Beate Hampe and Stefan Th. Gries

Syntax from and for discourse II: More on complex sentences as meso-constructions

https://doi.org/10.1515/gcla-2018-0006

> ... linguistic phenomena ... can be well explained by a cognitive understanding of grammar, and we will show how well the data and facts revealed by corpus linguistic analyses can be accommodated in a cognitive linguistic model" (Schönefeld 1999: 149)

Abstract: This paper presents a direct continuation of preceding corpus-linguistic research on complex sentence constructions with temporal adverbial clauses in a cognitive and usage-based framework (Diessel 2008; Hampe 2015). Working towards a more systematic construction-based account of complex sentences with *before-*, *after-*, *until-* and *once*-clauses in spontaneously spoken English, Hampe (2015) hypothesised that the morpho-syntactic realisations of configurations with initial adverbial clauses systematically diverge from those of configurations with final ones as a reflection of the specific functionality of each and that usage properties that are found across instantiations with a coherent functional load are retained in the schematisations creating constructions. This paper employs a multinomial regression in order to test to which extent each of eight closely related complex-sentence constructions with either initial or final *before-*, *after-*, *until-* and *once*-clauses can be predicted from the realisation of a few key morpho-syntactic properties of the respective adverbial and matrix clauses involved. The results support an analysis of complex-sentence constructions as meso-constructions that are not only specific about the subordinator and the positioning of the adverbial clause, but also retain "traces" of characteristic usage properties.

Keywords: temporal adverbial clause, usage-based model, complex-sentence construction, meso-construction, corpus linguistics, multinomial regression

1 Introduction: Complex sentences in Cognitive Construction Grammar

The usage-based commitment is one of the hallmarks of Cognitive Construction Grammar (henceforth CCxG), holding that constructions arise from the patterns

Beate Hampe, Universität Erfurt, beate.hampe@uni-erfurt.de
Stefan Th. Gries, University of California at St. Barbara & Justus-Liebig-Universität Giessen, stgries@gmail.com

of usage that language users experience. Constructions are thus characterised in terms of properties that are true to the linguistic surface (e.g. Goldberg 2002; Gries and Stefanowitsch 2004; Bybee 2006, 2013; Diessel 2015), and also retain discourse-pragmatic information that remains invariant across instances. Usage-based theorising in CCxG thus encompasses earlier functional approaches, such as "emergent grammar" (e.g. Hopper 1987, 1998), or, more generally speaking, the functional tenet to explain clausal or sentential syntax with reference to the surrounding discourse (e.g. McCarthy 1998).

Given that much corpus evidence has confirmed the Zipfian distribution of the tokens instantiating constructions, usage-based research has come to place special emphasis on constructions that present mid- or even low-level generalisations. These are partially lexically substantial and referred to as "item-specific", "meso-" or, if sense-specific, even "mini-" constructions (e.g. Boas 2003: 215–284; Langacker 2000; Tomasello 2003: 117–122; Traugott 2008a, b; Diessel 2014).

Although the constructicon has from very early on been envisaged as a highly dynamic network of *tightly interrelated* constructions (e.g. Langacker 2000; Diessel 2015), the usage-based commitment has also fed the tenet of "no synonymy" (e.g. Goldberg 2002, 2006), which has inspired linguists – especially those employing corpus methods – to focus on the divergent surface properties of constructions more than on issues of relatedness (for discussion, see, e.g., Cappelle 2006; Uhrig 2015).

This paper seeks to contribute to a construction-based account of complex sentences that is surface-true – i.e. that views each complex-sentence configuration as a construction in its own right without neglecting vital issues of constructional relatedness – not least because the latter are essential to an understanding of these constructions. Apart from belonging to a larger line of research on complex-sentence constructions with adverbial clauses (surveyed in Section 2), the study presents a direct continuation of previous corpus-based work on complex-sentence constructions with temporal adverbial clauses containing the subordinators *after*, *before*, *once* and *until* in spontaneous spoken English, provided by Diessel (2008) and Hampe (2015). While the former intended to determine the relative strength of the major semantic/cognitive, processing and discourse-functional factors motivating the positioning of temporal adverbial clauses in complex-sentence constructions, the latter focussed more on the discourse functionality of highly marked complex-sentence configurations.

In order to achieve a characterisation of the semantic pole of each complex-sentence construction in terms of both semantic and discourse-pragmatic properties, Hampe (2015) suggested that the syntax of complex sentences with temporal adverbial clauses is most effectively analysed in terms of meso-constructions, i.e. as syntactic constructions that *minimally* retain a lexical anchor (viz. their respective subordinator) and that are explicit about the position of the respective

adverbial clause vis-à-vis its matrix (ex 1). Complex sentences exhibiting clauses with the same subordinator in different positions (ex 2) would instantiate different but related meso-constructions (ex 3).

(1) a. [S matrix [S adverbial *subordinator* .....] _____]]
b. [S matrix _____ [S adverbial *subordinator* _____]]

(2) a. people are prepared to wait ten years *before* they get a job. [BNC-HYS]
b. *Before* you before you all vote, put your hands down a minute. [BNC JJE]

(3) a. [S matrix [S adverbial *before* .....] _____]]
b. [S matrix _____ [S adverbial *before* _____]]

The underlining in (1) and (3) stands for the schematic parts of the construction (i.e. the clauses without their respective subordinator), with double underlining indicating that the semantic pole of each construction also includes the foregrounding incurred in coding one of the two events as the matrix (e.g. Langacker 2008, 1991: 436). Importantly, the profiled content of the matrix gains "interactive prominence", i.e. carries illocutionary force; while the adverbial clause remains "interactionally deactivated", i.e. without an independent illocution (cf. Verstraete 2004: 824, 839). Furthermore, the dotted underlining marks that, with respect to information management (where the general default is to present given information before new one) the information coded for by the initial adverbial clause is usually given by the preceding co-/context, thus contrasting with the new circumstantial information expressed by the final adverbial clause.

However, the schematisation depicted in (3) is still simplified – not only because it does not show the many 'smaller' constructions that are required for building the finite clauses involved in the configuration, but also because Hampe (2015) further hypothesised that each meso-construction might exhibit functional and formal properties of its own, in addition to those explicated in (3). More specifically, it was assumed that morpho-syntactic properties which directly – and coherently – reflect the special functionality of any given complex-sentence configuration would occur with increased frequency across its instantiations and thus also be retained in schematisations (e.g. Ellis and Ferreira-Junior 2009: 194), i.e. become a part of the respective meso-constructions.[1] Based on a comparison of initial and final *before*-clauses, meso-constructions with initial ones, for

1 Their re-analysis of these as entrenched cognitive routines (Schmid, *this issue*), lends adequate emphasis to the fact that schematization does not produce stable cognitive products but consists in dynamic patterns that are "emergent", i.e. remain "in flux", constantly recycling previous experience (cf. Hopper 1998: 157).

instance, were, hypothesised to contain the following usage information (cf. Hampe 2015: 317):

(4) [S matrix: nondeclarative/modalised [S *before* [NP Subj: 1st person] .....] _____]

In addition, meso-constructions with the same subordinator (ex 3) were claimed to be related to one another by being the more specific versions (i.e., 'allostructions') of a superordinate meso-construction (i.e., a 'constructeme', cf. Cappelle 2006). Such a constructeme would retain the subordinator and the foregrounding of the matrix but remain unspecified for the position of the adverbial clause and the information-management choice incurred by its choice (ex 5).[2] It would thus be relatively "discourse neutral", in principle defining the semantics of the respective subordinator itself (cf. Hampe 2015: 317).

(5) [S matrix [S adverbial *before* .....] _____ [S adverbial *before* _____]]

Following up on this work, the present paper analyses a part of the data set used in Hampe (2015) in a more comprehensive, i.e. multifactorial, way. In order to further substantiate the above claims about complex sentences with adverbial clauses as meso-constructions, it asks whether the functional specialities of any given complex-sentence configuration are associated with the specific realisations of a number of selected morpho-syntactic features of the two clauses involved to such an extent that the configuration itself can be predicted from them. More exploratively speaking, the question is to which extent (or how reliably) the tokens of each of the complex-sentence constructions investigated can be classified as tokens of the construction they instantiate on the basis of their realisations of few key morpho-syntactic features.

