Grammaticalizing connectives in English and discourse information structure

Diana M. Lewis, Aix Marseille University

Abstract

The development of lexical expressions such as VP adverbs, matrix clauses and prepositional phrases into discourse markers and connectives expressing coherence relations has been well explored in the grammaticalization literature, under a broad view of grammaticalization, but there has been less emphasis on how the discourse information structuring functions of markers evolve during these developments. This paper investigates the relationship between discourse coherence marking and information structure by examining two developing discourse connectives: it suggests that the grammaticalization of a lexeme in its construction into a discourse marker may involve acquiring or strengthening discourse-level information structuring functions - indicating relative informational salience - and that (further) grammaticalization of markers may be at least partly directed by information structure.

Keywords

coherence relations; discourse connectives; English; grammaticalization

1. Introduction

The development of lexical expressions such as VP adverbs, matrix clauses and prepositional phrases into discourse markers expressing coherence relations ('discourse connectives' or 'relational discourse particles') has been well explored in the grammaticalization literature, under a broad view of grammaticalization. This work has intersected with recent renewal of interest in parenthetical and peripheral constructions, discourse markers being a major category of 'extra-clausal constituent'. In particular, attention has focused on the extent to which differing positions of markers correlate with different functions. But there has been less emphasis on the discourse information-structuring functions of markers and how these evolve. This chapter investigates aspects of the relationship between English discourse connectives and discourse information structure: it suggests that the grammaticalization of a lexeme-in-its-construction into a discourse connective may involve acquiring or strengthening information- structuring functions -- foregrounding, backgrounding and structuring units of discourse -- and that (further) grammaticalization may be at least partly motivated by information structure (cf. Lehmann 2008). A case study is presented of two English discourse connectives that can occur in different positions relative to the host unit. In each case, the meaning of the discourse connective, in the sense of the coherence relation that it expresses, is similar across initial and final positions but, it is argued, the
information structure of the overall 'discourse construction' results at least in part from the discourse connective and its position. If so, there may be implications for how models of discourse structure relate coherence to discourse information structure.

The rest of the chapter is organized into the following sections. The next section sets the study in the context of recent work on the grammaticalization of discourse markers and in particular on the significance of position in the utterance. It also outlines the relationship between coherence relations and information structure. Section 3 presents a case study of the recent evolution of English discourse connectives after all and in fact in initial and final positions. Section 4 contains a discussion of findings and Section 5 concludes the chapter.

2. Discourse connectives

2.1 Grammaticalization and discourse connectives

The terminological confusion surrounding discourse connectives (or discourse markers, pragmatic markers, linking adverbials, adverbial connectors and so on) reflects uncertainty over their categorial status as well as differing theoretical frameworks. For practical purposes, a discourse connective will be taken as an adverb or adverbial expression characterized by discourse-semantic, structural and lexicalization parameters. (‘Lexicalization’ is used here in the sense of loss of compositionality: that is, semantic fusion (Bybee 1985:37) or renunciation of internal analysis (Lehmann 2002: 15).) A discourse connective expresses a coherence relation, is structurally dependent on a host and constitutes a single- or multi-word lexeme (cf. Lewis 2011). This characterization makes no claims about discourse connectives or markers constituting a distinct word class. Insofar as discourse connectives refer back anaphorically to a previous idea, they cannot occur discourse-initially (i.e. at the start of a conversation or text, where no ideas have yet been expressed), except, of course, where a non-verbal previous idea is currently activated in the discourse context (cf. Blakemore 1996: 337-8).

The argument that the development of discourse markers, including discourse connectives, is best analysed as a case of grammaticalization dates back at least to Traugott (1995). Since then, grammaticalization scholars have been (increasingly, perhaps) split on whether such developments can be classed as grammaticalization. Roughly speaking, the arguments for viewing discourse marker development as grammaticalization centre around the semantic changes the expressions undergo (desemanticization/bleaching) and the categorial changes from a main constituent such as prepositional phrase (PP) or predicate to some sort of sentence adverb (Brinton and Traugott 2005, Hopper and Traugott 2003). The arguments against tend to focus on the syntactic changes, in particular the fact that emergent discourse markers show increased scope and more flexible positioning, whereas canonical grammaticalization developments show scope decrease, typically involving items becoming affixes. Those who adopt a narrow view of grammaticalization have proposed alternative labels for the type of change that subsumes discourse marking, including ‘pragmaticization’ (Erman and Kotsinas 1993, Aijmer 1997) on the grounds that the resultant expressions are ‘pragmatic’, or ‘co-optation’ (Heine et al 2015) on the grounds of sudden scope shift. In this chapter, a broad or extended view of grammaticalization is taken, encompassing changes where items or constructions become more grammatical, whether or not they become more syntactically fixed and more obligatory.
2.2 Positions of discourse connectives

The position of discourse connectives has attracted much recent attention. Questions that have arisen regarding the position of connectives include the frequency and origins of different positions, their motivation and correlations with function. The hypothesis will be put forward in sections 3 and 4 that the recent increase in English of final position connectives is at least partly motivated by the management of informational salience.

