“All that he endeavoured to prove was ...”: On the emergence of grammatical constructions in dialogal and dialogic contexts
Elizabeth Closs Traugott


1. Introduction

In the present paper I discuss some of the kinds of evidence we can find in historical texts for the emergence of grammatical constructions, and the kinds of discourse contexts in which these changes may have arisen. My particular example is the development of the WH- and ALL-“pseudo-clefts”, as in What/All I said was X, What/All I did was X. In the course of this discussion I also aim to provide an example of the ways in which combining the insights of grammaticalization and construction grammar can help account for micro-changes in morphosyntax (for some earlier studies with similar objectives, see Bergs and Diewald Forthc, Traugot t 2008, Forthc, Trousdale Forthc a, b).

I start with some background assumptions and terminology (section 2.). In section 3. I sketch out the development of ALL- and WH-pseudo-clefts, and discuss in which discourse contexts each construction appears to have arisen. Section 4. outlines what the dual perspectives of grammaticalization and construction grammar suggest not only for the pseudo-clefts but for the emergence of constructions in general.

2. Some background assumptions and terminology

Here I introduce the distinctions that I am assuming between innovation and change (2.1.), and between “dialogal” and “dialogic” contexts for language use (2.2.). I also outline the views of grammaticalization and construction grammar relevant for this paper (2.3.).

2.1. Innovation and language change

I assume that language change arises in language use (J. Milroy 1992, 2003, Croft 2000), i.e. in “practices of speaking” (Andersen 2006: 65). It starts as an ad hoc innovation by an individual in the speech (or writing) situation, but counts as a change only when the innovation is adopted by others (Weinreich, Labov, and Herzog 1968).

A further assumption is that innovations may occur throughout a language user’s life-time, not only prior to puberty. While more radical morphosyntactic innovations leading to structural changes may be privileged in early child language acquisition, some occur later, though at a less rapid rate (see Bergs 2005). Innovations are made by speakers as well as hearers, that is, in production as well as perception. In language change there is normally no intention to change some aspect of language (Keller 1994 [1990]). What speakers and addressees intend is to negotiate common ground (see Clark 1996, Croft 2000) and achieve certain ends such as getting each other to listen, exchange information, or to behave in certain ways, but not normally to change the language.

Negotiation of common ground is central to my view of discourse analysis. I will be adopting not only a primarily neo-Gricean information-structure-based approach to the data in question (see Traugott 2004), but also a more interactional approach (see e.g., Mann and Thompson 1992, Ford 1994).
2.2. Dialogal and dialogic discourse contexts

The focus of this volume is dialogue. The term can be understood in two ways. One is general speaker-hearer interaction, in which participants negotiate meanings through alignment, and accommodation (see Cooper, This volume). The second is more narrowly understood as turn-taking involving consecutive patterns of the type in (1) (see Sacks, Schegloff, and Jefferson 1974):

(1)  A. Current speaker selects next speaker.
     B. Next speaker (selected or self-selected) takes a turn.
     A. First speaker continues.

Since in historical work we are dealing with written texts, we need to think not only of a Speaker-Addressee (SP-AD) dyad, but also of a Writer-Reader (W-R) dyad. We may combine these as SP/W-AD/R. Unless we are to restrict ourselves entirely to represented conversations, we also need to think of SP/W as negotiating meaning through dynamic, interactive, discourse expressed by one individual (the writer, represented narrator, etc.).

Recently, as researchers in historical pragmatics have sought to identify contexts in which particular changes occurred, there has been considerable interest in turn-taking as a context for the development of micro-changes, as well as negotiation of meaning within a turn (see e.g., Detges and Waltereit 2003, Detges 2006), and this is one of the discourse contexts that I will discuss here. Since it involves interlocutors, it is language-external.

Another context, this time a language-internal one, that I will be discussing is dialogicity. As has long been recognized, especially in work associated with Roulet (1984) and Ducrot (1984, 1996) (see also more recently, e.g., Schwenter 2000, Nølke 2006), discourses may be relatively homogeneous in orientation (i.e. closely aligned toward some argumentative conclusion), or multiply perspectivized, in both cases either within or across turns. Multiply perspectivized expressions may be said to be “dialogic” or “polyphonic”. In the work cited above, attention is paid to expressions that code multiple perspectives, for example but vs. and, or focus particles like only within a clause. Here, however, I will be concerned primarily with prior contexts that either include such expressions and/or are intended to be contesting, and therefore introduce multiple perspectives.

Historically there are two issues to consider here. One is that cross-linguistically expressions that code multiple perspectives typically derive from relatively neutral or singly perspectivized expressions (e.g., but < butan ‘on the outside of’, see Nevalainen 1991), only < ‘singly’ (see Brinton 1998), in fact ‘epistemic adverb’ < ‘in practice’ (see Schwenter and Traugott 2000). The other is that such dialogic meanings can be shown in many cases to arise in linguistic contexts where SP/Ws are negotiating non-aligned perspectives, i.e. are presenting opposing arguments to others or to imaginary interlocutors. It is the latter issue that will be the second focus of my attention.

A useful framework within which to discuss dialogal and dialogic contexts is to distinguish numbers of participants and numbers of perspectives invoked. Drawing on Roulet and Ducrot, Schwenter (2000) proposed the distinction between monologal-dialogual interaction (this pertains to numbers of interlocutors), and monologic-dialogic perspectives (this pertains the orientation of the speaker’s move (alignment or disalignment, as in refutation, counter-expectation, or adversativity))
and to the numbers of perspectives invoked. Treating them as independent parameters allows us to account for the fact that relatively homogeneous or relatively polyphonic perspectives may be taken within or across turns. The intersection of the two dimensions may be modeled simplistically as in Table I. While Table I is based on Schwenter (2000), it should be noted that he is primarily concerned with dialogicity within the clause rather than in prior context:

Table I. Speakers and viewpoints (based on Schwenter 2000: 260)

<table>
<thead>
<tr>
<th>Number of speakers</th>
<th>Number of viewpoints in context</th>
</tr>
</thead>
<tbody>
<tr>
<td>One: monologual</td>
<td>One: monologual/monologic</td>
</tr>
<tr>
<td>Two: dialogal</td>
<td>One: dialogal/monologic</td>
</tr>
<tr>
<td></td>
<td>Two: dialogal/dialogic</td>
</tr>
</tbody>
</table>

To date, most of the historical work on interaction in dialogue has been focused on the turn. Over the years some hypotheses have been presented arguing that a particular construction has arisen out of turn-taking in dialogue. For example, in his article on conditionals as topics, Haiman (1978) draws on Jespersen’s (1965 [1940]: 374) suggestion that conditionals arose out of questions with implied positive answers to account for both the alleged topicality of conditionals, and their tendency to be expressed by forms that are interrogative. Haiman proposes a “mini-conversation” of the type:

(2) A: Is he coming?
    B: (Yes.)
    A: Well then, I’ll stay.

which “then functions as the basis for further discussion” (Haiman 1978: 571). This assumes that A or some overhearing participant takes the A/B pair and produces a new utterance that combines the pair, and that this combined pair later becomes conventionalized as a new structure in the linguistic system. Whether this kind of alignment and coordination among speakers and hearers actually results in the innovation of conditionals with interrogative forms remains to be tested, to my knowledge. vii

More recently, seeking to go beyond the rather general (and largely information-structure-based) theory of invited inferencing as a major source of change (see Traugott and König 1991, Traugott and Dasher 2002), Detges and Waltereit (2003), Detges (2006), and Waltereit (2006) have discussed the development of discourse markers like Italian Guarda! ‘look’ > ‘see’/’self-selection marker’, viii and of the cliticization of tonic subject pronouns in French and of tonic object pronouns Spanish. They argue that speakers self-select at turns by using attention-getters in “illegitimate” ways, e.g., Guarda! when there is nothing to look at, and of tonic clitics when there is no contrast. As the self-selection strategies are used more frequently, and become routinized, the original meanings are lost (“bleached”). The authors emphasize that the forms preexist and propose that it is the implicatures that arise specifically in turn-taking contexts that favor change, while the interactive turn-taking strategy motivates the change. Waltereit and Detges (2007) further distinguish turn-taking strategies (“What are we going to do next?”) from those negotiating common ground (“What do I believe that you believe concerning the felicity of my speech act?”). They propose that discourse markers are the product of the first strategy, whereas other types of change, such as the development of epistemic
uses of adverbs like *in fact* or French *bien* ‘surely’ have a different, contesting function, and are the product of the second. The distinctions I am drawing between dialogal and dialogic contexts are clearly related to those proposed by Detges and Waltereit in various writings, but my emphasis is on the way in which contesting contexts may give rise to a broader range of expressions than epistemics, and on how they may interact with turn-taking.\(^i\)

