Topics in Linguistics (2025), 26(1), pp. 20–40
10.17846/topling-2025-0002

Fear, *dread* and *horror aequi*: What motivates the non-finite complementation of apprehension verbs?

Joanna Podhorodecka*
University of the National Education Commission, Poland

Key words

non-finite complementation
apprehension verbs
usage-based Construction Grammar
logistic regression
COCA corpus

Abstract

Apprehension predicates, such as the verbs *fear* and *dread* or the adjectives *afraid* and *terrified*, can combine with a number of complementation patterns, including to-infinitives, bare and prepositional gerunds, as well as finite clauses. In many cases there is no immediately apparent difference in function between the alternative constructions. However, they are characterized by their own sets of collocations, syntactic idiosyncrasies and specific pragmatic functions, which suggests that the complement choice is conditioned by a number of probabilistic factors. The aim of this study is to investigate selected variables affecting the choice of the non-finite complements of the verbs *fear* and *dread*, namely the to-infinitive and the gerund. Logistic regression is applied to data from the COCA corpus in order to establish the relative importance and statistical significance of a range of factors, from strictly syntactic constraints, including the *horror aequi* principle and extraction, through semantic aspects such as agentivity, to pragmatics, specifically the contextual inferences associated with the constructions. The study investigates if the choice of non-finite complement constructions with both *fear* and *dread* is affected by the same set of factors, or whether there is a specific type of motivation – semantic, syntactic, lexical, or pragmatic – that prevails in the case of each of the two analysed verbs.

1. Introduction

The aim of this study is to examine the factors which account for the choice of the complement constructions of the apprehension verbs *fear* and *dread* in the COCA corpus. Both verbs allow a number of complementation patterns, including finite clauses with or without *that*, but the focus of this study is the contrast between the two non-finite same-subject constructions, the to-infinitive and the gerund, exemplified below:

* Corresponding author

Cite this paper: Podhorodecka, J. (2025). *Fear*, *dread* and *horror aequi*: What motivates the non-finite complementation of apprehension verbs?. *Topics in Linguistics, 26*(1), 20–40. https://doi.org/10.17846/topling-2025-0002

(1) a. He writes of the plight of the OD consultant who **fears to lose** his job.[1]
b. Commitment phobes **fear making** the 'wrong decision'.
c. "I **dread to imagine** what will happen if things continue like this," he said.
d. I know he **dreads going** to work every day.

As can be seen in the examples, the difference in function between the two non-finite complements can be rather subtle. In some cases, the constructions could easily be exchanged (as in 1a and 1b above), while in others this would arguably yield a less acceptable result (1c and 1d). The distinction between them is typically made in terms of volition: the infinitive contains an element of intentionality, while the gerund does not and is similar in function to an NP object (Huddleston and Pullum, 2002, p. 1243). Yet volition cannot be the sole motivation for complement choice, as evidence from the data is far from uniform (e.g., both 1a and 1b describe unintentional actions, while both 1c and 1d intentional actions). This indicates that the actual use of a specific form results from the interaction of various probabilistic factors, including, apart from the semantic content, a number of functional and contextual regularities and constraints. The multivariate statistical analysis applied in this study allows for an examination of this kind of variation and, subsequently, an estimation of the relative importance of the different motivations involved.

The study is set within the framework of usage-based Construction Grammar (Hoffmann and Trousdale, 2013; Perek, 2023; Fried and Nikiforidou, 2025) and shares the basic assumptions of its approach to argument structure (Goldberg, 1995; 2006; Perek, 2015): complementation patterns are regarded as constructions, pairings of form and schematic meaning, with sets of specific collocational preferences, syntactic idiosyncrasies and pragmatic functions. Constructions emerge from usage and since the process of their schematization is gradual, informative generalizations can be traced at all levels of abstraction. Sometimes it is the low-level generalizations, pertaining to specific groups of predicates, that are more explanatory of a construction's behaviour (Perek, 2015, pp. 105-139). This study sets out to examine if this is the case with apprehension verbs.

English complement constructions form a complex dynamic system, where some distinctions are relatively stable and categorial, i.e., motivated by clear functional differentiation between complement types, and others are non-categorial and more fluid, resulting from on-line speakers' choices, whose motivation is multifactorial and probabilistic. Such "unstable variation" may result from an ongoing diachronic change, in which one complementation pattern is gradually taking over the function of another (Cuyckens, D'Hoedt and Szmrecsanyi, 2014, pp. 183-185). The two verbs analysed in this study are a case in point, showing non-categorial variation between the more conservative to-infinitive complement and the gerund, a pattern which is gaining ground diachronically and dynamically expanding

[1] All examples are from the COCA corpus (Davies 2008-). Emphasis added.

into new contexts (De Smet, 2013, p. 2). Such a situation is worth exploring through multivariate analysis, as it allows one to investigate the various forces that are at play and estimate their respective impact on speakers' choices.