Connectives combine with their host unit and the previous unit (to which they refer back) to form a discourse-level structure of the type \([p] \text{ [connective]} [q]\) where the connective has scope over the host \(q\). English discourse connectives occur in the three sentence-adverb positions: initial, medial (that is, post first auxiliary or pre-verb if no auxiliary) and final. We will refer to these positions as left periphery (LP), medial and right periphery (RP). The term 'periphery' has been variously used in the functional literature to refer to a structural slot in relation to the clause, to the utterance or to the turn (Beeching and Detges 2014). In the context of discourse connectives, we use the term here in relation to 'host', to refer to a position preceding or following a discourse unit, of any syntactic status, that encodes an idea and acts as the host of a connective.

English discourse connectives are often assumed to occur predominantly in sentence-initial position (e.g. Huddleston and Pullum 2002:578), but many can occur at both LP and RP. Examples are actually, after all, anyway, even so, however, in fact, of course, otherwise, rather, really, surely, though, then. They thus align with 'comment clauses' and deverbal discourse markers such as I see, I mean, you know, I think, look, mind you, as far as I know, which have similar positional distribution, as well as with many evaluative sentence adverbs such as sadly, luckily, curiously and so on. It has been suggested that only English, among IndoEuropean languages, has a regular construction with adverbial connectives in final position (Lenker 2010: 198) and that this is "a comparatively recent syntactic change in English, i.e. a new position for adverbial connectors" (2010: 202). And final position (RP) for connectives appears to be increasing. Biber et al have found that for what they term 'linking adverbials', in the Conversation register initial position is the commonest, medial position is rare, and about 40% of tokens are in final position (1999: 891). Given also the findings that stance adverbials overall occur more frequently in final position in Conversation than in written registers (Biber et al 1999), it seems safe to infer that RP position is the more recent1.

Recent interest in utterance-final particles in general (e.g. Barth-Weingarten and Couper-Kuhlen 2002, Hancil 2014, Hancil, Haselow and Post 2015, Haselow 2012, 2013, 2015, Kim and Jahne 2011, Mulder and Thompson 2008, Traugott 2012, 2013) has led to hypotheses about the origin and the function of the 'utterance-final slot', and in particular the functional differences that may correlate with position. Final particles have, according to Haselow, "essentially the same function as conjunctions, establishing a two-place relationship between two structural units, conjunctions operating on the sentence-internal level, final particles at utterance-level and thus across clause-boundaries." (2015b: 210). Aijmer (2002) discusses RP actually as differing from LP actually in having interpersonal and positive politeness functions. Likewise, in her analysis of anyway in varieties of spoken English, Aijmer finds LP anyway and RP anyway to be distinct constructions with distinct meaning potentials (respectively a resumptive discourse marker or topic closer and modal/intersubjective) (2016:48-54). In similar vein, RP though, is analysed by Lenker as a concessive connector that has been reanalysed as marker of contrast in final position (2010:186,196). Others have argued that the left periphery favours subjective meanings, and the
right periphery intersubjective meanings (Izutsu and Izutsu (2013)\textsuperscript{2}; Beeching and Detges (2014)). However, counter-evidence has been found for certain English expressions in the peripheries (Traugott 2014, 2016), so that such a functional distribution may at best be a tendency. Not all scholars, however, have found functional differences according to position. In her analysis of Irish English \textit{like}, Schweinberger (2015:127) finds that the clause-final and clause-medial ‘like’ share functional properties, despite differing in direction of scope (backwards vs. forwards) and suggests that one is chosen over the other depending on the phrasal constituents that the marker modifies.

Given that a number of connectives can occur in either initial (LP) or final (RP) position in the same context, it is by no means immediately obvious that there are systematic functional differences between them.

2.3 Coherence relations and information structure

Grammaticalization and information structure are linked in at least two ways, according to Lehmann: "[o]n the one hand, information structure is present in the source constructions that undergo grammaticalization and may direct their course. On the other hand, information structure is itself susceptible of grammaticalization.” (2008:207). In the case of discourse connectives, their grammaticalization results in expressions that serve to background (or ‘depropositionalize’) a discourse relation. Connectives also participate in the information structure\textsuperscript{3} of the discourse (the sequence of related ideas) they are a part of, to signal the relative salience or informational prominence of adjacent idea units.

Several studies of RP English discourse markers have considered the implications of information structure. In her discussion of how English however and though have developed from subordinators into adverbial connectors, for instance, Lenker claims that "these changes are [...] induced by factors of information structure and may [...] also lead to changes in information structure. In this view, adverbial connectors are -- in contrast to subordinators -- very strong indicators of a great illocutionary weight of the second connect.” (2010:32). She suggests that information structure accounts for the development of though from subordinator (in a hypotactic construction) to sentence-final adverbial connector (in a paratactic construction) thereby changing from a marker of concession to a marker of contrast (2010: 201). The shift of the marker to final position, "a position which clearly differentiates hypotaxis from parataxis", is "triggered by the lack of other distinctive means” (2010: 213). The result is a new slot for connectives, and so a new connective construction.

In his studies of RP then and anyway, Haselow addresses the relative structural status of the conjuncts, the final position offering "a structural alternative to hypotaxis" (2012:154), and emphasizing "the paratactic nature of interactive speech” (2015:227) that gave rise to the construction. Haselow refers to RP then as being "based on an implied paratactic conditional construction" (2012:154) but does not discuss the relative informational salience of the conjuncts.