Although monologal-dialogic perspectivizing is well attested in older written materials, unfortunately many of these texts do not provide much direct evidence for the self-selection that is the focus of many of Detges’ and Waltereit’s papers. While they could be identified in spoken corpora collected in the last hundred years or so, and in novels, dramas, etc. that represent such speech, turn-taking that occurs in earlier texts may reflect conversational strategies only indirectly (Culpeper and Kytö 2000). Early courtroom trials can be a useful source of dialogue. While contemporary trials are of a largely ritualized nature, with pre-set rules for turn-taking where self-selection is limited, early trials present more oral features, and are more useful for analysis of dialogal turn-taking than might at first be thought (see Culpeper and Kytö 2000, Kryk-Kastovsky 2000).\(^x\) This is because in the seventeenth and eighteenth centuries defendants did not have defense lawyers to speak for them (and scribes appear to have represented speech fairly well). Basing her analysis on courtroom trials from 1640-1679, Archer (2006) provides figures for initiation, response, and follow-up by judges, defendants, and witnesses, and shows that defendants responded to direct elicitation by following it with another elicitation device almost as frequently as judges. This means that aspects of self-selection are far more likely to occur in early trials than in those of the present day. In what follows I will, draw examples not only from texts with represented dialogue (trials, drama, and narratives) but also from monologal reports.

2.3. Grammaticalization and construction grammar

I will be embedding my discussion of the development of the WH- and ALL-pseudo-clefts in the larger contexts of grammaticalization and construction grammar. Grammaticalization has been defined in various ways (see e.g., Bisang, Himmelmann, and Wiemer 2004, Brinton and Traugott 2005, Gelderen 2004, Heine, Claudi, and Hünnemeyer 1991, Hopper and Traugott (2003 [1993]), Lehmann 1995 [1982], Roberts and Roussou 2003). Here I take grammaticalization to be:

The change whereby in certain linguistic contexts speakers use parts of a construction with a grammatical function. Over time the resulting grammatical construction may continue to be assigned new grammatical functions. (based on Brinton and Traugott 2005: 99)

The developments under consideration in this paper concern an already grammatical construction which involves grammaticalization/constructionalization into another construction without lexical origins (see also Lehmann Forthc\(^ii\)), and therefore gives insight into alignments with topic and focus, in this case exhaustive focus patterns.

Some researchers on grammaticalization associate it with a variety of reduction processes (e.g., Givón 1979, Lehmann 1995 [1982]), or increasing dependency (e.g., Haspelmath 2004). The assumption is that “grammar” is restricted to “core” structures (syntax, semantics, phonology). However, if one regards grammaticalization as the morphosyntactic change whereby grammatical material is formed, and considers grammar to include “higher” discourse structures, a less
reductive perspective can be adopted (see Tabor and Traugott 1998). Himmelmann (2004) suggests that the following three types of expansion are criterial for grammaticalization:

(3) a. Semantic-pragmatic expansion, e.g., when articles arise out of demonstratives they may take on uniqueness functions dependent on the larger situation, including encyclopedic knowledge (the president), and may be used in associative anaphoric contexts (cf. a house – the front door, where the definite article is licensed by association with house); neither of these contexts is available for demonstratives.

b. Syntactic expansion, e.g., emerging articles develop first in core subject and object argument positions, only later, if ever, in adpositional ones.

c. Host-class expansion to more parts of speech, e.g., a grammaticalizing form will increase its range of collocations with members of the relevant category (the range of be going to expanded from activity to stative verb subclasses when it was reanalyzed from motion with a purpose to futurity).

Like grammaticalization, construction grammar has been defined in various ways (for overviews, see e.g., Croft and Cruse 2004, Fried and Östman 2004, Langacker 2005). Construction grammar is construed primarily in synchronic terms. Key for our purposes here are the following characteristics (see Goldberg 2003, 2006):

• Form and meaning are paired as equals,

• Grammar is conceived as holistic, i.e. no one level of grammar is “core”,

• Grammar is usage-based, i.e. grounded in speakers and utterances,

• Individual constructions are independent but related in a hierarchic system with several levels of schematicity and may intersect.

Importantly, construction grammar assumes that parts of a construction are not assembled on-line. Rather:

“[A] construction represents an automated, routinized chunk of language that is stored and activated by the language user as a whole, rather than ‘creatively’ assembled on the spot”. (De Smet and Cuyckens 2007: 188)

Since construction grammars treat all linguistic elements from morpheme and word to clause as constructions (Goldberg 2003), grammaticalization intersects only with the development of grammatical constructions (Noël 2007). Langacker (2005) has suggested that three factors are crucial in thinking about a grammatical construction: grammatical constructions must be schematic, productive (presumably in terms of tokens, since frequency of use is often mentioned as a criterion, see Goldberg 2006), and not fully compositional (see also Bybee, Perkins, and Pagliuca 1994).

It is useful to posit for each construction three type levels and one token level to capture similarities and differences between constructions under discussion in terms of schematicity (Traugott 2008, Forthc; also Fried Forthc, Trousdale Forthc a, b):
(4) a. Macro-constructions: high-level schemas, e.g., ditransitives,
b. Meso-constructions: sets of similarly-behaving micro-
constructions, e.g., give Obj Obj2, send Obj Obj2, which have
prepositional variants with to, as distinct from those that have
prepositional variants with for, e.g., buy Obj Obj2,
c. Micro-constructions: individual construction-types, e.g., give Obj
Obj2 as distinct from send Obj Obj2,
d. Constructs: empirically attested tokens of micro-constructions;
these are the locus of innovation.

The type levels are meant to characterize “family-resemblances”. They may
have further sub-types. Most importantly they form networks with other
constructions, allowing for partial matches across constructions (Trousdale Forthc b).
It should be noted that this view of constructions emphasizes the smaller parts out of
which a larger construction is constructed (subject to inheritance hierarchies). Some
degree of construction-internal accessibility is also assumed, i.e. the “chunk” is not a
rigidly fixed entity; if it were, it would not be subject to variation and change. It
should also be noted that constructs are characterized in (4d) as the locus of
innovation. When such innovations are conventionalized by some set of speakers, a
micro-construction emerges, and this is a change.

The term “schema” used above deserves discussion. Focusing on synchronic
cognitive systems available to individual language-users or members of a close-knit
speech community, Langacker (1987: 371) uses “schema” as an abstraction
compatible with all its members, and membership of which “is not a matter of
degree”. However, in the grammaticalization literature, which focuses on
generalizations about shifts of members of categories over time, the term is used very
differently. It is equated with general patterns of change (also referred to as “clines”).
An item may be more or less prototypical for its category over time and speech
communities, and therefore membership of a schema may be more or less
prototypical. For example, a number of binominal partitive expressions such as a bit
of, a lot of, a shred of, and a deal of, came over time to have quantifier and degree
modifier polysemies, as in a bit of pie (partitive) vs. a bit of a cheat (quantifier), a bit
prettier (degree modifier) (see Traugott 2008, Forthc). The first expression to undergo
this change was a deal of. In Old English dæl was a mem-
ber of the category of
ominals referring to parts.\textsuperscript{xii} In Middle English, especially in the context of
great, it came to be used to express quantity (‘amount’) rather than ‘part’ and, in the form a
deal came to function like adverbal degree modifiers such as quite, a lot:

(5) Why can't gentlefolks wroit like Ned Tiller oop at th'Red Lion - printin'
loike. It's easier to read, and a deal prettier to look at. (1863 M. E.
Braddon, Aurora Floyd, Ch. 22)\textsuperscript{xi}i

By the twentieth century a deal of became largely restricted to contexts with a
preceding adjective like good, great, and especially to a great deal of money/trouble
and similar formulaic, prefabricated expressions. A deal of participated in a general
change from binominal partitives such as a lot of, a bit of, a shred of to binominal
degree modifiers with quantifier readings, but only in a marginal way. Therefore it
participated in the schema partitive > quantifier > degree modifier only to a limited
extent. In Standard English it never became aligned to the meso- and macro-level
degree modifiers and eventually ceased to participate in the quantifier construction
except in fixed expressions, whereas *a lot (of)* and *a bit (of)* became well established and underwent changes to adverbial status (for example, they can be used as free adjuncts in response to questions, as can more prototypical degree modifiers like *quite*, and *indeed*).