The text is structured as follows: the present section introduces the scope of the study and its main assumptions, while the next reviews previous research on the complementation of apprehension predicates, in order to identify the potential factors in complement choice, which will then form a set of predictors to be used in the analysis. Section 3 introduces logistic regression as a multivariate statistical technique and explains the process of retrieving and coding the data, with further detail regarding the distinctions made. Section 4 describes the two logistic regression models built for the non-finite complements of *fear* and *dread*, the factors revealed to be significant in the choice of the construction and their relative salience, with a comparison between the two verbs. Section 5 discusses the implications of the results for the study of the complementation of apprehension predicates, while Section 6 offers some final remarks on the adopted methodology and perspectives for further research.

2. Non-finite complementation of apprehension predicates

In cognitive and functional complementation research (for an overview see Egan, 2008, pp. 45-88; Smith, 2009, pp. 363-366; De Smet, 2013, pp. 19-32; Kaleta, 2014, pp. 19-45) complement constructions are generally assumed to be motivated by various degrees of compatibility between the lexical meaning of the matrix predicate and the more schematic meaning of the complement structure. At the same time, the impact of a variety of other factors is acknowledged, not necessarily semantic in nature. In recent studies of English complementation emphasis has been placed on the role of discourse mechanisms (Ruohonen and Rudanko, 2021), the variation across registers and language varieties (Deshors and Gries, 2016), and the influence of diachronic development on the current usage (De Smet, 2013; Rickman and Rudanko, 2018). The need for finer-grained analysis has frequently been emphasized in order to capture the interplay between lexical and constructional meaning: how generic-level schemas are realized with specific groups of matrix predicates (Smith, 2009) and semantically related groups of complement verbs (Perek, 2015, pp. 105-139). This approach is also adopted in the present study, which investigates how more general syntactic and pragmatic regularities associated with the two non-finite structures motivate the choice of complement constructions of two semantically related verbs.

In a similar vein, Smith (2009) describes the contrast between the to-infinitive and the -ing complements in terms of the notions of conceptual distance and conceptual overlap, postulating a set of more specific constructional sub-schemas for different groups of matrix verbs. He links the to-infinitive with a set of polysemous subschemas, all of which "reflect at least some relevant aspect(s) of the source-path-goal image schema" (2009, p. 371). As variants of the notion of conceptual distance, they are motivated by semantic extensions of

different aspects of the basic spatial sense, including concepts such as futurity (motivated by the fact that in the literal sense of the schema, the goal is reached subsequent to traversing the path), volition (the goal is reached intentionally) and the holistic construal of the complement process (the goal is construed as a discrete entity). By contrast, the -ing form offers an internal perspective on the process, and its constructional subschemas involve the idea of overlap between the matrix and complement events, which may be either actual, prior, hypothetical, or subjectively implied (2009, pp. 376-380).

Smith (2009, pp. 373-374) counts *fear* among a set of predicates (*fail, refuse, forget, be afraid/be reluctant to*) whose use of the infinitival complement relies on multiple motivations and denotes a holistically construed, unrealized goal. On the other hand, *dread -ing* belongs among the "subjective overlap predicates", including verbs such as *avoid* or *resist*, which evoke "some kind of implied necessity or obligation" (2009, p. 380) as they profile a complement event which, from the speaker's perspective, was due to or likely to occur. This description accurately captures the sense of inevitability connected with *dread -ing*, but as opposed to other verbs in this category, *dread* does not necessarily imply the non-occurrence of the event (on the contrary, *dread -ing* often suggests that the event is unavoidable). In fact, corpus data shows that the usage context of both *dread* and *fear* may or may not give rise to the inference concerning the actual realization of the event, a fact that applies to both complement forms. This is a factor to be examined in the analysis.

The idea of agentivity and related notions of volition and control have been proposed as the main motivation for the use of the non-finite complements of *fear* and *dread* in many studies, including Egan (2008, pp. 249-252), Kaleta (2014, pp. 168-170) and Duffley and Fisher (2021), amongst others. The notion has been formalized as the Choice Principle (Rudanko, 2010, pp. 12-13), which associates the infinitive complement with [+Choice] contexts and the gerund form with [–Choice] contexts. The distinction refers specifically to the agentivity of the implicit subject: the *to*-infinitive complements tend to have agentive subjects, while the gerunds non-agentive subjects. The principle proved to be the main explanatory factor in a series of multivariate analyses of the two non-finite complements of the adjective *afraid* (*to-infinitive* and *of -ing* patterns) in the Strathy corpus of Canadian English, the BNC and the NOW corpus (Ruohonen and Rudanko, 2021). Rickman and Rudanko (2024) also show it to be an important factor in the complementation of *fear* in New Zealand English.