The next section gives an overview of the development of particular connective functions in \textit{in fact} and \textit{after all}. It will be suggested that, for each adverbial, while the rhetorical relation expressed through occurrences at LP and RP positions does not alter, there is an information structural difference. Initial connectives tend to act as presentatives, foregrounding the host idea that follows them, so that (leaving prosody aside for the moment) it functions as a further utterance carrying as much weight as the previous one. By contrast, final connectives bind their host to the previous idea in a discourse construction where the second conjunct has less informational salience than the first.
3. Grammaticalizing discourse connectives and position

3.1 Two recent English connectives

*In fact* and *after all* have followed similar trajectories to become discourse connectives, and have been studied as examples of grammaticalization at discourse level (e.g. Schwenker and Traugott 2002 on *in fact*; Lewis 2007 on *after all*). Both have gone through successive splits: first the prepositional phrase coalesced in some contexts into a complex adverb and later came to express a newer, connective use alongside the older VP-adverbial use. And in each case a further connective use developed, arguably at a more abstract or more subjective level. These developments are typical of grammaticalization and have resulted in polysemy as older, more lexical uses coexist with the newer, more grammatical ones. Both *in fact* and *after all*, as connectives, refer back anaphorically to a previous discourse unit or act as presupposition triggers to evoke an idea that is active in the discourse context. However, these expressions continue to evolve, and may currently be developing further uses.

Discourse connectives are typically multifunctional, and different senses may be differently distributed across syntactic positions. *After all* and *in fact* are interesting in that both initial and final positions, as well as medial position (post-auxiliary) are found for what seems to be one and the same function, although the positional split may also reflect an incipient functional split involving some further grammaticalization.

Section 3.2 looks at the development and current usage of *in fact*; section 3.3 focuses on *after all*. Section 3.4 summarizes the findings for these two adverbial connectives. Historical data on *in fact* and *after all* are taken from the period 1680-1920s and present-day data from the 1960s and 1980s. The data sources are listed in the appendix. The historical data were chosen to be representative insofar as possible of everyday English in the British Isles and to be as balanced as possible across periods. They consist primarily of personal letters, drama, diaries, and journals. The result is a corpus that provides small numbers of occurrences, and suffers from some 'burstiness' in the data, the stylistic quirks of authors being particularly apparent in usage of idiomatic expressions such as these two. For present-day English examples taken from the British National Corpus the text code is given along with the genre: BNC-CG stands for the contextually-governed part of the spoken section of the corpus, and BNC-DS stands for the demographically-sampled part, i.e. conversation.

3.2 In fact

3.2.1 Evolution of 'in fact'

*In fact* developed into a discourse connective from the prepositional phrase (PP). As a PP used as a VP-adverbial in Early Modern English, *in fact* gradually coalesced into an epistemic adverb emphasizing the veracity of its host. It is found in the eighteenth century in initial, medial and final positions. It often occurred in contexts which contrasted what was really the case 'in fact' with what was thought or said to be the case; its function was thus often epistemic (1) and can be paraphrased by 'in reality'.

(1) *I cannot help thinking .. that Earle's vanity has tempted him to invent the account of her former way of life ... - I dare say she was nothing but an innocent country girl* in
From such contexts arose present-day English contrastive *in fact*, illustrated in (2):

(2) *er you mentioned glucose in fact the one that is actually ... produced [...] Dextrose*  
[1980s, BNC-CG FLY, science lesson]

*In fact* here has become an adverbial marker of contrast, most frequently in initial position. With the development of the connective function came a move from the sentence-final or medial position typical for VP-adverbials to the initial position typical for connective adverbials.

Further development brought about present-day English elaborative *in fact*, which seems to have arisen during the late eighteenth and early nineteenth centuries from the epistemic ‘in reality / in truth’ sense, used to increase hearer belief in contexts where some further, potentially surprising claim follows an assertion⁴, as in (3)

(3) a. *The beast [a rhinoceros] .. kept on an even and steady course, which, in fact, was a kind of pacing* [1785, Sparrman, A Voyage to the Cape of Good Hope translated from the Swedish]  
   b. *and the accident has vexed me to the heart. In fact, I could not pluck up spirits to write to you, on account of the unfortunate business.* [1790, Burns, personal letter]

Such examples seem to have been the immediate precursors to elaborative *in fact*. At this stage *in fact* is unlikely to have had quite the same sense as the present-day English elaborative discourse connective, but rather to have still been epistemic (cf *in truth*). But by the late nineteenth century we find occurrences that are more clearly connective (4). An idea is followed by a rhetorically stronger or more specific idea. To say that a state of affairs obtains *‘in fact’* is to emphasize that one’s statement is true, so that hosts of *in fact*, whether expressing a correction of a false claim or an elaboration of the previous claim, carried emphasis. Here, *in fact* functions in much the same way as *indeed*, which had grammaticalized earlier in a very similar way.⁵

(4) *You have probably heard -- in fact, I have told you myself* [1894, Conan Doyle, *The Stark Munro Letters*]

The Elaborative use is the most frequent use of *in fact* in present-day English, exemplified in (5). The fact that the host of *in fact* is rhetorically stronger lends it greater informational salience within the discourse.

(5) a. *I hate Tech class ... I hate Music too .. in fact I hate most of my lessons* [1980s, BNC-DS, KPG, conversation]  
   b. *I said do you still have your late night on a Thursday so they said yes he said in fact we’re open every night now till 6 o’clock* [1980s, BNC-DS, KBC, conversation]

This elaborative construction is shown in (6).