While schemas of the synchronic type discussed by Langacker may, by hypothesis, be accessible to speakers, diachronic ones cannot, except indirectly (e.g., by drawing inferences from polysemy relationships or age-graded variation). Diachronic schemas are to be understood not as conceptual imprints on the mind but as analysts’ generalizations (Andersen 2001, 2006) concerning steps by which changes from one category to another may emerge over time, subject to general constraints of language acquisition. It is the business of the historian of language to seek out not only individual changes that might be accessible to speakers during their life-times, but also to find general patterns or schemas that, from the perspective of the linguist, can be seen to emerge over generations, centuries, even millennia. As McMahon has said:

“To understand language change as well as we can, we have to deal with two different levels all the time, that of the speaker, and that of the linguistic system … when it comes to language change, linguists need to stand outside what is going on to understand it. That is what historians are for”. (McMahon 2006: 148, 175)

3. A case study: WH- and ALL-pseudo-cleft constructions

WH- and ALL-pseudo-clefts, as in *What/All I wanted was a Ninja Turtle*, are micro-constructions of considerable interest for several reasons. For one, they involve discourse-structuring of focus elements, and are therefore closely related to information structuring and rhetorical strategizing. Almost all research to date on pseudo-clefts has been on WH-clefts and TH-clefts (the latter will not be discussed here). Some typical (for the most part constructed) examples are:

```
(6) a. What Martin ate was the banana. (WH-cleft)
b. What Martin did was (to) dance the mambo. (WH-cleft)
c. All Martin ate was the banana. (ALL-cleft)
d. All/*Everything that one has to do is to start training earlier. (ALL-cleft [BNC, Kay 2002])
```


- Two clauses, one of which is a relative, one of which involves a copula.
- Uniqueness and contrastiveness: elements named by the clefted constituent are construed in terms of an exhaustive, exclusive listing (Martin ate only the banana, not the peach, grapes, etc.); they are dialogic.
- Givenness: some part of the construction (typically the relative) must be given or at least recoverable (Martin ate something).
- Specificational/identifying focus: the complement of the copula is specific and referential (not ascriptive or non-referential) (see Patten 2007 for detailed discussion of specificationality); compare (6) with (7):
(7)  a. What she did is a shame. (non-referential nominal)  
     b. What she did was laughable. (ascriptive adjective)  
     c. What he said was laughed at. (ascriptive passive)  
     d. All/Everything that I command is yours now. (= ALL-cleft;  
                    ascriptive [BNC, Kay 2002])

There are also “reverse clefts” such as:

(8)  a. The banana was what Martin ate. (reverse WH-cleft)  
     b. The banana was all Martin ate. (reverse ALL-cleft)

Depending on their context, these may have a “given-focus” (“topic-comment”) or all  
“focus” structure (Hedberg and Fadden 2007, discussing WH-cLEFTs). They therefore  
are information-structurally different from canonical clefts.

WH- and ALL-cLEFTs differ from IT-clefts in that their focus may be a clause  
as well as an NP or PP. WH-cLEFTs differ from other cLEFTs in that they can be used  
cataphorically to introduce an upcoming clause (e.g., What I am going to talk about is  
...) in contexts such as a lecture where there is an expectation that the speaker will  
talk about something. More importantly for purposes of this paper, they are often  
associated with questions, for example, (6a) could be thought of as an answer to the  
question in (10):

(9)  What did Martin eat?

Indeed Den Dikken (2006) rejects the relative clause analysis of WH-cLEFTs, arguing  
that WHAT is not a free relative but an embedded question.

Most work on pseudo-cLEFTs has been conducted from the point of view of  
(relatively) formal semantics or of information-structurinG. However, investigating  
WH-cLEFTs from the perspective of interactional conversation, Kim (1995) and Hopper  
(2001) have suggested that WH-cLEFTs have rather different characteristics than those  
identified above. Kim argues that in conversation WH-cLEFTs often do not refer to  
information retrievable from prior context. They are primarily used to expresses  
counteractive stance, specifically disagreement with the addressee, or topic-shift. He  
also argues that they are closely related to left-dislocations. Hopper argues that they  
play a significant role in turn-taking: WH-cLEFTs are used “to delay an assertion for  
any of a number of pragmatic reasons” (p. 111, italics original). He suggests that,  
since most examples are not “complete”, and do not have a fully developed focus,  
they are not primarily motivated by the desire to highlight the focus-constituent, as is  
usually claimed. Basing his analysis on the COBUILD corpus, he says the ideal locus  
for using the WH-cLEFT construction is the turn, where the listener’s attention must be  
held, or shortly after, where keeping the floor is important, as in (10):

(10)  There’s n there don’t seem to be a r- real need. And in defence I mean  
what snakes or what animals try <pause> like what most animals try to  
do is if they tha have got a poisonous property is another animal  
attacks them they give them er a dose of venom… (Hopper 2001: 115;  
italics original)
Hopper suggests that speakers’ reasons for using WH-clefts in the data include not only delaying an assertion in order to keep the floor, but also:

“impressing the listener with the “social” significance of something about to be said, and making the listener aware that what follows is part of a considered argument worthy of attention and not a casual comment”. (Ibid.: 124)

Similar findings have been made with respect to conversational Brazilian Portuguese (Lilian Ferrari, p.c.) and French (Jullien and Müller 2007).

Unlike WH-clefts, ALL-clefts have received very little attention, but it is standardly noted that they differ from WH-clefts in being evaluative. *All* is not equivalent to ‘everything’, but imposes a “below expectation” scalar reading on what follows (Kay 2002). It is understood as convertible with *only* (i.e. it is “downward inferential”, Horn 1996: 18).

In the one detailed study of ALL-clefts known to me, Bonelli (1992), using the COBUILD corpus argues that they correlate with “change of posture”, especially when there is a shift in subject, as in:

\[(11)\]

\[a. \] then, when the call was finished, *all we could hear on lifting was* … (Bonelli 1992: 32 [COBUILD: 10 Million Corpus])

\[b. \] they owed me plenty, but *all I wanted was the right to develop my ideas* (Ibid.)

She also identifies the following discourse moves: retrospection (encapsulating prior text), and prospect (pointing to subsequent text, i.e. cataphora) (p. 32). With respect to attitudinal standpoint, she identifies “positive attitude” (a course of action is simple, perhaps unimportant and not worth worrying about), negative (evaluating a fact or result as unfortunate, undesirable, or insufficient), and damage limitation (admission, with the suggestion that there were no other alternatives) (p. 33). She does not mention correlation with turn-taking. Rather, the functions she identifies often occur within the speech of one individual, and may be embedded in dialogic contexts, as in (11b).

Given these analyses of WH- and ALL-clefts in contemporary speech, it is reasonable to investigate earlier texts to determine whether the same kinds of discourse functions can be identified, and to hypothesize about what role they might have had in the development of the constructions, even though our data are written, not spoken. Since WH- and ALL-clefts invoke an exclusive reading, they are semantically dialogic. One question is how this semantic dialogicity arose. The hypotheses that WH-clefts are synchronically related to questions (Den Dikken 2006), to turn-taking (Hopper 2001), and to counter-active stance (Kim 1995) make WH-clefts a particularly interesting site for investigating whether there is evidence that dialogual and/or dialogic contexts were crucial in the development of the construction. The observation that ALL-clefts are downward entailing also raises the question how this characteristic arose.

A search of the MED, OED, and Early Modern English texts shows that the specification ALL-cleft arose in the second half of the Early Modern English period, around 1600, and the WH-cleft by 1680.\textsuperscript{xvi} This is a period for which we fortunately have a considerable amount of represented dialogue, for example in drama and in trials. I therefore investigated a number of data bases, including drama as represented in the LION: Early English Drama data base for the two periods Jacobean and
Caroline (1603-1660), and Restoration (1660-1700), and trials as represented by the Old Bailey Proceedings Online from 1678 to 1743, and the Old Bailey Speech Set from 1732-1743.Ⅷ Trials up to August 1731 in the Proceedings are summaries reported in the third person, with some first person quotation. Those after that date are reported in first person. Trials in the Speech Set are lengthy first person transcripts from which the shorter Proceedings were drawn.