We regard this as a good/adequate test of the ideas advanced, both in this paper and elsewhere before, because it (very roughly) approximates a simulation of the task faced by language users/learners schematising over (i.e. classifying) experienced instantiations. In this sense, the predictiveness of the complex-sentence configurations from very few usage properties would lend support to claims about the cognitive reality of the meso-constructions depicted or, more generally, to an entrenchment-based account of constructions.[3]

2 The two occurrences of the *because*-clause indicate alternative positioning options and the information-management choices that go with it. For the purpose of this work, the much rarer clauses in mid-matrix position are entirely ignored. Hampe (2015) reported that less than 3% of the 1,554 tokens in the original dataset exhibited a clause in mid-matrix position.

3 Our sampling strategy (see Section 3) allows us to assume that, minimally, our data are compatible with what any one speaker's input would be.

2 Complex sentences with temporal adverbial clauses in CCxG

Adverbial clauses containing the subordinating conjunctions *before, after, once* or *until* provide an ideal test case for inquiries into the diverging functionality of complex sentences with temporal adverbial clauses, because they can only introduce clauses denoting events that either precede (*after, once*) or follow (*before, until*) the event expressed by the respective matrix. Depending on their position relative to the matrix in the resulting complex-sentence configurations (final vs. initial), they thereby create either iconic (ex 6, 8) or uniconic (ex 7, 9) representations of event sequences (cf. also Diessel 2008, testing a wider range of clauses and clause positions).[4]

(6) a. people are prepared to wait ten years *before* they get a job. [HYS]
b. they [projects] used to filter through the organization, *until* they landed on somebody's desk who was actually supposed to carry out the work. [H48]

(7) a. A little man turned up Friday afternoon, *after* they got home [KBW]
b. the men that I worked beside were quite prepared to honour an agreement *once* it was established. [GYV]

(8) a. So *after* you've finished training, you're going down to the library, ... [KGL]
b. *once* they've given us the go ahead then we can go to the solicitor in Lancaster [KB9]

(9) a. *Before* you came home the white cat came back! [KD8]
b. *Until* you've known loss you've never known what you really have got. [G4G]

According to Hampe (2015), the complex sentences illustrated in examples (6)–(9) instantiate the meso-constructions depicted in (10)–(13):

(10) a. [$_{S\ matrix}$ _____ [$_{S\ adverbial}$ *before* _____]] → iconic
b. [$_{S\ matrix}$ _____ [$_{S\ adverbial}$ *until* _____]] → iconic + telic

(11) a. [$_{S\ matrix}$ _____ [$_{S\ adverbial}$ *after* _____]] → uniconic
b. [$_{S\ matrix}$ _____ [$_{S\ adverbial}$ *once* _____]] → uniconic + telic

4 All examples are from the BNCII (world edition), file names are given in parentheses throughout the paper.

(12) a. [$_{S\ matrix}$ [$_{S\ adverbial}$ *after* .....] _____]] → iconic
b. [$_{S\ matrix}$ [$_{S\ adverbial}$ *once* .....] _____]] → iconic + telic

(13) a. [$_{S\ matrix}$ [$_{S\ adverbial}$ *before* .....] _____]] → uniconic
b. [$_{S\ matrix}$ [$_{S\ adverbial}$ *until* .....] _____]] → uniconic + telic

To survey the most relevant dimensions of these configurations again: Initial adverbial clauses with *before* and *until* (ex 13) are united in creating uniconic clause orders in initial position – contrasting with *after*- and *once*-clauses (ex 12). Initial adverbial clauses with *once* and *until* (ex 12b, 13b) are united in imposing telic/completive aspectual contours on the event depicted – contrasting with both *after*- and *before*-clauses (ex 12a, 13a). The two conjunctions paired in each block define event sequences in the same way but differ in that only the subordinators *once* and *until* impose a telic/completive aspectual contour on the event depicted by the adverbial clause. It will be discussed below (Section 2.1) how these variations lead to considerable differences in the functional potential (and resulting morphosyntactic realisations) of the resulting clause configurations.

Overall, Diessel's (2008: 478–483) multifactorial analysis, a binary logistic regression predicting the positioning of the adverbial clause (*initial* vs. *final*), showed that final adverbial clauses are generally preferred over (i.e. significantly more frequent than) initial ones, which is in line with both the amount of processing effort incurred (initial adverbial clauses create larger recognition domains) and information-management requirements (new information is expected at the end of a syntactic unit). He also found that semantically motivated clausal sequences (i.e. iconic clause orderings) are preferred over uniconic ones. Especially in the case of (rarer) initial adverbial clauses, iconic sequences are vastly more frequent than uniconic ones.[5] It could thus be said that the sentence configurations depicted in (10) and (13), respectively, present the extreme ends of a markedness scale, with those in (13) being least frequent due to their being dispreferred along *all* the dimensions investigated, i.e. processing effort, motivation and information management.

Much independent functional work on the diversity of discourse functions associated with adverbial clauses in different positions (cf. e.g. Diessel 2005; Ford 1993; Givón 1990; Lehmann 1988; Thompson 1985, 1987; Thompson and Longacre 1985; Verstraete 2004) has yielded an answer to the question of why even such (multiply) dispreferred configurations occur with any notable frequency in the first place. It has firmly established the special discourse-organising function of

5 Diessel (2008: 462, 472) found that this effect is stronger in *once*-clauses and attributes this to their conditional interpretation.

initial adverbial clauses, i.e. their capacity to act as framing devices that perspectivise the matrix event in line with discourse requirements, and also contrasted this function with the more strictly local scope of sentence-final clauses, which mostly provide new, circumstantial information directly related to the matrix event. Though the discourse-organising (re-)framing of matrix clause content by final adverbial clauses is not excluded, it is generally held to be more unusual than either final clauses with a local function or initial clauses with a framing function (cf. e.g. Verstraete 2004: 843–44). The claim that initial adverbial clauses have constructional status (e.g. Thompson 1985: 55; Verstraete 2004: 837–846) builds directly on these findings.

Hampe (2015) assumed that the specific ability to frame matrix information is what makes highly marked configurations the constructions of choice – i.e. the *unmarked* syntactic realisation – *in these specific functions*. She argued that an empirical/quantitative investigation of complex-sentence configurations needs to consider (the formal reflections of) the functionality of specific clause configurations because instantiations of dispreferred configurations frequently differ in usage from what might be called default realisations of complex sentences with adverbial clauses. More specifically, in such configurations, the adverbial is often related to the matrix clause "at the speech-act level" (Diessel 2008: 472–473). They thus exhibit a "pragmatic matrix", i.e. a matrix that refers to linguistic or non-linguistic aspects of the ongoing social interaction/speech situation itself, typically by carrying the illocutionary force of a request (including a request for information or permission). Formally, this is reflected by non-declarative (i.e. imperative or interrogative) clause types (ex 14a, b), or by declarative clauses expressing want-statements (ex 14c). This is also possible with, but not thought typical for, configurations with final adverbial clauses (ex 14d).