(6)  
[Claim or stance]  
  [[in fact]  
  [Elaboration of claim/stance]]

A further move is beginning to be apparent, whereby *in fact* introduces a discourse unit that has no immediate coherence connection with the preceding one; rather, *in fact* becomes a type of presentative introducing a new idea or topic (7).

(7)  
*Oh I’ll do it myself ... hundred and fifty ohm .. think I’ve got one of them. ... yeah ... in fact*
Grammaticalizing connectives in English

this particular chassis I've never had er ... never done any work on [1980s, BNC-DS, KC1, conversation]

This seems to be a recent shift and it remains to be seen if and how it develops.

3.1.2 Position of *in fact*

The earliest occurrences in the data of final *in fact* that cannot plausibly be interpreted as the VP-adverbial, but rather connect back to the previous unit, date from the second half of the nineteenth century (8).

(8) Well—the truth is—that our firm has got some dealings with these students—a long account *in fact*—and as a settlement's approaching... [1869, Bernard, *The man of two lives*]

In the data from 1880 to 1920, 13 out of 64 occurrences of *in fact* are in final position as in (9) and (10).

(9) they are beginning to bore me horribly those estimable personae of mine. I am very much annoyed *in fact* -- because they have all got (those few who are left) into such a distressfully lofty atmosphere .. [1889, Dowson, personal letter]

(10) a. the average man's conscience does not begin work till eight or nine o'clock -- not till after breakfast, *in fact*. [1891, Jerome, *Diary of a pilgrimage*]
    b. I came back to-day - finding it supremely triste: did not go near Queen's at all -- nowhere *in fact*. [1889, Dowson, personal letter]

It seems likely that RP connective *in fact*, as in (9) and (10), developed not from the epistemic adverb (cf. ex. (1)) but after LP and medial uses had acquired a connective sense, or at least strong connective implicatures. We hypothesize that connective *in fact* split into LP and RP constructions only once the elaborative sense had become established. This is because, whereas there is a period when initial and medial uses are vague across epistemic and connective, this does not seem to be the case for RP uses, which are all connective. This hypothesis will need to be tested on a larger corpus, because the present data set is too small to draw any firm conclusions, and with other connectives having a similar history.

In both positions, the host unit seems to bear the same type of relationship to the previous segment, expressing a more accurate, specific or stronger formulation. The different positions may therefore correlate with some other factor. The most noticeable difference is that only one RP occurrence (ex. 9) has a full-clause host, the remaining twelve attaching to sub-clausal units as in (10). This finding suggests an association between RP position of the connective and a discourse construction where the second conjunct (the host of the connective) is informationally backgrounded and subordinate to the first conjunct (the previous idea linked back to).

In present-day English, elaborative *in fact* occurs overwhelmingly in initial position. Only 4 occurrences in the LLC corpus are clearly in final position. Of the occurrences in the demographically-sampled section of the BNC (conversation), only 5% are clearly discourse-unit-final (24 occurrences). They occur in contexts where the host of *in fact* can be interpreted as either a greater precision of the previous idea (11) or as a correction of it to a point higher on a scale (12).
Grammaticalizing connectives in English

(11) a. A - I will say it's cut shorter at the back
    B - [unclear] sort of waistcoaty in fact [1980s, BNC-DS, KBK, conversation]
b. we went with Traffens one year the first year in fact we went with Traffens [1980s, BNC-DS, KE2, conversation]

(12) a. so come Monday morning Sunday night in fact Noel said gosh ... [1980s, BNC-DS, KE2, conversation]
b. they had hundreds in there .. thousands in fact I would say [1980s, BNC-DS, KE6, conversation]

In terms of the coherence relation holding between the in fact host and the previous unit, again we find very similar relations with LP in fact (13). In each case the host expresses a rhetorically stronger idea that elaborates on the previous idea. So that in each case in fact introduces a scale on which its host is higher than the idea referred back to.

(13) a. A - can I have a piece of paper please
    B - in fact you can have two [1980s, BNC-DS, KPG, conversation]
b. lots of people claim .. in fact ever such a lot of people claim that they've got communication with the dead [1980s, BNC-DS, KBX, conversation]
c. A - have you got any stamps?
    B - no I don't think I have .. in fact I know I haven't [1980s, BNC-DS, KCX, conversation]
d. I'm tired .. I'm very tired .. in fact I think I'll go to sleep [1980s, BNC-DS, KSV, conversation]

This use is to be distinguished from Contrastive in fact as in (14), where the host is not stronger, but rather denies the proposition in the previous unit altogether and expresses the contrary.

(14) A - you were supposed to do six and you only did four!
    B - in fact erm I'm not supposed to do any number! [1980s, BNC-DS, KST, conversation]

While fragments and clauses continue to be connected by final position in fact to the previous segment, no examples were found of fragments being connected by initial in fact. Examples (15a) and (15c) are both acceptable, while (15b), with initial in fact, is odd:

(15) a. I shouldn't really be here now .. but we had a very quiet surgery ... no one there in fact [1993, BNC, JYE, fiction]
b. ??I shouldn't really be here now .. but we had a very quiet surgery ... in fact no one there [manipulation of (16a)]
c. I shouldn't really be here now .. but we had a very quiet surgery ... in fact there was no one there [manipulation of (16a)]

It seems that the reduced-clause hosts that are compatible with RP in fact are designed to be pragmatically dependent on and informationally subordinate to the previous discourse unit. The connective binds its host to the previous unit and marks the end of the sequence. By contrast, initial connectives tend to act by virtue of their position as presentatives to put their hosts into focus.
In positional terms, *in fact* has thus come 'full spiral', so to speak, insofar as its discourse-connective functions have led it from end position (VP-scope) to medial and initial position and thence 'back' to end position but at higher level (wide scope, RP), in a new function. Out of the prepositional-phrase adverbial, an elaborate discourse connective has grammaticalized to host-initial (LP) position and more recently to host-final position (RP). In PDE, *in fact* can function on any of three levels as VP-adverbial (now rare), as epistemic adverb and as discourse connective.