Miller (2006) says English is unusual in having IT-, WH- and reverse-WH clefts. He points out that not all languages in the European area have clefts of the IT-type, or this whole set of pseudo-clefts. Indeed, he claims that Finno-Ugric languages, Turkic, Russian, Bulgarian, Serbo-Croatian, and Polish have no clefts, and German has no IT-clefts, although most other Western European Indo-European languages have clefts. Old English had equivalents of “NP-focus IT-clefts” with _æt_ or zero but no IT- or pseudo-clefts (Ball 1994a), nor did it have any specificational constructions with the cleft structures of Present Day English. Ball (1994a, 1994b) discusses the development of specificational IT-clefts in Middle English (she calls them “informative-presuppositional”):

(12) The kniht bad speke and seith,
“Vilein, Thou schalt me telle, er that I go.
It is thi king which axeth so”.
“The knight commanded him to speak and says: “Villain, you shall tell me before I leave. It is your king who asks this”. (c.1393 Gower, Confessio Amantis 3.1244 [Ball 1994b: 186])

Visser (1963: 49-50) associates this Middle English development with the shift from relatively free word order to relatively fixed word order with syntactic subject and with changes in the pronominal system (more recent work regards this change as the reanalysis of clitic pronouns as NPs, see e.g., Kemenade 1987).

In what follows I sketch the histories of the ALL- and WH-constructions, and discuss the discourse contexts in which the developments appear to have occurred. As in other areas of change, constructional emergence can be construed as change either within a construction (intra-constructional change, e.g., developments internal to the Transitive construction, Trousdale Forthc a), or from one construction to another (inter-constructional change followed by intra-constructional change, e.g., the development of Partitives into Degree Modifiers, Traugott 2008, Forthc). A subtype of inter-constructional change is the emergence out of an extant construction or constructions of a new construction that did not exist before as a productive type, followed by intra-constructional change. This is the kind I will be investigating.

### 3.1. A brief history of ALL-specificational-clefts
Early examples in the data base with the string ALL – NP – V – BE – X are ascriptive (13a, b) or purposive (13c, d). In all cases _all_ can be understood to mean ‘everything’:

(13) a. I haue heard as much, and _all_ thou hast said is true. (1615 Bedwell, Mohammedes imposturae [LION: EEBO])
b. I haue made him happie by training you forth: In a word, _all_ I said was but a traine to draw you from your vow: Nay, there's no going backe.
‘I have made him happy by drawing you forth: in a word, all I said was only a trick to draw you from your vow. No, there is no going back’. (1606 Chapman, Monsieur D’Oliue [Ibid.])

c. Her nathelesse
Th’enchaunter finding fit for his intents,
Did thus reuest, and deckt with dew habiliments.
For all he did, was to deceiue good knights.
‘Nevertheless, finding her fit for his intentions, the enchanter dressed her again and decked her with appropriate clothes. For everything he did was in order to deceive good knights’. (1590 Spenser, Fairie Queene Bk. II [Ibid.])

d. I loue thee dearer then I doe my life,
And all I did, was to aduance thy state,
To sunne bright beames of shining happinesse.
(1601 Yarrington, Two Lamentable Tragedies [Ibid.])

There are also some strings that might be “reverse” ALL-clefts prior to 1600; here all means ‘only’:

(14) a. Candidus. Why crau'st thou then my Verse, & dost anothers bowndes inuade?
Siluanus. I reaue (‘deprive’) thée not thy Muse, …
But to thy Musicke for to lende an eare, is all I craue.
‘C. Then why do you crave my verse, and invade another’s territory?
S. I don’t deprive you of your Muse … but to your music to bend an ear is all I crave’. (1567 Baptista, Eglogs [LION: EEBO; this text is a translation])

b. But happinesse I had not as I thought,
… things begunne in ioy, were parting sad,
And yet that present ioy was all I had,
In recompence of all my trau'll and paine.
‘But I didn’t have happiness as I thought I would … things begun in joy turned sad at parting, and yet the joy of the moment was all I had in recompense for all my work and pain’. (1597 Lok, Ecclesiastes [Ibid.])

Note (14a) is a complex example with a topicalization paraphrasable as ‘All I crave is for to lende an ear to thy Musicke’. If examples in (14) are reverse clefts, they are structurally different from early “canonical” ALL-clefts in that the verbs used are different: crave and have, whereas those of the earliest clefts are say and do, as illustrated in (15) below.

Around 1600 we begin to find ALL NP V BE X strings with the pseudo-cleft meaning: all means ‘only’, not ‘everything’, and the focus is understood to be exhaustive and specification, i.e. dialogic. The NP is a personal pronoun (it was, however, not found in the data). Most occur with a verb of speaking, usually say (15a, b), some with do plus infinitive marker (15c):

(15) a. [The nymph Melliflora kisses Faunus]
Faunus thought oft Loues fire for to display,
Desire was bolde, but Shamefastnesse said nay.
If he began to come but somewhat neare her,
His body quak't as though his heart did feare her,
*All that he said was, Nymph when you are at leasure,*
*Faine would I speak,* he might haue spoke his pleasure …
(1600 Weever, *Faunus* [LION: EEBO])

b. [A “confutation” between a Jesuit (S.R.) and Bell]
Our slanderous and rayling Iesuite, reporteth my wordes in this manner; for saith *Bell* it is a thinge proper to God, to make something of nothing in al cases, and at al times. So then, *all that I said was this;* (viz) That though man can at sometime in some cases, make one thing of another; yet to make of nothing something, is proper to GOD alone, neither is man able to *performe the same.* (1608 Bell, *The Jesuits Antepast* [Ibid.]) (Note so then, and orientation to a conclusion contrary to that of the Jesuit; the referent of *all* is anaphoric, the focus repeats a prior statement with minor modification)

c. there is no possibilitie of overthrowing the new election which shalbe made when the place is voyd, and if it be so allready, or shalbe so, *all you can doe is to do some good for the tyme to come,* which if you can doe conveniently, and without much trouble, it wilbe woorth your labour… Further then this I see not to be done. (1624 Oliver Naylor to John Cousin [ICAME: CEECS])

*Say* often appears in the formula *all I can say,* as in (16). Whether in this formula or not, the verb of speaking in the ALL-clefts is followed directly by a finite clause (15a, 16a), or indirectly by the deictic *this* (15b), (16b) or the quotative *viz* (15b).

This means that virtually all ALL-clefts in the seventeenth century have a clausal focus.

(16) a. *Medloc.* [on bees] But as a bow continually bent, doth lose his strength: so Salomon wisheth that in havung found hony, we should but eate that is sufficient, lest other wise it fall out, we vomit it vp. ...

*Malcon.* And *all I can say is,* *Wisedome wil be iustified of her Children,* when *Follie will not depart from a foole, though he were brayed in a morter with a pestell.* (1608 Clapham, *Errour on the Left* [LION: EEBO])

b. What have you to say concerning the cause of the flowing and ebbing of the Sea?

*Answ.* To that, *all I can say is this,* that Aristotle himselfe for all his cunning was so perplexed in following that doubt, that he died for grieue because he could not understand it aright. (1635 Person, *Varieties* [Ibid.])

The clearest examples of the ALL-cleft construction include those used in negative contexts like (17a) (they block an ‘everything’ reading, and trigger an ‘only/nothing but’ or ‘below expectation’ reading), and clauses with *do + bare Verb* (17b):
a. But as for my self he doth me notorious wrong, I did not mention any Principles of Vnity in this place, nor so much as dream of them, … All I said was this, That we doe not separate from other Churches, but from their Accidentall Errours. (1658 Bramhall, Schisme Garded [LION: EEBO])

b. When any bow'd to me with Congees (= ceremonious bow) trim, All I could do, was stand and laugh at him. (1681 Baxter, Poetical fragments [Ibid.])

Examples like (17b) show that by the later part of the 17C the ALL-cleft construction had become conventionalized and morphosyntactically differentiated from the purposive string of the same abstract form. ALL-clefts with do do not have a purposive reading when V is modalized by can/could. By hypothesis, to in the originally purposeful constructions with do (13c, d) was reanalyzed as an infinitive marker in the context of modalized do. Having no semantic significance, to could be treated as optional (bare infinitives were, however, not the norm until the end of twentieth century).