(14) a. Well, *before* you get carried away remember there's lots to do! [KD2]
b. So *before* we break to our groups okay can I have the old step one critiques that you haven't used back, thank you [K70]
c. *before* we pass to other business, I would like to thank you very much indeed for all the work [F86]
d. Are there any further matters arising *before* we pass onto the next item on the agenda? [D95]

Diverging from the methodology employed in Diessel (2008), Hampe (2015) therefore did not exclude instantiations with a pragmatic matrix from her data. Supporting the functionalist finding that initial adverbial clauses occur with a pragmatic matrix more frequently than final ones, this led to a significantly higher frequency of the most strongly dispreferred configurations with initial

before- and *until*-clauses (ex 13) in Hampe's data, which otherwise replicated Diessel's proportions of sentence configurations (cf. Hampe 2015: 318).[6]

2.1 Functionality and usage characteristics of sentences with initial adverbial clauses

The following section is meant to contextualise the multifactorial study reported in Sections 3 and 4, both with respect to preceding work on the specific functionality of the complex-sentence configurations investigated and concerning the syntactic features included as independent variables in the quantitative investigation. In particular, we demonstrate why we assume the (very few!) morpho-syntactic properties chosen to reflect the specific usage/functionality associated with a particular clause configuration. As most previous work focussed on the specific framing-functions of the four configurations with initial adverbial clauses (ex 12, 13), these will be surveyed in considerable detail, starting out from those that are the least dispreferred of these, namely those containing *once*- and *after*-clauses, creating iconic event sequences. Quantitative findings reported in either Diessel (2008) or Hampe (2015) that have inspired or do support our own considerations will be reported in the notes.

Initial *once*-clauses create iconic event orderings and impose a completive aspectual contour on the event depicted. This strongly motivates a shift to a conditional interpretation, which actually makes the occurrence of the adverbial clause in initial position *unmarked*, i.e. expected and frequent (cf. Diessel 2008: 470, 477).[7] More specifically, initial *once*-clauses code for events whose resulting state coincides with the *beginning* of the matrix event and are thus readily understood as defining an *enabling condition* (cf. Hampe 2015: 299). We therefore assumed that their matrix clauses should occur in the positive polarity (ex 15a–e) as well as modalised (ex 15a–c). Furthermore, both the temporal and the conditional interpretation of initial *once*-clauses cohere with the occurrence of the perfect in the adverbial clauses, explicitly marking anteriority (cf., e.g., Kortmann 1991) (ex 15b, c). They also cohere with *then*-insertion in the matrix

6 More precisely, the relative frequency of initial *before*-clauses (out of all *before*-clauses) increased from 0.07 (in Diessel's data) to nearly 0.2 (!), and that of initial *until*-clauses (out of all *until*-clauses) from 0.05 (in Diessel's data) to nearly 0.11.

7 In both Diessel's (2008) and Hampe's (2015) data, initial *once*-clauses were roughly four times more frequent than final ones.

(esp. in matrix-initial position), likewise serving to reinforce the sequentiality of the events expressed (ex 15b, c).

(15) a. but *once* they get one rope ashore they could do it, the job [H5H]
b. *Once* the change has been made, then many advantages, of course, would be found for hydrogen economy [KRF]
c. *once* they've given us the go ahead then we can go to the solicitor in Lancaster [KB9]
d. Now, *once* you've, *once* you've finished this penicillin stuff the hospital have given you, Mm. I want you to go onto this. [H57]
e. And *once* you pay it I'll put it back. [KDV]

The special kind of framing provided by conditionals does of course not preclude the occurrence of pragmatic matrix clauses. Examples (15d, e), for instance, realise a request in the form of a want-statement and a promise, respectively. However, this might turn out to be less typical for this configuration than for others with initial adverbial clauses – and perhaps even than for those with final *once*-clauses – because initial *once*-clauses are the only initial clauses in our data set that are actually *less* marked and more frequent than their final counterparts. Non-declarative matrixes marking a range of direct and conventionally indirect speech-acts are thus not assumed to be a specific hallmark of this configuration.

Though not quite as unmarked as the conditional use of initial *once*-clauses, configurations with initial *after*-clauses are comparatively inconspicuous as they create iconic event orderings without imposing a completive aspectual contour. As the shift to a conditional interpretation cannot be observed with *after*-clauses and as the previously observed shift to a causal interpretation is not known to motivate occurrence in initial position (cf. Diessel 2008: 472, 477), initial *after*-clauses are assumed to mostly provide a purely temporal framing by defining the matrix event as subsequent to the secondary one (ex 16a, b). As with initial *once*-clauses, this kind of framing invites matrix clauses in the positive polarity, the use of the perfect in the adverbial clause as a marker of anteriority (all tokens in ex 16 except 16d) as well as (matrix-initial or even multiple) *then*-insertion in the matrix additionally emphasising the event sequence expressed (ex 16b, e, f). As the adverbial clause does not act as a conditional, matrix clauses should not attract modalisation as a matter of course (though modals can occur as part of conventionally indirect speech-act routines, see below).

(16) a. and *after* I'd stamped and sealed the envelopes he rang back [KBK]
b. er *after* they'd revealed this in the Tet offensive of nineteen sixty eight, then President Johnson announced ...

c. *After* you'd had this erm punishment were you expected to go back to your lessons straight away, ... [FY5]
d. *After* you finish your first tape can we rewind it [KP0]
e. and then *after* you've done that then present it to the February sales meeting [JN6]
f. then *after* you've watched that, then you can go to bed after Thunderbirds [KD6]

If this simple kind of temporal framing is exploited for the purpose of managing ongoing social interaction, pragmatic matrix clauses will occur, realising, for instance, information questions (ex 16c), requests (ex 16d, e), or permissions (16f). Given the general considerations about the framing functions of (marked!) initial adverbial clauses, this should regularly happen (thereby also increasing the frequency of modalised matrix clauses). *Once*-clauses excepted, it should happen more frequently than in configurations with final adverbial clauses, though whether it should also happen more frequently than in other configurations with initial ones is unclear. No expectations about the typicality of non-declarative matrix clauses, defining a range of direct and conventionally indirect speech act formula, will thus be made at this point.

It is noteworthy, though, that one special construction with initial *after*-clauses that occurs with some frequency in our data fully exploits their capacity to provide purely temporal framings by defining the temporal distance between the two events coded in addition to the sequence of their occurrence (ex 17):

(17) a. a few weeks *after* the quarry men first came out on strike a few of us went up to the picket line, a few of the women. [HUX]
b. Two to four years *after* you left for instance the Pru or one or two other companies they can still claw back commission [J9Y]

Until-clauses impose endpoint focus on the events expressed but differ dramatically from *once*-clauses – not only because they create iconic clause orderings only in final position (ex 6b), but also because they code for a secondary event whose culmination point coincides with the *end* (not the beginning) of the matrix event. Hampe (2015) focuses on the precise functional affordances that this semantics creates if the *until*-clause occurs initially (ex 9b, 18) – considering that, in this position, it uses strictly given information in order to frame a matrix event by defining (*anticipating!*) a point in time at which that event ends before it is even expressed.

Though neutral uses are certainly not ruled out (ex 18a), the overall contrast to initial *once*-clauses expressing an enabling condition could not be sharper: Apart from inviting speakers to emphasise the temporary character of the matrix

event, an initial *until*-clause can be understood as expressing an event that, as long as incomplete itself, effectively *blocks* the completion of the matrix event, so that any change (or completion) of the matrix situation appears as impossible or difficult (ex 18b–e).[8]

(18) a. *Until* he was seven his life at home though poverty stricken had not been too bad. [J9A]
b. ... *until* I have been trained to do it perhaps P C should continue using his skills. [K6W]
c. *Until* that budget is actually er, settled, it's difficult to be definitive about the actual amount of money that's available [J3R]
d. ... *until* you can do something about this then you're not going to fight the crimes that occur [GY4]
e. *Until* the Conference Office say yea or nay, there's nothing I can do literally there. [FM2]
f. and *until* you come up with an alternative policy, don't you start knocking us for having a policy which we are prepared to debate. [JT8]

Pragmatically, this yields a framing potential that is relevant to the organisation of ongoing social interaction in highly specific ways. Not only does it allow the speaker to present an undesirable or problematic situation (that is in the foreground of the communicative interaction), as temporary (e.g. ex 18b–e), it also allows him/her to point to the event depicted by the *until*-clause as a potential reason for the unpleasant matrix situation (esp. ex 18c–e). This is especially useful if interlocutors need to express that they are unable or unwilling to act in a required way.