### 3.3 After all

#### 3.3.1 Evolution of *after all*

In present-day English, adverbial *after all* is found as a temporal VP-adverb, but rarely. It mainly occurs as a sentence adverb encoding counterexpectation (16) and as an adverbial connective signalling a justificative relation (17); that is, its host unit is presented as justifying, or reinforcing the validity of, the previously expressed idea. In this use it expresses 'because' on the speech-act plane.

(16) *oh she’s gone to sleep *after all* [1980s, BNC-DS, KBH, conversation]

(17) *Don’t get your hopes too high or let yourself get too carried away* *after all* you know what people are like [1980s, BNC-DS, KBE, conversation]

The typical justificative construction with *after all* can be sketched as in (18).

(18) [Stance or claim] [[*after all*] [Justification]]

Adverb *after all* originates in a prepositional phrase (19) that slowly coalesces into an adverb over the sixteenth and seventeenth centuries and eventually becomes connective (Lewis 2007, Traugott 1997).

(19) *doctur Whyt bysshope of Lynkolne dyd pryche at the sam masse; and *after all* they whent to his plasse to dener* [1555, Machyn, *The diary of Henry Machyn*]

'Doctor White bishop of Lincoln preached at the same mass; and *after all* they went to his place for dinner'

With the grammaticalization of *all* towards determiner status, and its partial replacement by *everything* and *everyone*, the loss of compositionality becomes clearer (cf. also *overall, in all*, etc.) together with semantic narrowing as the adverb acquires implicatures from its typical contexts. Yet it remains in PDE close to its prepositional use. The data for the historical usage includes all occurrences of the sequence *after all*, since there is no discernable dividing line between the prepositional phrase and the emergent adverb (i.e. almost all occurrences are analysable as *after* + NP).

Counterexpectational and justificative senses of *after all* emerge slowly after a period when *after all* has the senses of temporal 'finally' and the more abstract 'in the end' (cf. 'when all's said and done'). It acquires a modal-epistemic quality, based on the notion that time (events) produces outcomes, and that over time (after events) firmer judgments can be made about states of affairs. A
recurrent context for after all in the late seventeenth and eighteenth centuries is following a conjunction and, naturally given its meaning, introducing a topic closure or the end of a turn. Over the period 1680-1839, just under half of occurrences follow a conjunction, equally distributed between addition or reason (and, for, then) and contrast (but, yet, though) (table 1). After all has no connective sense of its own during the period; it is not tied to any coherence context.

### Table 1. Proportions of 'after all' collocating with conjunctions

<table>
<thead>
<tr>
<th>Period</th>
<th>Proportion</th>
<th>and, for, then</th>
<th>but, yet, though</th>
</tr>
</thead>
<tbody>
<tr>
<td>1680-1719</td>
<td>11/25</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>1720-1759</td>
<td>8/15</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>1760-1799</td>
<td>13/32</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>1800-1839</td>
<td>21/63</td>
<td>10</td>
<td>11</td>
</tr>
</tbody>
</table>

Both the counterexpectational and the justificative uses develop out of recurrence in contrastive and justificative contexts. The first to emerge is the counterexpectational adverb from contexts typically involving a situation where after some time, effort or activity there is no result or an unexpected result. It emerges from adverbials in initial (20a), final (20b) and medial (20c) positions, this last position being typical for adverb placement.

(20) a. I have been studying all this Night long to save Charges; but after all, I find you must be at the Expence of a new Bed [c1700, Ward, The whole pleasures of matrimony]
   b. - if you are in Earnest you are Undone.
      - I am afraid not, says he, for I am really afraid she won't have me, after all my Sisters huffing and blustering. I believe I shall never be able to persuade her to it. [1722, Defoe, Moll Flanders]
   c. I have revolved this Sentence in my Mind till I have quite tired myself, but cannot, after all, find any Meaning in it. [1739, Anon, Review of Hume's 'A treatise of human nature']

After all is temporal in these examples, the contrast supplied by but or by inference. In contrastive contexts the PP/adverb develops into the modal adverb that it seems to be in (21), where a temporal interpretation makes less sense than a counterexpectational sense of 'despite everything / despite indications to the contrary'.

(21) a. let not the unhappy Wretch, who, after all, is your Daughter, want those Necessaries of Life .. [c1741, Richardson, One hundred and seventy-three letters]
   b. There is no reason to suppose that Miss Morgan is dead after all. [1798, Austen, personal letter]
   c. It is more than a fortnight since I left Shanklin chiefly for the purpose of being near a tolerable Library, which after all is not to be found in this place. [1819, Keats, personal letter]

Although (21a) now looks like a justificative use, there is no evidence that it was used as such at that time; the context might nevertheless invite that inference. (21b) is an early example of the modern counterexpectational usage; after all at final position is not in the periphery but it has scope over the whole sentence as it would in initial position, countering a presupposition that there was
reason to believe X. Likewise, (21c) counters the presupposition that the writer expected to find a library.