Later developments of the ALL-cleft include use with gerundive complements after do (18), with it and full NP subjects, and with NP and PP focal elements.\(^{xxi}\)

(18) What I did I was driven to, as any one can see. It takes a real shock to make the average Familey wake up to the fact that the youngest daughter is not the Familey baby at seventeen. All I was doing was furnishing the shock. If things turned out badly, as they did, it was because I rather overdid the thing. (1917 Rinehart, Bab: A Sub-Deb [UVa])

Assuming that the origin of the ALL-pseudo-cleft was in ascriptive and purposive clauses like those in (13), there was:

(19) a. Semantic-pragmatic expansion in that:
   • the string ALL PRO V BE X is now polysemously used for semantically dialogic specificational focus constructions as well as for semantically monologic ascriptive and purposive constructions,
   • all is assigned downward inferential meaning,

b. Syntactic expansion, as illustrated by use of bare infinitives after do, and of gerundives,

c. Host class expansion, in that more Vs come to be used. At first, V is mainly a verb of saying or do; later we find statives, e.g., desire, know, mean, want. Other expansions include those of the subject to full NPs and of the focal element to NPs and PPs.

These changes are clear cases of grammaticalization in Himmelmann’s (2004) sense, albeit grammaticalization without lexical bleaching. Likewise, they are clear cases of constructionalization, since the new pseudo-cleft is understood holistically.

3.2. Discourse contexts for the development of specificational ALL- clefts

I turn now to evidence that can be gleaned from the data for discourse contexts for the development of ALL- clefts. The first question is what kinds of contexts would
have allowed for the semantic reanalysis of ascriptive and purposive constructions into specificational clefts. In the case of ALL-clefts, everything that one person says or does may not be enough for some other person or may be interpreted as mistaken/inadequate. This is because of the quantificational meaning of all. In (20) Henry V explicitly discusses this dilemma:

(20) More will I do; Though all that I can do is nothing worth, Since that my penitence comes after all, Imploring pardon. (1599 Shakespeare, Henry V IV.i.319 [UVa])

In (21) we see two perspectives interwoven dialogically: that of the person referred to, and that of the speaker:

(21) By all which your Honours may perceive, how he hath falsly traduced the Commissioners of the Navie, … and all he drives at, is by his unjust aspersions to bring the Parliament and them at ods, that so he might accomplish his own ends. (1646 mscb [ICAME: Lampeter])

From the perspective of the subject (he), everything he does is for the purpose of causing discord. From the perspective of SP/W, however, what he does is cause conflict (note the tense difference here), and is to be evaluated negatively (falsly and unjust are grounded in SP/W, not he, since he would not have characterized his own remarks as ‘false’ or ‘unjust’).

All the examples in (14) (reverse constructions) and (15) (canonical constructions) occur in dialogic and at least partially contesting contexts. If the subject is third person, SP/W typically evaluates an individual’s action, conveying counter-expectation (the ‘everything’ was not as much as could be expected, not good enough), as in (15a) and (22a). In the case of first person subjects, the speaker typically complains (everything they said or did was misunderstood, undervalued, or did not meet their own expectations), as in (15b) and (22b):

(22) a. The Trial was very tedious; a Cloud of Witnesses being called on either Side … But there being a Parson in Dod's Company, and he charged for being aiding and abetting, he discover'd where his Haunts were, so that the same Night he was taken. All that he endeavoured to prove, was, That there was no former Malice between them; so that upon my Lord Chief Baron Mountague's summing up the Evidence, the Jury gave in their Verdict, That he was guilty of Manslaughter. (12th July 1682, Trial of Robert Dod [BAILEY: s1682071216820712001])

b. There was no possibility of my leaving the Army to fetch her out of that Convent … I never could obtain Leave to be absent, but remain'd most part of the Winter there; all I could do was to order some Soldiers, that went for France, to call at Charleville, but I never heard from them since. (1697 Evremont, Female Falsehood (translation) [LION: EEBO])

Since dialogicity is in the context, it is not necessarily semanticized at this stage. To establish semanticization of the dialogicity we need to look for examples that occur
independently of dialogic contexts. These include use at turns. In the data ALL-clefts are rarely found at a turn, except in the formulaic expression *all I can say* in (16). (23) is a relatively late example from the Old Bailey trials of the ALL-cleft used by the speaker to initiate a turn:

(23) Sarah Clayton. I happened to go that Night to see what a-Clock it was, by Mrs Burges's Dial … I saw him go into the same House, and speak to a Woman, - … Alexander Watson. *All I know of the Matter, my Lord and Gentlemen of the Jury, is this. On Tuesday I saw this Moody walking about our Court, with his Hat on;* - I live in Windsor Court over against Mrs Burges's. - (Jan. 1740, Trial of Sarah Burges and Ann Hill [BAILEY: s17400116-84174001160001])

We can hypothesize that semanticization of dialogicity occurred in the early part of the eighteenth century. Here and in other uses at turns they mainly express unwillingness or inability to give information.

ALL-clefts occur in contexts what are “given” or “recoverable” in the sense that the prior context indicates that the subject was saying or doing something, or was attempting or expected to do so. However, in most cases the most important function appears not to be to fill an open proposition, but rather to highlight an upcoming statement as salient, and impose an exhaustive reading on it. This is particularly clear in the case of (15b). Here the focus, *That though man can at sometime in some cases, make one thing of another; yet to make of nothing something, is proper to GOD alone…*, is only a minor reformulation of what Bell a few lines quoted himself as saying: *as man can in some cases at some time make one thing of another; so in all cases, at all times, to make something of nothing, is proper to God alone.*

We may conclude that ALL-clefts arose as discourse moves of the type Bonelli (1992) identified for Present Day English, namely retrospection (referring back to given or recoverable contexts), although the salience of the focus may suggest prospection (or forward, cataphoric orientation) as well. Bonelli regards “change of posture” as a major characteristic of ALL-clefts. Understood strictly as switch-reference (which is what she intends), this does not appear to have been a significant factor in their development. However, understood more broadly as dialogicity, it is unquestionably crucial in their development.

With respect to attitudinal standpoint, from the beginning ALL-cLEFTS conveyed what Bonelli identifies as “negative attitude”. There are no examples of what she calls “positive attitude” (“evaluating a course of action as very simple”) or the related “damage limitation”. Interestingly, she lists “damage limitation” first under attitudes (Bonelli 1992: 23), but no clear examples of this appear in the historical data until the nineteenth century (an early example is (24)). They appear to be largely associated with the progressive, as in *I am saying is*, xxii and their absence in the Early Modern English data is therefore not surprising, given that progressive examples are rare at that period.

(24) “… Tell me, do you never relax from this very correct behaviour? " "I do not pretend people in general are without imperfections," Charlotte said stiffly. "*All I am saying is that goodness and foolishness are so often combined to such an extent that it is sometimes impossible to
separate them on a short acquaintance." (1817 Austen, Sanditon [UVa])

In sum, ALL-clefs arose in dialogic contexts. Only ALL-clefs in the formula all I can say is are dialogal in the sense of being associated with question and answer, or any other kind of turn-taking. The contesting, adversative meaning of the original context, though often still present, has essentially become semanticized into the construction, so that even out of context, all in ALL-con structs with the appropriate syntax are understood as meaning ‘only’, not ‘everything’. This has also occurred in expressions like albeit, although, all the same, after all.

