 30, 260–303.

Wolk, C., Bresnan, J., Rosenbach, A., & Szmrecsanyi, B. (2013). Dative and genitive variability in Late Modern English: Exploring cross-constructional variation and change. *Diachronica* 30, 382–419. https://doi.org/10.1075/dia.30.3.04wol.

Zehentner, E. (2019). *Competition and language change: The rise of the English dative alternation*. De Gruyter Mouton. https://doi.org/10.1515/9783110633856

Zehentner, E., & Traugott, E. C. (2020). Constructional networks and the development of benefactive ditransitives in English. In L. Sommerer, & E. Smirnova (Eds.), *Nodes and networks in Diachronic Construction Grammar* (pp. 168–211). John Benjamins. https://doi.org/10.1075/cal.27.05zeh

Names index

Subject index

This book is a contribution to the growing field of diachronic construction grammar. Focus is on corpus evidence for the importance of including conventionalized pragmatics within construction grammar and suggestions for how to do so. The empirical domain is the development of Discourse Structuring Markers in English such as *after all, also, all the same, by the way, further* and *moreover* (also known as Discourse Markers). The term Discourse Structuring Markers highlights their use not only to connect discourse segments but also to shape discourse coherence and understanding. Monofunctional Discourse Structuring Markers like *further, instead, moreover* are distinguished from multifunctional ones like *after all* and *by the way*. Drawing on usage-based work on constructionalization and constructional changes, the book is in three parts: foundational concepts, case studies, and currently open issues in diachronic construction grammar. These open issues are how to incorporate the concepts subjectification and intersubjectification into a constructional account of change, whether position in a clause is a construction, and the nature of constructional networks and how they change.

ISBN 978 90 272 1091 3

John Benjamins Publishing Company