Unambiguous occurrences of connective (justificative) after all are found only towards the turn of the twentieth century, as perhaps in (22).

(22) I would much prefer to get the War completely over than get leave. After all, in my present job I am not worried by monotony, and I find the work of absorbing interest [1917, Jones, War letters of Paul Jones]

Apparently justificative uses continue in present-day English to be preceded by another connective, usually 'cos / because (23).

(23) I gave mum thirty five pound because after all you know I think she needs it [1980s, BNC-DS, KDN, conversation]

Many occurrences are still vague as to whether the adverb really links back to the previous idea or can be glossed as 'in the end'. These observations are compatible with the justificative after all being a recent innovation that is incomplete.

3.3.2 Position of 'after all'

Counterexpectation after all can no longer occur in initial position, final position being typical (it is not in the periphery, but in end focus). Connective (justificative) after all occurs mainly in initial position (24), which suggests a division that would correspond to the polysemy. But the connective also occurs at RP (25).

(24) Firms will often see merger as an 'easy way out'; after all, nobody in business prefers to face competitive pressure [1990s, public speech]

(25) a. A: tomorrow
    B: yes OK
    A: why not . why not you're free after all [1960s, LLC, conversation]

b. I mean how do you view it? you're a professional after all [1980s, BNC-CG, KRL, radio talk show]

Like RP elaborative in fact, RP justificative after all is relatively infrequent (fewer than 1 in 10 of occurrences). Unlike RP in fact, it does not seem to be acceptable with a sub-clausal host. But it resembles RP in fact in that a relationship is unambiguously marked -- of justification in the case of after all -- between the two conjuncts, which is not always so for after all in initial position. It is also the case, as for in fact, that the LP and the RP occurrences express the same rhetorical relation between the conjuncts that are linked.

3.4 Findings for 'in fact' and 'after all'

The developments so far of in fact and after all can best be viewed in the light of the broad approach to grammaticalization mentioned in section 2.1. In this approach, grammaticalization is not restricted to items becoming more fixed and obligatory, but encompasses items becoming more
grammatical while undergoing extension and scope expansion. All cases lead to greater abstraction and increased productivity: expansion of the host class, and thereby opportunities for the item to occur, favours increased frequency.

Traugott (2015) argues for a gradient distinction between 'core clause' and periphery. Data on the evolution of *in fact* and *after all* clearly support the gradient hypothesis and a gradual evolution. This suggests they cannot be considered to have undergone 'co-optation' in the sense of Heine (2013). From the perspective of hindsight it appears that a reanalysis occurred, for example, from epistemic VP-adverb *in fact* to (wide-scope) sentence adverb *in fact* with discourse-connective function. Such reanalysis is sometimes discussed as though it progressed incrementally, with increasing frequency of the new analysis as it becomes established alongside the old one. But tracing the evolution 'forwards', so to speak, often reveals, rather, a period of over-extension or over-generalization which only later may settle into a pattern. During the actualization period, both the old and the new analyses obtain simultaneously; that is, "the speaker makes both (or many) analyses" (Harris and Campbell 1995: 82). As has been seen above, *after all* cannot be said to have neatly reanalysed from temporal adverbial to modal adverb and connective. It is not so much that it retains aspects of its origin in its newer uses, as that the older use and the newer one even now can both be seen to obtain in many occurrences. Such dual analysis can persist over many decades or more, and both *in fact* and *after all* illustrate this phenomenon.

4. Grammaticalization and information structure beyond the sentence

Final adverbial discourse connectives are a comparatively recent phenomenon. The earliest, according to Lenker (2010: 200), was *however* in the seventeenth century. Final connective *then* and *though* date from around the same time. Lenker suggests that the construction may be specific to present-day English (2010: 202). The overall RP 'slot' is well established for a range of non-connective but speaker-oriented and interactional discourse markers and comment clauses. So it is plausible that the new connective construction is becoming productive and that further existing connectives are aligning analogically with those early ones. The trajectories of both *in fact* and *after all* from PP to VP-adverb and then modal adverb and connective are similar. And both connectives have recently started to occur at the right periphery of their host unit, where they become 'retrospective' markers. But there are important differences between the two developments. *After all* has developed a little more recently than *in fact*. It remains closely linked to its origins as a temporal PP adverbial and retains structural and functional similarity to *after all that, after all's said and done*. In frequency terms, it is overshadowed by its temporal cousin. Moreover, the connective function is not as robust as that of *in fact*, but is often bolstered by, if not carried by, another connective. Connective *in fact*, by contrast, vastly exceeds in frequency VP-adverb *in fact*; and its meaning has now become highly bleached. What are the possible motivations for the development of *in fact* and *after all* as connectives at RP as well as LP? It is suggested that processing and information structurte are factors in the emergence of the two patterns, and that the pattern with a RP connective (e.g. [p. q *in fact]*) constitutes a more complex and arguably more grammaticalized construction than the LP pattern.