### 3.3. A brief history of WH-specificational-clefs

Prior to around 1680 the only examples in data base with the structure WHAT – NP – V – BE – X are of the ascriptive and purposive types in (25):

(25)  

<p>| | | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>Weigh well my words, and perswade thy selfe, that what I haue said is true. (1631 Mabb, Spanish Bawd I [LION: English Prose Drama]) (ascriptive adjective)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>I am ty’d to nothing In this businesse, what I doe is meerely recreation, Not constraint. (1630 Middleton, Chaste Maid [ICAME: Helsinki cedud3a]) (non-referential nominal)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td>his hate to you unjustly, did not grow so fast, as my esteeme waranted by vertue, since what you did, was in defence, both of your Prince and Countrey. (1639 Carlell: Arviragvs and Philicia II [LION: English Prose Drama]) (reason adverb)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d.</td>
<td>Shal. Will you, upon good dowry, marry her? Slen. I will do a greater thing than that, upon your request, cousin, in any reason. Shal. Nay, conceive (‘understand’) me, conceive me, sweet coz. What I do is to pleasure (‘please’) you, coz. Can you love the maid? (?1597 Shakespeare, Merry Wives of Windsor I.i.250 [LION: Shakespeare]) (purposive)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(26) illustrates some possible reverse-clefs prior to 1680. Like the possible reverse ALL-clefs in (15), they have verbs that are atypical of the canonical clefts (which have mainly verbs of saying and do, as in (28) below):

(26)  

<p>| | | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>Some things I haue, which here I will not show; Some things I want, which you shall neuer know: … That, which to treat of, I now purpose (therefor,) Is what I neither haue, nor want (‘lack’), nor care for. (1621 Wither, Wither’s Motto [LION: EEBO])</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>There remaineth yet a third [objection], which may be answered … Who will forbid them to supply in such a case, that by a voluntary and arbitrary forme, which the Church could not provide for in a set forme? And this is what I intended to say of this argument. (1642 Mede, Diatribae [Ibid.])</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A few WH-Left-dislocations are also found in the data. Most have indefinite (non-given or inferrable) what(soever) (27a) but some are definite/referential (27b). (27b) follows several stanzas representing the speech and describing the actions that are the referents of what:

(27)  
a. Christ was conceiued for vs, … & **whatsoever** he did, he did it for our profite.
    “Christ was conceived for us … and whatever he did, he did it for our profit’ (1581 Baker, Lectures on Christian Faith [LION: EEBO])

b. For **what** he spake, for you he spake it, Dame,
    And **what** he did, he did himself to saue.
    (1596 Spenser, Fairie Queene, VI,2.13.5-6)

WH-Left-dislocations were rare in Early Modern English (and had been in decline from Middle English on, see Pérez-Guerra and Tizón-Couto Forthc). All the same, examples like (27b) may have served along with other specificalional constructions, such as specificalional IT-clefts, TH-clefts, and ALL-clefts as partial models for the new construction which emerged toward the end of the seventeenth century. As mentioned above, Kim (1995) sees a close relationship between WH-clefts and WH-left-dislocations in contemporary English.

Early examples of the WH-specificationlal cleft with a definite focus element include:

(28)  
a. Good Lord! Sir **Anthony**, you need not be so purty (**pretty** = ‘proper’); **what I say, is the Discourse of the whole City**, how lavishly you let him live, and give ill Examples to all young Heirs. (1685 Behn, City-heiress I.i [LION: English Prose Drama])

b. If it be objected that I preached to separate Congregations; my Answer is, That I preach’d only to some of many Thousands that cannot come into the Temples, many of which never heard a Sermon of many years. And **what I did, was only to preach to such as could not come to our Churches.** (1697 Baxter, Mr. Richard Baxter's Last Legacy [LION: EEBO])

In the case of WH-clefts, NP focal elements appeared early (e.g., (28a)), unlike in the case of ALL-clefts, but were rare. None of the historical data bases used show syntactic expansion to bare verbs after *do* within the Early Modern English period. This is apparently a twentieth century phenomenon. The first example I have found so far is (29a); the parenthetical (*Mr. Werner said*) and the light verb *let* are suggestive of a possible entry-point for the construction.

(29)  
a. Werner said yesterday that operations continued through the weekend. **What he did**, Mr. Werner said, **was let manual laborers go home Tuesday for some rest.** (1961 Keat, Baltimore Sun [ICAME-Brown])

b. **What an impro of this kind does is confront the actors with a situation where they have to act truthfully.** (1991 So You Want to Be an Actor [BNC])
Nevertheless, in general, the WH-cLEFTs followed the same line of structural development as the ALL-cLEFTs. Overall, Himmelmann’s criteria for grammaticalization are met. Semantic-pragmatic expansion is clear: the examples in (28) are specificational, not ascriptive or purposive as in (25). In so far as WH-cLEFTs eventually came to have a bare infinitive after do, they participated in syntactic expansion, and in so far as they came to be used with NP subjects and more Vs, they participated in the normal types of host-class expansion that grammaticalizing constructions typically undergo. As in the case of ALL-cLEFTs, they illustrate grammaticalization without lexical bleaching and the development of a micro-construction.

3.4. Discourse contexts for the development of WH-cLEFTs

Like ALL-cLEFTs, WH-cLEFTs arose in contexts where the initial part is given or retrievable. Like ALL-cLEFTs they also appear to be used primarily to highlight the focus as salient, and impose an exhaustive reading on it. But they differ somewhat from ALL-cLEFTs in three other significant respects:

- the extent to which dialogic contexts are involved,
- the preponderance of first person contexts,
- even sparser occurrence at turns.

With respect to the first difference, ascriptive and purposive precursors of ALL-cLEFTs have non-dialogic contexts, while contexts for reverse (14) and canonical (15) ALL-cLEFTs are dialogic. Over time, although ALL-cLEFTs have come to be used in non-dialogic contexts (the dialogicity of their context has been semanticized into the construction), for the most part they continue to occur in contexts of adversativity and refutation. In contrast, contexts for ascriptive, purposive, and reverse-cLEFT precursors of WH-cLEFTs are all dialogic. Once the WH-cLEFT arose, most contexts continued to be dialogic, for example in (28a) there is argumentative countering (Good Lord! … you need not be so purty; what I say, is the Discourse of the whole City), as there is in (28b) (If it be objected … my Answer is … And what I did was only to preach …). In (30) the context is argumentative reporting on the relationship among various chemicals:

(30) So that I see no reason or necessity, from this Phaenomenon, to assert the existence of Vitriol in the Sand of the Bath … If any shall affirm this Ochre to be Vitriol, I have not deny’d it, having formerly supposed it might be Terra Vitrioli, but what I here question is, whether any Vitrioline saline body, different from the Ochre, be contained in the Sand, or can lye undissolved there. (1676 ICAME: Lampeter: scib1676])

But there is no obvious dialogicity in the contexts of (31). (31a) is the beginning of the deposition by Paul Crispin, a shop-keeper and silver-smith who has accused the defendants of stealing a silver dish (the transcriber records Crispin’s lisp by substituting ‘th’ and ‘sh’ for ‘s’), and (31b) is the deposition of Mrs. Exton, wife of the owner of the dwelling from which linen was allegedly stolen:

(31) a. What I have to tha ith thith, I lotht a thilver Dith belonging to Brigadier Churchill, out of my Grate in Compton-Threet, the Corner of Greek-Threet, but I can’t tell when, becauth it ish impothible. (June 1733, Trial of Alexander Watson and William
Howard [BAILEY: s17330628-441173306280001])

b. I heard a Noise, and came down Stairs, but all the Things were
gone: I wash Linnen, and **what I lost was the Property of Mr. Gold.**
(May 1736, Trial of Christopher Freeman and Samuel Ellard
[BAILEY: s17360505-463173605050])

We may hypothesize that the long association with dialogic contexts, and perhaps the
model of ALL-clefts, allowed for semanticization of dialogicity very early.
WH-clefts are favored in the context of first persons (30, 31). In the third
person reports of the Proceedings of Old Bailey, there are no examples with third
person subjects. However, in first person reports, there are several examples,
including those in (31). These first-person examples occur at a turn (31a) or shortly
after (31b). The first example of a WH-cleft in the first person reports is (32), from
the testimony of one of the defendants, Paine, in response to a witness called French.
Like (31b) it occurs shortly after the beginning of a turn):

(32) *French.* I keep an Alehouse in Cross-Lane. P - has been at my House
in Company with the Prisoners.
*Pain.* I could have but a short Acquaintance with I - , for I am but just
come from Sea. **What I know of him is, that he is a Cooper by Trade,**
**and that when he was taken up, he deny'd that he knew any thing of the Watch.**
(Dec 1731, Trial of Samuel Cole and Edward Paine of St.
Sepulchres [BAILEY: s17311208-428173112080001])