In view of the blockage-interpretation motivating this usage, we expect matrix clauses in this configuration to be characterised by negative polarity markers (ex 18a, d–f) as well as modalisation (ex 18b, e, f). As this functionality does not seem to attract matrix clauses with non-assertive illocutionary force, at least not in any obvious way – (ex 18f) is used rhetorically – there is no reason to assume that non-declarative matrix clauses should characterise this configuration (vis-á-vis all others). As *until*-clauses do not code for anterior events, perfect markers should not occur in the adverbial clause. Along the same lines, *then*-insertion in the matrix (though sometimes found, ex. 18d) does not help to make the temporal structure of the clause more transparent, hence is not expected to be frequent either.

8 Comparing configurations with initial vs. final *until*-clauses, Hampe (2015: 319–320) found significantly more negative-polarity markers in the matrix clauses of initial *until*-clauses.

Underlining the plausibility of these considerations, there is a special cleft-construction in our data, which reinforces both the telic nature of the adverbial clause and its interpretation as a blocking event. Since it clefts and negates the adverbial clause, the matrix codes for a desirable change and appears in the positive polarity (ex 19).

(19) a. And it's not *until* that economic climate changes that people are … able to to relate to what the politicians say. [K6A]
b. And it wasn't *until* I phoned the co-ordinators that he was able to tell me that [DCH]
c. It isn't *until* you make a conscious effort to listen, you realise how much noise you are subject to in this modern world [JK1]

Initial *before*-clauses, though likewise creating uniconic event orderings, are a bit more straightforward in that they do not impose an aspectual contour but merely frame the foregrounded matrix temporally by locating it prior to a secondary event, which the speaker considers as given and expresses first. Example (9a, repeated here as 20a) illustrates a neutral (narrative) use of this configuration.

Hampe (2015) demonstrated how this configuration is frequently exploited for organising ongoing social interaction – especially in more formal, semi-structured speech-situations such as meetings or lessons, creating configurations in which the clauses involved are clearly related at the speech-act level. More precisely, the two clauses in this configuration often refer to two events both of which involve speaker or addressee(s) and are about to occur in the immediate situational context. The later event of the two (coded by the initial adverbial clause) is given by the co-/context, while the earlier event (coded by the subsequent matrix) is literally (!) in the foreground of the interaction, depicting an event that the speaker wishes to happen next and judges to be unexpected to the addressee(s) (ex 20b–c):

(20) a. *Before* you came home the white cat came back! [KD8]
b. All right now *before* we do any more dates put your hands down … [JA8]
c. well, *before* you take off or turn left or turn right, or slow down or stop, use your mirrors, you should always look behind [KBM]

This takes on a specific metalinguistic/discourse-organising quality if either or both of the events expressed are communicative subevents of the ongoing speech situation itself (ex 21). That, in this case, the configuration is employed to manage multi-party interaction across turns is most obvious in tokens with interrogative or imperative matrix clauses, which pose a question or a request requiring an immediate response by the interlocutor (ex 21a, b), but also the case with declarative matrix clauses realising want-statements (ex 21c).

(21) a. *Before* we go onto that, erm how far do you think Jane Eyre supported this idea? [K60]
b. Chairman *before* we do that could I just make an observation on Yes. what er [the] Councillor said. [JA5]
c. Now *before* we move further I'd like to take this opportunity of expressing on your behalf our thanks to Hugh here. [F84]

Following from these functional considerations, complex-sentence configurations with initial *before*-clauses should be characterised by non-declarative and also modalised matrix clauses, given that the latter are a part of nearly all conventionally indirect speech acts, including want-statements.[9]

In our data, there is a special variation of tokens with pragmatic matrix clause highlighting the centrality of interactive uses to this configuration. There, the matrix appears reduced to a noun phrase headed by a communication noun or time noun (ex 22). Pragmatically, these "matrix-NPs" are the equivalents of non-declarative matrix clauses ("*can/could we have NP*"), or of want-statements ("*I/we would like/need to have NP*").

(22) a. Now, er, *before* I hand over to Frank er, a word about the dividend [HM7]
b. Mr Brighton, erm *before* I turn to another speaker, your comment about the location of a new settlement, ... [HVK]
c. *Before* you end, a few minutes, er, Mr President, ... [JNK]

2.2 Sentence configurations with final adverbial clauses

Configurations with final adverbial clauses (ex 10, 11) are all united in presenting the matrix event before the secondary one, which is assumed to mostly function locally, i.e. provide new circumstantial information directly related to the matrix. They differ semantically in accordance with the choice of the subordinator, creating either iconic (*before, until*) or uniconic (*after, once*) event sequences, and, as before, imposing telic interpretations of the adverbial clause only in the case of *once*- and *until*-clauses.

It has been mentioned above that the specific framing functionality of each of the four configurations with initial clauses, though not entirely ruled out

9 Hampe (2015: 319–320) found that markers of non-assertive illocutionary force, such as non-declarative matrix clauses but also deontic modals, were nearly twice as frequent in the matrix clauses of initial *before*-clauses than in the matrix of final ones.

for configurations with final ones, is generally assumed in the literature to be found with these only in exceptional cases, especially if the adverbial clause and its matrix are related at the speech-act level. That is not to say that (at least some) configurations with final adverbial clauses could not also develop special usage patterns with pragmatic matrixes, especially given that not all patterns with final adverbial clauses present the unmarked option (it was mentioned above that final *once*-clauses are marked), and given that not all patterns with initial adverbial clauses attract pragmatic (esp. non-declarative) matrix clauses more often than their final counterparts (it was reported above that initial *until*-clauses may not).

Two things should be noted, however. Firstly, the presence of a pragmatic matrix in sentences with final adverbial clauses does not by necessity entail that the final adverbial clause also carries out a specific, interactive or discourse-organising framing function exactly paralleling that of the corresponding configuration with an initial adverbial clause. Secondly, the fact that tokens with pragmatic matrix are unlikely to be typical for the usage of configurations with final adverbial clauses implies that morpho-syntactic properties reflecting these uses may not even be relevant to the prediction of the complex-sentence configurations instantiated.

To flesh out the first of these two points, we will briefly discuss the functionality of sentences with final adverbial clauses that show a pragmatic matrix. We will especially focus on tokens with a non-declarative, i.e. "openly pragmatic", matrix clause, though tokens with a declarative matrix clause realising a non-assertive illocutionary force, such as a promise (ex 23a, b) or a request in the form of a want-statement in (ex 23c), can also be found.

(23) a. I'll tell you my own view on this *after* you've voted [FL9]
b. I'll go through it again *once* we've seen the film [KS6]
c. And I need you to sign a consent form *before* you go as you have spoken at the meeting, okay. [FYB]

Irrespective of the specific configuration instantiated, non-declarative matrix clauses of final adverbial clauses in our data are frequently interrogatives. The resulting configurations mostly realise a request for information (rather than, indirectly, a request for action), with the adverbial clause overwhelmingly functioning locally (ex 24). The difference between the two uniconic configurations with final *after*- and *once*-clauses lies primarily in the fact that only the former can acquire a causal connotation (ex 24b).

(24) a. How long did he come in then *after* I phoned? [KBF]
b. What did P C do *after* the man was lying outstretched on the floor? [JNE]

c. Did you notice much differences in the work *once* you got promoted to a chargehand? [GYU]
d. What had she done *before* she came to you? [KGK]
e. Did you er wait *until* your kids were that wee bit older ...? [H01]

If both clauses refer to the ongoing social/communicative interaction, an information question posed by the matrix opens a sequence that organises the ongoing social interaction. From a more conversation-analytic point-of-view, this information question can be seen as a part of a pre-sequence, i.e. realises a speech act indirectly, functioning, for example, as a pre-offer/-invitation (ex 25a, c), or a pre-request (ex 25b, d).

(25) a. d'you fancy coming round for lunch *after* you drop Oliver off on Wednesday [KDE]
b. What goes in the end column *after* you've got decimal multiply or divide? [JJS]
c. Ann do you want a smoke *before* you go in? [KB8]
d. ... and will you keep on making that sound repeatedly *until* I clap like that and tell you to stop, okay [KP2]

Although it is correct to say that the matrix and adverbial clause are related at the speech-act level in these cases, it is likewise true that – if taken at the face-value of the information question itself – the adverbial clause can still be said to function locally.