To explain the linear ordering of main and subordinate clauses in complex sentences, competing motivations have been proposed, including discourse planning advantages for the speaker, ease of processing by the hearer, and the organization of information flow (including thematic structure and relative weight). While 'adverbial clause last' has processing advantages for the speaker (it requires less utterance planning) and hearer (it requires lower memory load), 'adverbial clause first' results in
better information flow, because it links back to the previous discourse and provides the ground for the following main idea (Diessel 2005). If we compare now the LP and RP connective constructions, it seems that LP position will have processing advantages for the hearer, insofar as the connective will guide his interpretation of the host unit (Lenker 2010:198). Final connectives seem to run counter to this natural order, as the hearer must hold the host unit in suspense, only at the end discovering how the speaker makes it relevant to the previous discourse. For the speaker, however, the idea comes first, so that there may be a processing overhead for the LP construction, which is likely to have developed in dialogic contexts where the link is to the previous turn (is the speaker's reaction to her interlocutor's turn).

A second possible factor in the ordering is the speaker's structuring of the discourse information. A number of models of discourse structure address the relationship between coherence relations and discourse-level information structure (relative informational prominence). Some models, such as Rhetorical Structure Theory (RST) (Mann and Thompson 1987), conflate coherence relations with information structure. The early version of RST thus posits two types of coherence relation: symmetrical, relating two discourse units that carry equal prominence, and asymmetrical, where one unit acts as a 'satellite', or support, to the 'nuclear' unit. A large majority of the relations identified by RST are of this second, asymmetrical type. An example is the Justify relation, where a claim (the 'nucleus') is supported by a 'satellite' expressing the speaker's warrant for the claim, thereby increasing the hearer's acceptance of it. In practice, such conflation has been found to be too constraining. To remedy this, Carson and Marcu's (2001) RST model allows some relations to be symmetrical or asymmetrical, and allows some asymmetrical relations to express the relation as nucleus or satellite. Others have gone further to decouple coherence relation from information structure, as in the 'multi-level analysis' model proposed by Stede (2008), where nuclearity is removed from the coherence relation and is redefined in terms of optional support relations between discourse units. Constructions such as those we have seen with in fact and after all seem to support this last type of model, where relation and relative salience are separate levels, and one relation can co-occur with more than one information contour. As has been seen, LP and RP connectives in fact and after all each expresses a particular rhetorical relation, respectively elaboration and justification, independently of position. But the relative informational salience arguably does change. And at RP the connective is more integrated prosodically into the host unit than at LP. These findings are in line with those found, in a rather different rhetorical context, by Gentens et al (2016) for the development of no wonder; they also emphasize "the central role played by larger rhetorical structures in the grammaticalization" (2016: 151).

The pattern, or discourse construction, that involves a connective at LP overlaps with the complex sentence construction containing a final-position adverbial clause. Unlike initial adverbial clauses, final ones need not beinformationally subordinate to the main clause. Subordinating conjunctions such as because, so that, although, whereas and so on, when the clause they introduce follows the main clause, can introduce an idea of equal weight to that of the main clause. In his analysis of finite adverbial clauses, Diessel (2005: 464) finds that "final because-clauses ... basically function like independent assertions: they tend to provide new information". Connectives like in fact and after all at LP function very much like because in a final adverbial clause. They link their host to the previous idea and at the same time they are points of departure, opening a turn or move or a new idea, and effectively functioning as presentatives that put the following idea(s) into focus. Their hosts are overwhelmingly finite clauses. Like other initial, pre-subject adverbials, these connectives provide grounding for the idea that follows. They have grammaticalized to encode a relatively backgrounded idea (the coherence relation) and they place what follows in focus.
In final position the connectives have the same linking function, but they close a sequence of at least two ideas, the second acting as a comment on the first. It has been seen that RP \textit{in fact} (like some other RP connectives) is found with sub-clausal units which do not occur with LP \textit{in fact}, and which are 'tacked on' to the previous unit, on which they elaborate or comment. Final position may therefore be partly motivated by information structuring. Such a development is arguably to be seen as a further grammaticalization. It results in a 'subordinating discourse construction' that links two units (two independent clauses or a clause and a fragment).

Such a distribution, with both LP and RP positions available for certain connectives, would offer speakers a choice of a presentative or a backgrounding structure. The RP construction is more complex than the LP one in that the host unit seems to be more tightly bound to, or more pragmatically dependent on the previous idea than is the case for hosts of initial connectives. If the two constructions continue to co-exist, over time the LP and RP connectives may start to drift apart towards further polysemy.

5. Conclusion

Discourse-level information structuring involves not only thematic progression (old/new) but also relative informational salience: how information is marked as foregrounded or backgrounded with respect to some other information. Initial position, including the left periphery, has been shown in the past to be associated with particular roles in discourse information structure: it is used for markers of new discourse frames, including topic change, and can also have an attention-seeking and presentational function, serving to place what follows in end-focus position, thereby foregrounding it. By contrast, a connective at RP marks the end of a comment on (or a modification of) an idea that is pragmatically subservient to the previous idea. The hypothesis explored here must be tested on other connective and non-connective peripheral expressions, both of clausal origin such as \textit{I shouldn't wonder} or \textit{mind you} and non-clausal, with a view to better identifying the role of information structure in their development and ongoing evolution.

Many wider questions remain open regarding, first, the interplay between coherence relations and discourse information structure in grammaticalization, and, second, the grammaticalization of information-structuring functions. For future work on the diachrony of English connectives, there is scope for exploring the interactions among their connective, information-structuring and interactional functions.