Nevertheless, WH-clefts were found at turns in only two other trials in the data
(August 1740, Trial of Mary Ray, and July 1742, Trial of Stephen Price and John
Clark). Most examples prior to the trials do not occur at a turn (see (28b) and (30)).
The turn does not appear to be a prime context for the emergence of this construction.
Its primary use appears to be identify salient information as part of a narrative of
events (or in the case of (30) as part of the development of a research strategy.
It is particularly interesting in light of the fact that (16b) illustrates **all I can say**
used in a question-answer pair that there are no WH-clefts used in answer to a
question of the type *What did you say/do?* in the data, Most notably, in the Old Bailey
trials there are examples of *X asked what it was?*, but WH-clefts do not appear in
response. Instead we find answers such as those in (33):

(33) a. *Brown.* What did you say to Mr. Fisher, at Paddington, when he
bid you have a Care, or you'd hang yourself?
*Curtis.* I said, if I knew any Thing of you before, it was the Day
before Michaelmas, when you came to our Door with a Couple of
Geese, and I shew'd them to my Master.
(Dec. 1733, Trial of William Brown and Joseph Whitlock, of
Paddington [BAILEY: s17331205-44317331205000])

b. *Q.* What did you do next?
*O Bryan.* We bound him with a Penny Cord, and turn'd him into
the Ditch. (Jan. 1737, Trial of James Ryan, Garret Farrel, and Hugh
Macmahon [BAILEY: s17370114-469173701140001])

It appears that at the time of their origin, WH-clefts were used primarily to signal a
counter-active stance (see Kim 1995), or to make “the listener aware that what follows is part of a considered argument worthy of attention and not a casual comment” (Hopper 2001: 124). However, they are not originally privileged at conversational turns.

In sum, both types of clefts occur at their origins in dialogic contexts. Only ALL-clefts in the formula *all I can say is* are dialogal in the sense of being associated with question and answer, or any other kind of turn-taking. There is no evidence that WH-clefts were construed as related to questions.

4. From micro-construction to macro-construction

The emergence of specificational clefts can be considered an example of the intra-constructional emergence of a new cleft construction with a specific form-meaning pairing out of preexistent building blocks. First there was the development of the IT-cleft drawing on impersonals of the happen-class (Ball 1994a), then the development of the ALL-pseudo-cleft out of ascriptive and purposive, etc. constructions, and that of the WH-cleft out of similar material, micro-construction by micro-construction. It appears that at first ALL-clefts were used exclusively, and WH-clefts almost exclusively with clausal foci. Together they formed a meso-construction separate from IT-clefts by virtue of their clausal complements. IT-clefts in turn came under some circumstances to have clausal foci, and so finally all three came to form a cluster. Amanda Patten (p.c.) has pointed out that while (34a) is ungrammatical, according to her broad construal of the category of IT-clefts, (34b) is grammatical. An earlier example is (34c).

(34)  
   a. *It's that he’s an idiot that I don’t like him.
   b. No, *it's because he's an idiot that I don't like him.
   c. It was the faith and the persistence of Columbus that discovered America and opened the way for the millions who now call it their home. *It is because of these qualities that we honor him to-day; it is because this faith and persistence ended as they did in the discovery of a new world, that to-day his fame is immortal.* (1892 Brooks, *Christopher Columbus* [UVa])

With respect to clefts, we can think of micro- to meso- to macro-constructional changes of IT-, ALL- and WH-clefts as in Figure I:
While a macro-construction developed early, as soon as multiple clefts came into being, it was not configured in its current form until about 1900.

Patten (2007, In preparation) questions whether clefts as such should be considered to have a special status, and proposes instead that specification is the overarching concept to be attended to. Patten’s argument is that various syntactic strings can be used to express definite descriptions, and clefts are only one type of specificational construction. In clefts the relative clause is inherently restrictive and contributes to identification of the definite phrase following the copula. However, within the larger domain of specificational constructions, they are presumably only a sub-type. On the assumption that this view is correct, the macro-specificational cleft construction could be regarded as part of an even larger specificational construction. What its history was is a subject for further research.

Working with construction grammar from a historical perspective focuses attention on alignment and matching of micro-constructions with each other, resulting in their incorporation into meso-level constructions, and eventually reconfiguration of macro-constructions. Alignment and matching are analogical processes. The association of new structure with analogy may on first thought appear to be contrary to received practices in grammaticalization. As is well known, Meillet said:

“Tandis que l’analogie peut renouveler le détail des formes, mais laisse le plus souvent intact le plan d’ensemble du système existant, la “grammaticalisation” de certains mots crée des formes neuves, introduit des categories qui n’avaient pas d’expression linguistique, transforme l’ensemble du système”. ‘While analogy can renew details of forms, but usually leaves the structure of the existing system intact, ‘grammaticalization’ of certain words creates new forms, introduces categories that had no linguistic expression beforehand, transforms the system as a whole’. (Meillet 1958 [1912]: 133)

Meillet’s concern with changes that lead to systemic reconfigurations has resonated with generative interest in “catastrophic” changes in I-language, and parameter-
settings (see e.g., D. Lightfoot 1979 and passim). Nevertheless, work on grammaticalization in both parameter- and usage-based research has been largely concerned with initial changes, step-by-step, local realignments—small-scale reanalyses and analogical adjustments that may eventually lead to a major shift. In recent years there has been increasing convergence between formal and functional theories of syntax over grammaticalization (Fischer 2007 provides an overview of the different perspectives). Importantly, bringing construction grammar and grammaticalization together provides for partial convergence with Kiparsky’s (Forthc) proposal in terms of constraint-based Optimality Theory and UG that analogical changes are reanalyses at the local level (for fuller discussion see Traugott Forthc).

Cognitive systems are highly plastic—not only sensitive to frequencies in the linguistic environment as Bresnan et al’s OT analysis (This volume) assumes, but also open to experiment and fine-grained local realignments.xxv Each step is a small step and “locally abrupt”, whether within or across constructions (D. J. Lightfoot 2005). This does not, however, deny the correctness of the observation that over time divisions of labor occur, and that prototype, macro-constructional constructions lead to strong alignment effects, resulting in the marginalization of non-prototypical members (Denison 2001). However, since innovations are always occurring and only some of them become accepted by a speech community, and sometimes only for a short time, categories and especially the more complex meso- or macro-constructions can never be expected to be completely aligned. Emergent structures are essentially unstable in nature (Bybee and Hopper 2001).

5. Conclusion

In this paper I have shown that some basic assumptions about constructions based on constructed examples may be very specific either to the contemporary linguistic situation, or even to constructed examples. Specifically, I have shown that WH-pseudo-clefts, which are often said to have question-answer properties, do not appear to have them in the early data, at least in so far as they do not normally appear in question-answer turns. It would be worth investigating whether they actually have them in contemporary data of similar genre as were used for the present study. I have also shown that projecting back from conversational data to earlier (admittedly written) data can provide useful questions to explore, but what appears in contemporary conversations may not necessarily have motivated how a construction arose in the first place.

The morphosyntactic changes discussed in this paper suggest that distinguishing dialogic and dialogal (especially turn-taking) contexts is useful in attempting to assess the relative importance of different types of context in which new constructions emerge. This in no way denies the importance of other types of context including the sociolinguistic dynamics behind particular choices (L. Milroy 2007), or shifts in cultural values (Wierzbicka 2006). Ultimately these and many more issues would ideally contribute to hypotheses about how a particular change occurs, but integration of the various approaches cannot be achieved until the particular components of change are better understood.

The present paper has only scratched the surface of one subset of constructions. To achieve a more fine-grained account of the types of dialogic contexts in which the ALL- and WH-specification constructions arose, and an assessment of what SP/W attitudes they were used to convey, a quantitative collostructional analysis would be useful such as is modeled in work by Gries and
Stefanowitsch (2004) for synchronic analysis and by Hilpert (2007) for diachronic analysis, if it could be extended to assess inter- as well as intra-clausal distributional preferences.