This is not the case for the adverbial clauses in the very rare tokens with imperative or hortative matrix clause, i.e. tokens that realise a directive speech act directly (ex 26), which clearly (re-)frame the matrix in the manner of an afterthought (cf. Verstraete 2004) or of a pragmatic tag – after the speech act has already been carried out.

(26) a. and don't be so long *before* you come again [K66]
b. Well wait *until* he's gonna come one weekend [KCY]
c. Let's not have sniping from the Tory benches *once* we've taken this decision. [KGM].

To be more precise, in the two iconic configurations instantiated, the adverbial clauses either simply locate the period at which the illocutionary force holds before the secondary event (ex 26a) or specify a point in time from which the matrix illocution ceases to hold (ex 26b). In the uniconic configuration (ex 26c), the final clause defines a point in time at which the matrix illocution becomes relevant.

Though the general sparsity of tokens with imperative matrix clauses prevents us from fully exploring the pragmatic options that this opens for language users, we will at least take a closer look at the non-declarative uses of matrix

clauses with final *once*-clauses, which are marked, as well as at those with final *until*-clauses, which also might be pragmatically special, as the corresponding configuration with initial clauses is not expected to attract non-declarative matrix clauses (see Section 2.1).

To begin with the former, the matrix clauses of all tokens with final *once*-clauses in our data except the one given in (26c) are interrogative (ex 27). All of them realise questions seeking information or confirmation, none presents an indirect request.[10] The non-declarative uses of this pattern in our sample are thus more inconspicuous than expected, at least when compared to other tokens of configurations with final adverbial clauses that exhibit a pragmatic matrix.

(27) a. How long does it take for income support to come through? *Once* a claim has been made. [KE3]
b. They'll look nice won't they, *once* they're all filled up? [KDE]
c. So do we phone them *once* we get to the hotel or do they contact us? [KB3]

In contrast, the non-declarative matrix clauses of final *until*-clauses in our data are different in that five of the eight tokens are imperatives (ex 28a, b), three are interrogatives. Of these, only one realises the direct speech act of a request for information, the other two present indirect requests (ex 28c). The (relatively more marked) uses of configuration with final *until*-clauses thus tends more strongly towards truly interactive (i.e. directive) matrix clauses than the configuration with final *once*-clauses. Beyond that, the qualitative inspection of these directive uses suggests that they are strongly drawn towards negative polarity and mostly express that something should *not* be done in the (limited) time span till the completion of the secondary event.

(28) a. but don't move on *until* you feel the guy is either totally not committed to do something or sufficiently disturbed to do something. [K70]
b. Well don't start *until* grandma comes. [KBW]
c. why don't you wait *until* you move, ... [KD1]

However, non-assertive matrix clauses are generally dispreferred in configurations with final adverbial clauses (and also comparatively infrequent in our data). We are thus not able to predict any further *more specific* tendencies in the way users may actually exploit these configurations. Apart from that, we must state that, from the qualitative investigation of the remaining uses of the

10 Declaratives with question tags (posing requests for confirmation) were subsumed under interrogatives.

configurations with declarative matrix clauses, we are not able to discern any *more specific* usage tendencies either. This is not to say, of course, that we rule out their existence.

At any rate, we abstain from formulating any specific form-related hypotheses about the morpho-syntactic realisation of any of the complex-sentence constructions with final adverbial clauses, apart from those that are dictated by the semantics of their respective subordinator: Paralleling our expectations for the corresponding configurations with initial clauses, we expect perfect markers in the adverbial clause and/or *then*-insertion in the matrix only in the configurations with *after*- and *once*-clauses, where they are even less redundant (i.e. more needed to make the expression of the event sequence coded more transparent) than in the corresponding iconic configurations with initial clauses.

Otherwise, we hope that, even in the absence of a more detailed investigation of the functionality of configurations with final adverbial clauses, the decreased frequency (or near-absence) of those usage features that reflect the highly specific functionality of the sentence configurations with initial adverbial clauses alone suffices to distinguish not only between the allostructions belonging to the same constructeme, as shown in Hampe (2015) for configurations with initial *before*- and *until*-clauses, but between all eight configurations.

3 Corpus methods and quantitative analysis

This study continues to investigate a part of the data first explored in Hampe (2015). These were retrieved from a subset of the BNC II (world edition) totalling about 9.42 million words compiled from all files containing spontaneous spoken language – not only those collected in the "spoken demographic" files, but also files providing unscripted spoken language from more formal spoken genres, such as classroom or courtroom interactions, science demonstrations, spontaneous commentary and business meetings.[11] For the sample used in the 2015 study, Hampe had retrieved all instances of the four subordinators *after*, *before*, *once* and *until*, and chosen one random hit per file (i.e. the first according to the randomiser). False hits (such as prepositional uses of the subordinators),

11 This BNC subset contains a total of 812 files and was created with the help of Kilgariff's (1996) survey of the BNC, encompassing the genres labelled as follows: *Brdcast_discussn; classroom; consult; conv; demonstr; interview; interv_or_hist; meeting; parliament; pub_debate; tutorial; unclassified* (together 740 BNC files); *courtroom, speech unscripted, sportslive* (together 72 BNC files).

utterances that were (partially) unanalysable due to fragmentation or entirely without a matrix, adverbial clauses with ambiguous relations to preceding or subsequent clauses as well as tokens exhibiting special constructions (such as *it*-clefts) were removed from the data.

For this study, we used nearly identical samples, but limited the sample sizes to 70 tokens for each of the eight sentence configurations depicted in (10) to (13), following the same randomiser as in the first study. The sample size was determined by the frequency of the most infrequent configuration, i.e. that of initial *until*-clauses. The samples generally contain one hit per file (thus closely approximating 1 hit per speaker), which obviates the need for any random effects (intercepts or slopes by speaker(s)). In the coding procedure, any remaining coding errors overlooked in the first study were cleaned up. In addition, all tokens with adverbial clauses in mid-matrix position as well as all tokens exhibiting one of the special constructions discussed in Section 2 were also excluded from the data to prevent the additional constructions involved from bringing in effects that might distort the picture. The latter are again surveyed in (29). The resulting sample gaps were filled up with hitherto unused first tokens from new files, following the old randomiser.[12]

(29) i. configurations with initial *after*-clauses in which the adverbial clause appears in the complementation of a time noun (ex 17);[13]
ii. configurations with initial *until*-clauses in which the *until*-clause itself is clefted (ex 19);[14]
iii. configurations with initial *before*-clauses containing a matrix in the form of a noun phrase headed by a communication/interaction noun (ex 22).[15]

All of the resulting 560 tokens were then coded for the variables listed in (30) below. Crucially, variables (i) to (iii) do not code for usage characteristics but are invariant for all of the 70 tokens in each of the eight samples. The first two

12 We had to diverge from this 1-token-per-file strategy in the case of the rarest configuration (i.e. the one with initial *until*-clauses, which originally had no more than 70 tokens) when forced to bring in six new tokens (1.07 %) to replace six old ones removed in the second round of coding. Following the same randomizer as before, we took these six tokens from files we had already taken one token from.

13 These uses accounted for about 7 % of the data points (i.e. 5 of the first 75 true hits following the randomizer).

14 These uses accounted for about 11 % of the data points (i.e. 9 of the first 81 true hits), compared to only a single token clefting an initial *after*-clause: *It'll only be after I've m met a few people and photographed them well, that maybe my name will be start to be mentioned around* (HEN).