Endnotes

1. A distinction is to be drawn, from the point of view of language change, between discourse markers that have become 'stranded', so to speak, at the end of an utterance due initially to ellipsis of their host unit, and those that occur in end-position immediately following their host unit. The former are exemplified in the concessive constructions of (1). \textit{But still} in (1a) links its host, \textit{that's life}, to \textit{We'll miss each other}, and is at LP. The second conjunct (the host unit) may be ellipted, resulting in occurrences such as (1b) (spoken) and (1c) (written), where the marker is a 'hanging implication' (Thompson and Suzuki 2011: 670). The marker's host unit has at first to be pragmatically inferred, but the marker soon becomes autonomous (in its own tone unit) and ends up at RP. The \textit{but still} of (1a) may then be analysed as being at LP and/or RP.

(1)  
   a. "We'll miss each other, \textbf{but still}, that's life". [1980s, BNC CFY, fiction]  
   b. \textit{I still can't carry anything heavy in it \textbf{but still}.} [1980s, BNC-DS KBB, conversation]
c. Worrying in advance can stave off disaster. Sounds better in Latin, perhaps, but even still. [2015, The Financial Times, 14/08/2015, feature article]

No claims are made here about 'stranded' markers, other than to point out that over time they can of course become integrated and indistinguishable from markers that have 'moved'.

2. Izutsu and Izutsu (2013:232) go so far as to suggest that utterance-final position "may be somehow exploitable for (peripheral) intersubjectivity marking in most languages"

3. Information structure is usually taken to involve intrasentential information relations such as theme, rHEME, topic, focus, background, contrast and so on. Information relations beyond the clause or sentence have traditionally been dealt with separately and under different labels. Polanyi et al (2003), among others, distinguish 'information structure' operating at clause level from 'discourse structure' at the level of interclausal information relations. Other labels are 'discourse relations', 'coherence relations' and 'rhetorical structure'. However, the dividing line between sentence and discourse is not so clear-cut: grammar and discourse are interdependent (Mithun 2005, 2016). Many connectives function at the interface between syntactic dependency and discourse structure. Both spoken and written discourse is made up of both units that have 'standard' sentence structure and those that do not, including multi-conjoined clauses or concatenated clause complexes, isolate non-finite clauses and all kinds of fragments that cannot be explained away by ellipsis. Information structure, whether at sentence or discourse level, concerns the relative salience given by speakers to the elements of their discourse, so we will prefer the term (discourse) information structure to 'discourse structure' which sounds analogous to 'sentence structure'.

4. Schwenter and Traugott (2000) suggest that elaborative in fact developed directly out of Contrastive in fact, via a reanalysis on to the rhetorical plane where elaborative in fact expresses a rhetorical contrast, but our data did not reveal any evidence for this (cf. Lewis 1998); rather, there was evidence of use of the still-epistemic adverb in contexts such as those in (4).

5. Fischer (2007:285) claims that it is likely that the change of in fact into a "sentence-adverbial/pragmatic marker was via analogy with indeed rather than via any form of grammaticalization, as the suddenness of the development (in comparison to indeed) suggests". But Fischer adduces no data to substantiate the suddenness claim, and our data suggest on the contrary that the development was gradual (cf. Lewis 1998). Fischer suggests that a major mechanism for the development of initial-position (LP) pragmatic markers (discourse connectives) is the topicalization (2007: 285, 287, 294-6) of a VP-adverb which later, by analogy with other LP expressions, acquires scope over the following proposition by virtue of its initial position, but she does not discuss final-position (RP) connectives / pragmatic markers.

6. Heine (2013) mentions in fact, but cites only one example, which he takes from Traugott 1995. This example is problematic: only the host is given, and when we look at the Hume text, and especially when we take account of other contemporary (mid-eighteenth century) occurrences, it is clear that this occurrence of in fact is not connective; nor can it be an example of 'cooptation' in Heine's definition. Indeed, Heine seems to suggest that such connectives are different from the kinds of delexical expressions with which he illustrates co-optation (2013: 1234).

7. Harris and Campbell (1995; 82-89) discuss three types of evidence that multiple analyses continue in individuals' grammars for some time. See also Hankamer (1977) and Dowty (2003).

**Appendix: Data sources**

**Historical English:**

The Helsinki Corpus of English Texts (distributed by ICAME)
A Corpus of English Dialogues 1560–1760. 2006. Compiled under the supervision of Merja Kytö (Uppsala University) and Jonathan Culpeper (Lancaster University).
A Corpus of Late Modern English texts, v. 3, compiled by H. De Smet, H. Diller, and J. Tyrkkö
Archer: A Representative Corpus of Historical English Registers
London-Lund Corpus of spoken British English (LLC). (Distributed by ICAME)
Additional historical texts C18th-C20th

**Present-day English:**

References


Heine, Bernd. 2013. On discourse markers: grammaticalization, pragmatization, or something else? *Linguistics*
51(6), 1205-1247.

Heine, Bernd, Gunther Kaltenböck and Tania Kuteva. 2015. Some observations on the evolution of final particles. In Hancil et al. (eds.), 111-140.


Traugott, Elizabeth C. 2013. 'I must wait on myself, must I?’ On the rise of pragmatic markers at right periphery of the
clause in English. Talk given at Lund University, 4 September 2013.

Traugott, Elizabeth C. 2014. On the function of the epistemic adverbs 'surely' and 'no doubt' at the left and right peripheries of the clause. In Beeching and Detges (eds.), 72-91.