---

**Notes**

i Many thanks to Ruth Kempson for comments and to Scott Schwenter and Graeme Trousdale for ongoing discussion of the issues. Ulrich Detges, Amanda Patten, and Richard Waltereit made several helpful comments on an earlier draft. Liz Coppock and Harry Tily helped me access some of the electronic data bases. To all my deepest appreciation.

ii Contrast the Chomskyan generative position that language change is grammar change (see Kiparsky 1968), with the concomitant assumption that a change is a change in the individual child’s grammar. The two points of view are discussed in Croft (2000: 42-63) and Hopper and Traugott (2003 [1993]: 43-50).

iii Not all interaction is cooperative, or seeks to achieve common ground; however, distinctly non-cooperative interactions are not evidenced in my data.

iv The terminology can be problematic, e.g., the term “dialogic” is sometimes used to cover both dialogal and dialogic interaction (Ford 1994, Taavitsainen, Härnäs, and Korhonen 2006).

v Much of this work ultimately goes back to Bakhtin, see Holquist (1981).

vi Various taxonomies of such moves have been proposed, among them Mann and Thompson (1992) where antithesis, concession, and contrast are distinguished, and Rudolph (1996), where adversativity and concessions are distinguished.

vii Rossari and Cojocariu (2007) provide diachronic evidence that questions such as French *la raison/la cause?* ‘the reason/the cause?’ (used either dialogally across turns or monologically within a turn) came to be used as routines for introducing explanations and elaborations, but these do not involve the development of complex clauses as in the case of conditionals proposed by Jespersen and Haiman.

viii For similar developments see Brinton (2001) on *Look!*, Lindström and Wide (2005) on Swedish *Hör du!* ‘Listen’ and other imperatives that have become discourse markers.

ix Although there is brief mention of the contesting dialogicity of the interaction in passages used in Detges (2006) to argue for turn-taking as the strategy leading to obligatory subjects in French, it appears to have been a far stronger motivation than is implied.

x Archer (2007) provides a detailed bibliography of Early Modern English courtroom trials as well as an assessment of their value for linguistic analysis. Huber (Forthc) discusses the *Proceedings of the Old Bailey* at the end of the Early Modern English and the beginning of the Modern English period.

xi In an early version of this paper, Lehmann referred to “grammaticalization without lexical bleaching”. In the current version he refers to manipulation of structures “that do not denote anything”.

xii The following approximate periods of the history of English will be referred to: Old English 650-1150, Middle English 1150-1500, Early Modern English 1500-1750, Modern English 1750-1970, Present Day English 1970-.

xiii Thanks to Graeme Trousdale for this example, drawn from a Google search.

xiv This is in contrast to the alleged position that Newmeyer (1998) and Janda (2001) attribute to several proponents of directionality in grammaticalization.
It is not clear why these are “positive” as opposed to “neutral”.

I searched each file for all/what I/you/he/she/it/we/they and for said/did was. The latter strings were selected because they were the most frequent verbs found in the earlier development of both cleft constructions, in order to find utterances which might have NP rather than pronominal subjects (none was found). It should be noted that Bonelli restricts the subject to I, you, we, one; also she does not use a test for specificationality, and many examples are cut off after the be-verb, so her analysis of ALL-clefts is not strictly comparable with the one presented here.

Many thanks to Merja Kytö for introducing me to this data base, and to Tim Hitchcock for permission to use it. A corpus for the historical sociolinguistic study of spoken in English in the late seventeenth to early nineteenth centuries, based on the Proceedings of these trials, is under preparation (Huber Forthc). Proceedings of The Old Bailey available to me date from 1678-1805, and those in the Speech Set from 1732-1834. Since I was investigating the period when WH-clefts arose, the earlier texts seemed most useful.

Many thanks to Mirjam Fried for suggesting investigation of the reverse clefts.

‘Viz’, short for videlicet ‘that is to say’, is a quotative grammaticalized in legal texts and restricted to writing according to Moore (2006).

Heine (2002) and Diewald (2002) call such unambiguous contexts “switch contexts” and “isolating contexts”, respectively.

When the use of it and full NP subjects and of NP and PP focal elements occurred, remains to be investigated; they appear to be largely nineteenth century developments.

In some present-day contexts it has political significance deriving from John Lennon’s All we are saying is give peace a chance. The discourse origins of this anti-Vietnam War slogan include covert contesting of the FBI’s attempt to deport Lennon on grounds of his being a potential terrorist: ‘No, we are not advocating violence, all we are saying is give peace a chance’ (U.S. v. John Lennon 2006).

Thanks to Carol Kaske for introducing me to this example.

Den Dikken (2006), working with a formal syntactic and semantic model, also seeks to connect IT- and WH-clefts to wider specification functions.

The fine-grainedness of these realignments suggests that it is impossible from a historical perspective to maintain the position Aarts (2004) has put forward that intra-category (“subsective”) changes are gradient whereas inter-category (“intersective”) changes are abrupt.

Sources


BNC  British National Corpus, http://www.natcorp.ox.ac.uk/

COBUILD  Collins Birmingham University International Language Database, http://www.collins.co.uk/Corpus/CorpusSearch.aspx

ICAME  ICAME Corpus Collection, http://nora.hudib.no/corpora.html

LION  Chadwyck Healey website, http://lion.chadwyck.com

References
Andersen, Henning. 2001. Actualization and the unidirectionality of change. In
(Current Issues in Linguistic Theory, 219.) Amsterdam/Philadelphia: Benjamins.
Andersen, Henning. 2006. Synchrony, diachrony, and evolution. In Ole Nedegaard
Thomsen, ed., Competing Models of Linguistic Change, 59-90. (Current Issues
in Linguistic Theory, 279.) Amsterdam/Philadelphia: Benjamins.
Archer, Dawn. 2006. (Re)initiating strategies: Judges and defendants in Early Modern
English courtrooms. Journal of Historical Pragmatics 7: 181-211.
Archer, Dawn. 2007. Developing a more detailed picture of the English courtroom
(1640-1760): Data and methodological issues facing historical pragmatics. In
Susan Fitzmaurice and Irma Taavitsainen, eds., Methods in Historical
Pragmatics, 185-217. (Topics in English Linguistics, 52.) Berlin/New York:
Mouton de Gruyter.
Ball, Catherine N. 1994a. The origins of the informative-presupposition it-cleft.
Ball, Catherine N. 1994b. Relative pronouns in it-clefts: The last seven centuries.
Bergs, Alexander. 2005. Social Networks and Historical Sociolinguistics: Studies in
Morphosyntactic Variation in the Paston Letters (1421-1503). (Topics in
English Linguistics, 51.) Berlin/New York: Mouton de Gruyter.
Bergs, Alexander and Gabriele Diewald, eds. Forthc. Constructions and Language
Grammaticalization - A Look from its Fringes and its Components. (Trends in
Linguistics: Studies and Monographs, 158.) Berlin/New York: Mouton de
Gruyter.
Bonelli, Elena Tognini. 1992. ‘All I’m saying is …’: The correlation of form and
Brinton, Laurel J. 1998. ‘The flowers are lovely; only, they have no scent’: The
evolution of a pragmatic marker. In Raimund Borgmeier and Andreas H.
of the German Association of University Teachers of English, 19.) Trier:
Wissenschaftlicher Verlag.
Brinton, Laurel J. 2001. From matrix clause to pragmatic markers: The history of look-
Change. (Research Surveys in Linguistics.) Cambridge: Cambridge University
Press.
Bybee, Joan and Paul J. Hopper, eds. 2001. Frequency and the Emergence of
Linguistic Structure. (Typological Studies in Language, 45.)
Amsterdam/Philadelphia: Benjamins.


Detges, Ulrich and Richard Waltereit. 2003. Turn-taking as a trigger for language change. Presentation at the panel on Diachronic Micropragmatics of the Romance Languages, Ninth International Pragmatics Conference (IPra), Toronto.


Fried, Mirjam. Forthc. Constructions and constructs: Mapping a diachronic process. In Bergs and Diewald, eds.


Hopper, Paul J. 2001. Grammatical constructions and their discourse origins: Prototype or family resemblance? In Martin Pütz, Susanne Niemeier, and René


29

101-159. (Cognitive Linguistics Research, 32.) Berlin/New York: Mouton de Gruyter.


Milroy, Lesley. 2007. Off the shelf or under the counter? On the social dynamics of sound changes. In Cain and Russom, eds., 149-171.


Rossari, Corinne and Corina Cojocariu. 2007. Constructions of the type la cause / la raison / le preuve + utterance; Grammaticalization, pragmatization or something else? MS.


Traugott, Elizabeth Closs. Forthc a. Grammaticalization, constructions and the incremental development of language: Suggestions from the development of degree modifiers in English. In Regine Eckardt and Gerhard Jaeger, eds., Variation, Selection, Development. Probing the Evolutionary Model of