15 These uses account for about 8 % of our data points (i.e. 6 of the first 76 true hits).

together suffice to define the clausal configurations depicted in (10) to (13) above, the third specifies a property that is an inherent part of the semantics of two of the subordinators and absent in the other two. Variables (iv) to (vii) code for selected morpho-syntactic characteristics of the clauses involved. Although we are aware of the fact that, potentially, a much larger amount of relevant usage properties could characterise each of the eight complex-sentence configurations, the relatively small number of data points forced us to restrict the number and kind of variables to be included in our further analysis to those we were able to formulate hypotheses for. Despite precise ideas about its occurrence, *then*-insertion in the matrix, for instance, was not included in the analysis, as it can in principle occur in all configurations that would also encourage the use of the perfect as an anterior marker in the adverbial clause, but is less reliable than the latter, due to its tendency to be (over-)used, in the way of an insert.

(30) i. subordinator of adverbial clause: *after, before, once, until*
ii. position of adverbial clause relative to matrix: *initial, final*
iii. telicity imposed by subordinator: *telic, atelic* (stable for each of the four subordinators)
iv. clause type of matrix: *declarative* (comprising declarative and subordinate matrix clauses) vs. *non-declarative* (comprising interrogative and imperative clause types as well as clauses with question tags in interrogative format)
v. modal verb in matrix: *modal* (including all modal verbs except future-*will*) vs. *non-modal*
vi. negative polarity markers in matrix: *negated* (including tokens exhibiting negation related to either the verb or its complements, negating affixes belonging to lexemes heading phrases complementing the verb, strong negative-polarity adverbs or negated question tags) vs. *non-negated*
vii. perfect in adverbial clause: *perfect* (any instantiation of the perfect construction *HAVE + past participle of VERB*) vs. *no perfect*

To briefly comment on the operationalisations underlying some of the variables, variable (iv) is intended to capture pragmatic matrix clauses in a strictly form-oriented way, subsuming only non-declarative formats used for carrying out non-assertive speech acts, but not want-statements, to ensure conservative coding. However, want-statements are usually characterised by the use of modal verbs like *would*, which is captured by variable (v). In addition, modal verbs are relevant to configurations in which the secondary event can be viewed as either enabling or blocking the matrix event. Variable (vi) comprises not only negation with scope over the verb phrase, but a broad range of strong negative-polarity markers. We did not include the polarity of the adverbial clause as a predictor

variable in the final analysis because, rather surprisingly, there was not a single strong negative-polarity marker in any of the adverbial clauses in the entire sample of 560 tokens. All sampling and coding decisions are documented; the data are archived at *TROLLing* (The Tromsø Repository of Language and Linguistics) and can be accessed via <https://dataverse.no/dataverse/trolling>.

In her first exploration of a part of the data, Hampe (2015: 319–320) compared only sentences instantiating allostructions of the same constructeme, restricting her analysis to configurations with *until*- and *before*-clauses in initial vs. final position. Although these within-constructeme comparisons showed at least two of the configurations with initial clauses to differ from those with final ones, they were not able to take into account the combined effects of all variables in the entire data set, i.e. to check whether distinct sets of feature values are actually characteristic for (in the sense of 'predictive of') each of the complex-sentence configurations postulated as meso-constructions. Hampe (2015: 315) thus judged her results encouraging, but insufficient by themselves to support her own claims about these constructions.

We therefore employed a multifactorial analysis testing to which extent the realisations of the morpho-syntactic features that we assume to mirror distinct usage patterns are actually *different enough* to be able to distinguish between the eight complex-sentences configurations instantiated in our data. To this end, we submitted the 560 tokens in the data set to a *multinomial regression*, viewing variables (iv) to (vii) as independent/predictor variables, and the eight configurations formally defined by variables (i) and (ii) as the levels of the dependent variable (i.e. the complex-sentence configuration to be predicted). Given that the data used have been partially analysed before, the strong explorative character of this investigation should perhaps be emphasised by saying that the multifactorial analysis was performed in order to see whether the 560 tokens making up our final data set can be *classified* with sufficient accuracy into eight categories corresponding to the eight configurations postulated as meso-constructions above (ex 10 to 13).

4 Results and discussion

The multinomial regression predicting the eight configurations started out from an initial model with the following predictor variables:

(iv) clause type of matrix (CLAUSE_TYPE_MATR);
(v) modal verb in matrix (MOD_MATR);
(vii) negative-polarity markers in matrix (NEG_MATR);
(ix) perfect in adverbial clause (PERF_AVCL).

A forward stepwise model selection process (with a check for collinearity) led to a final model that is nearly the same as the initial one but has one more interaction, viz. that of variables (v) and (vii): modal verb in matrix with negative-polarity marker in matrix. This model is highly significant ($LR = 377.8265$, $df = 35$, $p < 0.0001$). The predictors' significance values are as follows:

(iv) clause type of matrix: $p = 0.023$
(ix) perfect marker in adverbial clause variable: $p < 0.0001$
(v):(vii) interaction of modal verb in matrix with and negative polarity in matrix: $p = 0.029$

R-squared with 0.49 is acceptable, but a higher overall classification accuracy would of course have been desirable. Note, however, that the one achieved (0.3179) is still significantly better than baseline ($p < 10^{-20}$). A higher accuracy was probably unfeasible given (i) that the dependent variable had eight levels, (ii) that there were four predictor variables, and (iii) that the sample size of 560 was relatively small in view of the classification task posed. Apart from this, important usage tendencies (and thus relevant predictor variables) may have been completely overlooked, especially given the lack of precise hypotheses about configurations with final clauses.

Figure 1 shows how well each of the eight configurations is classified in terms of precision, recall and accuracy. Note that, with the exception of final *before*-clauses, which are classified with the lowest accuracy (about 0.7), classification accuracy for all configurations is above 0.8. It is especially noteworthy that the most marked (and hence functionally most interesting) configurations, i.e. initial *before*- and *until*-clauses, are also those that get classified/identified best by far in terms of both precision and recall,[16] with precision scores around 0.4 and recall scores around 0.6.

Table 1 presents the combinations of independent variables that, if they were attested in the data, would lead to the highest predicted probabilities of each of the configurations. Saying that they correspond to the "prototypes" of the eight configurations means these combinations of characteristics lead to the model making its most confident predictions. (Just like humans would consider a schematic of even a non-existing bird prototypical of the category 'birds' if that schematic showed all or most of the most salient feature of birds, i.e. the gestalt of a bird body with feathered wings, a beak, etc.) Table 1 lists the configurations in the order of decreasing confidence of prediction/prototypicality, showing that

16 Precision is the percentage of cases of Y correctly classified as "Y" out of all cases classified as "Y"; while recall is the percentage of all cases of Y the system finds (out of all cases there are).

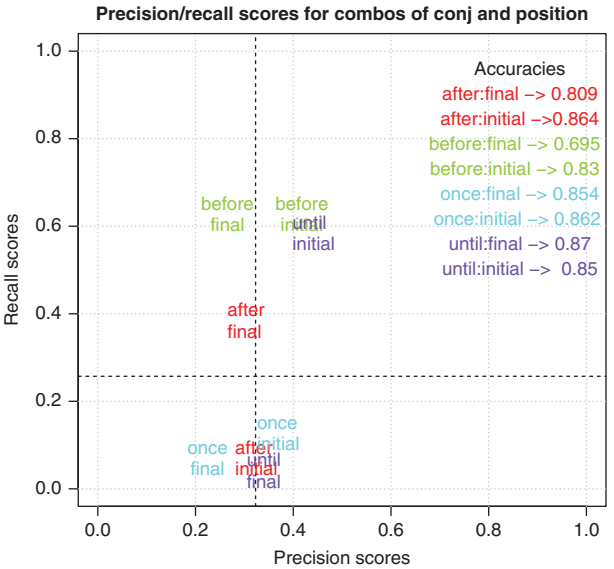

Figure 1: Precision, recall and accuracy of the classification into the eight complex-sentence configurations.

predicted probabilities tend to be higher for the more marked configurations – and highest again (above 0.5) for initial *before*- and *until*-clauses, which we also happened to have the most explicit functional hypotheses about. Note in this respect that configurations with final *once*-clauses, which were classified with a relatively high predicted probability as well (0.42), do not really present an exception to this general observation, as configurations with final *once*-clauses realise the marked positioning option for a clause with strong conditional interpretations.

Entirely in line with our functional considerations about their highly specific discourse-organising potential, configurations with initial *before*-clauses are classified/identified quite reliably on the basis of their tendency to exhibit a non-declarative (i.e. pragmatic) matrix of positive polarity and to not mark anteriority by a perfect in the adverbial clause (which non-iconically codes for a subsequent event). The tendency towards modal verbs in matrix clauses is motivated by the fact that all conventionally indirect speech acts, including most want-statements in declarative form, are modalised.

The same applies to configurations with initial *until*-clauses. Supporting an interpretation of these clauses as specific framings which bring in an event effectively blocking the matrix event from completion, they are characterised by modalised declarative matrix clauses of negative polarity. The perfect tends to

Table 1: Predicted probabilities for the eight complex-sentence configurations investigated. Predicted probabilities < 0.3 indicate very weak prototypes.[17]

Complex-sentence configuration	Predicted probability	Scenarios (combinations of independent variable levels) leading to the highest predicted probability
before:initial (ex 13a)	0.6443154	CLAUSE_TYPE_MATR: interrog/imp & PERF_AVCL: absent & MOD_MATR: present & NEG_MATR: positive
until:initial (ex 13b)	0.5564182	CLAUSE_TYPE_MATR: decl/subord & PERF_AVCL: absent & MOD_MATR: present & NEG_MATR: negative
once:final (ex 11b)	0.4174666	CLAUSE_TYPE_MATR: interrog/imp & PERF_AVCL: present & MOD_MATR: absent & NEG_MATR: negative
after:initial (ex 12a)	0.3523375	CLAUSE_TYPE_MATR: interrog/imp & PERF_AVCL: present & MOD_MATR: present & NEG_MATR: positive
once:initial (ex 12b)	0.3293354	CLAUSE_TYPE_MATR: decl/subord & PERF_AVCL: present & MOD_MATR: present & NEG_MATR: positive
after:final (ex 11a)	0.3275697	CLAUSE_TYPE_MATR: interrog/imp & PERF_AVCL: present & MOD_MATR: present & NEG_MATR: negative
until:final (ex 10b)	0.2874066	CLAUSE_TYPE_MATR: decl/subord & PERF_AVCL: absent & MOD_MATR: absent & NEG_MATR: negative
before:final (ex 10a)	0.2511407	CLAUSE_TYPE_MATR: decl/subord & PERF_AVCL: absent & MOD_MATR: present & NEG_MATR: positive

be absent from the adverbial clause again, which likewise does not code for an anterior event. What the multifactorial analysis has added to previous findings about configurations containing initial adverbial clause with *before* and *until* is thus that the observations initially made on the basis of two isolated comparisons do in fact characterise these two clause-configurations vis-à-vis all others in the data set.

Despite lower predicted probabilities, nearly the same also applies to configurations with initial *after*- and *once*-clauses. The interpretation of the latter as an enabling condition made us expect modalised matrix clauses of positive

17 The value of 0.3 results from the fact that the scenarios of the final two configurations are those that lead to the highest predicted probabilities for those complex sentence configurations. However, these scenarios do not also actually predict the listed configurations. For instance, the highest predicted probability of *before*:final results from the last scenario in Table 1 (which is why that scenario is listed there), but that scenario's predicted configuration is in fact the same as for the first scenario, i.e. *before*:initial. Put differently, if the configuration is *before*:final, then the scenario that was most likely is the one listed in the last row of Table 1, but if the scenario is the one listed in the last row of Table 1, then the analysis actually considers *before*:initial more likely.

polarity, but – in view of the unmarked nature of initial conditionals – not also non-declarative matrix clauses, both of which turned out to be largely correct. With respect to the purely temporal framing provided by the former (i.e. initial *after*-clauses), we were not sure about the typicality of non-declarative matrix clauses, but the finding that these are actually characteristic for this configuration makes sense given that initial *after*-clauses are more marked than both initial *once*-clauses and all final adverbial clauses except *once*-clauses. In accordance with the iconic clause orders to be found with both configurations, both configurations are also (and as expected) characterised by the presence of the perfect in the adverbial clause.

Of the four configurations with final adverbial clauses, those with *once*- and *after*-clauses are classified better, and – surprisingly! – configurations with *once*-clauses even better than configurations with initial *after*-clauses (0.42). With the possible exception of final *once*-clauses, however, we did not expect non-declarative matrix clauses to be typical of any of the configurations with final adverbial clause, although their increased occurrence with final *once*- and *after*-clauses is generally in line with the relatively higher markedness of the uniconic clause ordering in both configurations and, in the case of configurations with final *once*-clauses, even additionally motivated by the fact that these are more marked than initial ones. As assumed, the uniconic clause ordering in both configurations also motivates the use of the perfect in the respective adverbial clauses, coding for anteriority. The precise feature levels of the two remaining variables, however, are unexpected and deserve further discussion.

For configurations with final *once*-clauses, they predict modalisation as absent from and negation as present with matrix clauses, thus determining the tokens in (ex 31) as instantiations of its predicted prototype.[18] Note that the positive polarity of the matrix is in line with the interpretation of the adverbial as expressing an enabling condition and that interrogative matrix clauses which truly ask questions are also in line with the more local function of final adverbial clauses generally assumed.

(31) a. How long does it take for income support to come through? *Once* a claim has been made. [KE3]
b. Now er what happened *once* you'd been captured then? [FYJ]

18 Tokens with non-modalised and non-negated imperative clauses would also do that (due to the fact that the variable clause type only had two feature levels), but the only token with an imperative matrix in our data is negated (ex 27.c)

Configurations with final *after*-clauses have not been specifically discussed beforehand at all, not least because we did not expect them to be characterised by non-assertive matrix clauses. Their predicted prototype should show non-declarative, modalised matrix clauses of negative polarity, but there is *not a single token* instantiating this feature-level combination in our sample (even though occurrences of this kind may well be attested in other/larger samples). It is thus impossible even to explore the functionality of this configuration post-hoc.[19]

With a view to the weakness of (some of) the predicted prototypes, it should finally be added with hindsight (and on a slightly speculative note) that any of the following might have contributed to making potentially characteristic usage patterns appear less pronounced in the analysis than they could have been, or perhaps even entirely invisible: (i) the near-absence of specific hypotheses about the (mostly unmarked) configurations with final adverbial clauses, (ii) the necessity to keep the overall amount of feature levels to a minimum (interrogative and imperative clauses, for instance, were subsumed under "non-declarative") as well as (iii) the strictly form-oriented way in which the data were coded (excluding, for instance, want-statements or promises from the group of non-assertive matrix uses).

5 Final considerations

Despite these critical remarks, and given that this study presents the first attempt to go "backwards" from the selected usage features of observed tokens to the categories (i.e. the syntactic constructions) that these tokens instantiate, we regard it as a satisfactory result that the multinomial regression has managed to classify a decent (i.e. above chance) number of tokens in line with expectations – on the basis of a *very small* amount of morpho-syntactic usage characteristics. In other words, the number of pertinent/coherent clues that language users can draw on when schematizing over usage is probably much larger, such that our simulation of their classification task must remain a rough approximation. This notwithstanding, we view the assumption of meso-constructions as supported by our results *in principle*.

Moreover, we are also confident that the feature values of the predictor variables provide a significant part of the usage information that the meso-constructions depicted in a simplified manner above (ex 10–13) should be

19 Scenarios with very weak predicted probabilities (i.e. < 0.3) are not discussed any further.

specified for, at least for those configurations predicted with a probability larger than 0.3 and also instantiated in our data. (32) presents the elaborated versions of the meso-constructions for those configuration:

(32) a. [S matrix: non-declarative, modalised, [S adverbial *before* .....] _____]]
b. [S matrix: declarative, modalised, negative polarity [S adverbial *until* .....] _____]]
c. [S matrix: non-declarative, negative polarity _____[S adverbial: perfect *once* _____]]
d. [S matrix: non-declarative, modalized, _____[S adverbial: perfect *after* _____]]
e. [S matrix: declarative, modalised [S adverbial: perfect *once* .....] _____]]

Concerning the claim that *all* of the eight sentence configurations have the status of meso-constructions, two different conclusions can be drawn from our results at this point: It is not implausible, on the one hand, that configurations are the less prone to developing specific usage patterns the less unmarked they are. Echoing the results of much functional research, this would imply that, out of all possible usage configurations, only some acquire special status as meso-constructions. Cognitively speaking, this would amount to saying that users will only remember those formal hallmarks (or only entrench those cognitive routines) that are associated with the most salient special uses that the respective configurations lend themselves to.

On the other hand (and for the above-mentioned reasons), we cannot exclude a partial failure on our side to capture those morpho-syntactic features that would have been most directly relevant to the description of the functional specialties of (some of) the configurations investigated – and hence provided the best predictor variables for these. Though the results of the regression analysis have also produced at least one meaningful suggestion (viz. for final *once*-clauses) as to what this unexplored potential might look like, we leave it to future research to determine – on the basis of larger samples and perhaps an adequate number of improved predictor variables – which of the two conclusions is the one to go for.

References

Boas, Hans C. 2003. *A constructional approach to resultatives*. Stanford University: CSLI Publications.

Bybee, Joan. 2006. From usage to grammar: The mind's response to repetition. *Language* 82(4). 711–733.

Bybee, Joan. 2013. Usage-based theories and exemplar-based representations of constructions. In Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford handbook of Construction Grammar*, 49–69. Oxford: Oxford University Press.

Cappelle, Bert. 2006. Particle placement and the case for 'allostructions'. Special issue of *Constructions* 1–7 (*Constructions all over: Case Studies and Theoretical Implications*), < www.constructions-journal.com >.
Diessel, Holger. 2005. Competing motivations for the ordering of main and adverbial clauses. *Linguistics* 43(3). 449–470.
Diessel, Holger. 2008. The iconicity of sequence: A multifactorial analysis of clause order in complex sentences. In Arne Zeschel (ed.), *Usage-based approaches to language processing and representation*. Special issue of *Cognitive Linguistics* 19(3). 465–490.
Diessel, Holger. 2014. Frequency and lexical specificity in grammar. In Heike Behrens & Stefan Pfänder (eds.), *Experience counts: Frequency effects in language*, 209–238. Berlin & Boston: de Gruyter Mouton.
Diessel, Holger. 2015. Usage-based Construction Grammar. In Ewa Dąbrowska & Dagmar S. Divjak (eds.), *Handbook of Cognitive Linguistics*, 296–321. Berlin & Boston: de Gruyter Mouton.
Ellis, Nick & Fernando Ferreira-Junior. 2009. Constructions and their acquisitions: Islands and the distinctiveness of their occupancy. *Annual Review of Cognitive Linguistics* 7. 187–220.
Ford, Cecilia E. 1993. *Grammar in interaction: Adverbial clauses in American English conversations*. Cambridge: Cambridge University Press.
Givón, Talmy. 1990. *Syntax: A functional-typological introduction*, Volume II. Amsterdam & Philadelphia: John Benjamins.
Goldberg, Adele E. 1995. *Constructions: A Construction Grammar approach to argument structure.* Chicago: The University of Chicago Press.
Goldberg, Adele E. 2002. Surface generalizations: An alternative to alternations. *Cognitive Linguistics* 13(4). 327–356.
Goldberg, Adele E. 2006. *Constructions at work: The nature of generalizations in language*. Oxford: Oxford University Press.
Gries, Stefan Th. & Anatol Stefanowitsch. 2004. Extending collostructional analysis: A corpus-based perspective on 'alternations'. *International Journal of Corpus Linguistics* 9(1). 97–129.
Hampe, Beate. 2015. Syntax from and for discourse: Adverbial clauses as item-specific construction in spontaneous spoken English. In Thomas Herbst & Susen Faulhaber (eds.), *Constructions: Aspects of Construction Grammar*. Special issue of *Zeitschrift für Anglistik und Amerikanistik* 63(3). 295–322.
Hopper, Paul. 1987. Emergent grammar. *Berkeley Linguistic Society* 13. 139–157.
Hopper, Paul. 1998. Emergent grammar. In Michael Tomasello (ed.), *The new psychology of language: Cognitive and functional approaches to language structure*, 155–176. Mahwah, NJ: Erlbaum.
Kilgarriff, Adam. 1996. BNC database and word frequency lists. URL: < http://www.kilgarriff.co.uk/bnc-readme.html >.
Kortmann, Bernd. 1991. The triad "tense-aspect-aktionsart": Problems and possible solutions. *Belgian Journal of Linguistics* 6. 9–30.
Langacker, Ronald W. 1991. *Foundations of Cognitive Grammar, Volume II: Descriptive Applications*. Stanford, CA: Stanford University Press.
Langacker, Ronald W. 2000. *A dynamic usage-based model.* In Michael Barlow & Suzanne Kemmer (eds.), *Usage-based models of language*, 24–63. Stanford: CSLI Publications.
Langacker, Ronald W. 2008. Subordination in Cognitive Grammar. In Barbara Lewandowska-Tomaszczyk (ed.), *Asymmetric events,* 137–149. Amsterdam & Philadelphia: John Benjamins.

Lehmann, Christian. 1988. Towards a typology of clause linkage. In John Haiman & Sandra A. Thompson (eds.), *Clause combining in grammar and discourse*, 181–225. Amsterdam & Philadelphia: John Benjamins.

Lehmann, Christian. 2007. Motivation in language: Attempt at a systematization. In: Peter Gallmann, Christian Lehmann & Rosemarie Lühr (eds.), *Sprachliche Motivation: Zur Interdependenz von Inhalt und Ausdruck*, 100–135. Tübingen: Narr.

McCarthy, Michael. 1998. *Spoken language and Applied Linguistics*. Cambridge: Cambridge University Press.

Schönefeld, Doris. 1999. Corpus Linguistics and cognitivism. *International Journal of Corpus Linguistics* 4(1). 137–171.

Thompson, Sandra A. 1985. Grammar and written discourse: Initial vs. final purpose clauses in English. In Talmy Givón (ed.), *Quantified studies in discourse*. Special issue of *Text* 5(1–2). 55–84.

Thompson, Sandra A. 1987. 'Subordination' and narrative event structure. In Russell Tomlin (ed.), *Coherence and grounding in discourse*, 435–454. Amsterdam & Philadelphia: John Benjamins.

Thompson, Sandra A. & Robert E. Longacre. 1985. Adverbial clauses. In Timothy Shopen (ed.), *Language typology and syntactic description*, *Volume II*, 171–234. Cambridge: Cambridge University Press.

Tomasello, Michael. 2003. *Constructing a language: A usage-based theory of language acquisition.* Cambridge, MA, & London: Harvard University Press.

Traugott, Elizabeth Closs. 2008a. The grammaticalization of NP of NP patterns. In Alexander Bergs & Gabriele Diewald (eds.), *Constructions and language change*, 23–45. Berlin & New York: Mouton de Gruyter.

Traugott, Elizabeth Closs. 2008b. 'All that he endeavoured to prove was …': On the emergence of grammatical constructions in dialogic contexts. In Robin Cooper & Ruth Kempson (eds.), *Language in flux: Dialogue coordination, language variation, change and evolution*, 143–177. London: Kings College Publications.

Uhrig, Peter. 2015. Why the Principle of No Synonymy is overrated. In Thomas Herbst & Susen Faulhaber (eds.), *Constructions: Aspects of Construction Grammar*. Special issue of *Zeitschrift für Anglistik und Amerikanistik* 63(3). 323–337.

Verstraete, Jean-Christophe. 2004. Initial and final position for adverbial clauses in English: The constructional basis of the discursive and syntactic differences. *Linguistics* 42(4). 819–853.