In a recent study focusing on the complementation of *fear* and *dread* in the COCA corpus, Duffley and Fisher contrast the to-infinitive and the gerund in terms of, respectively, negative predisposition towards the occurrence of the event and the emotional state whose source is the event (2021, pp. 91-94). The authors associate the infinitive with the notion of volitionality, but their understanding of the term seems to be fairly broad, subsuming, apart from the typical agentive contexts, also examples of indirect volition (i.e., an unintentional effect of the action performed by the subject), and even cases of passive or non-volitional

complement verbs. The authors argue that all these structures encode "negative predisposition on the part of the subject towards moving to the actualization of the infinitive's event" (2021, p. 91). However, such a "negative predisposition" seems to result from a negative evaluation of the matrix verb as much as from the schematic meaning of the complement construction. In this study the definition will be narrowed down, in an attempt to provide a more objective measure, potentially capable of distinguishing between the alternative complement constructions.

Another important aspect of complement variation is the immediate syntactic context. A specific factor that has been found to influence the complement choice is the *horror aequi* principle. The term refers to the well-attested aversion to grammatical repetition, leading to the avoidance of adjacent identical forms, in this case those of the main verb and the complement verb (Vosberg, 2003, p. 315). Thus, an -ing form of the matrix verb will generally discourage the gerundive complement and favour the infinitive, while the opposite tendency can be observed for the infinitives of the matrix verbs.

The choice of the complement form may also be affected by the structural complexity of the clause. The Complexity Principle states that in cognitively more complex environments, which require more processing effort, more explicit structures are typically preferred (Rohdenburg, 2003, p. 217). A related regularity, formalized as the Extraction Principle, pertains specifically to the influence of structural discontinuities in the matrix clause on the choice of the non-finite complement: in sentences where an element of the subordinate clause is situated beyond the clause boundary due to structures such as wh-questions or relative clauses, more firmly entrenched infinitive complements are generally preferred to the diachronically newer gerunds (Vosberg, 2003, p. 308). Recent research offers ample evidence for this extraction effect; for instance, Ruohonen and Rudanko (2021, p. 82) show that it influences the complementation of *afraid* in the NOW corpus. Another complexity-related factor that may prompt the use of the infinitive complement is negation in the matrix clause (Ruohonen and Rudanko, 2021, p. 75).

The behaviour of the two complement constructions is also influenced by their diachronic development. The spread of the gerund at the expense of the infinitive constitutes a major part of the set of changes referred to as the Great Complement Shift, which has affected English in recent centuries (Rohdenburg, 2006). The expansion of the gerund, although widespread, is not uniform: there are contexts and constructions in which the infinitive resists change, which may be conditioned syntactically (Rohdenburg, 2006) or lexically (De Smet and Cuyckens, 2007). In such situations, the constructions may acquire more specific senses through the conventionalization of pragmatic inference, as is the case with *like to-infinitive*, which gained the meaning of a habitually performed action (De Smet and Cuyckens, 2007, pp. 13-14). Gerunds, on the other hand, show a tendency to generalize in function across various contexts of usage, to the point of becoming semantically neutral clausal complements,

equivalent to nominal direct objects (De Smet 2013: 141). We are thus dealing with a change in progress, motivated by a number of factors that are not entirely predictable.

The two recent studies that directly inspired the present research are Duffley and Fisher (2021) and Rickman and Rudanko (2024). The former examines the same verbs and corpus, but uses no multivariate statistics, while the latter applies logistic regression to the analysis of the verb *fear* in New Zealand English. However, Rickman and Rudanko focus on the Choice Principle and exclude examples with extraction and *horror aequi*. By contrast, this study includes all the above-mentioned factors within the same multivariate model in order to be able to compare their impact.

The present study also builds on Podhorodecka (2025), where the non-finite complements of *fear* and *dread* in the COCA corpus are subjected to distinctive collexeme analysis, which characterizes the two alternative constructions in terms of the complement verbs which occur in one of them significantly more often than expected, given the frequency of both constructions in the corpus. The results reveal certain regularities in usage: *fear to-infinitive* tends to combine with verbs of cognition, perception and communication, typically describing short events with a clearly defined endpoint; *fear -ing* strongly correlates with verbs and structures profiling non-agentivity: *be* passives and adversative *get* passives, *have to* for obligation, involuntary action verbs (*fail*, *lose*) and link verbs describing a change of state. *Dread to-infinitive* typically occurs with a small but uniform group of mental verbs centred around the key collexeme *think*, while the most characteristic complements of *dread -ing* denote deictic motion, possession and obligation. The results show that the features potentially affecting the complement choice are agentivity, the semantic domain of the complement verb, its transitivity type (also suggested in Ruohonen and Rudanko, 2019), telicity (i.e., well-defined endpoint) and duration. These factors will be included in the analysis.

3. The method and data

The statistical tool applied in this study is logistic regression, a technique which estimates to what extent the occurrence of a specific binary variable (in this case either the to-infinitive or the gerundive complement) is impacted by the presence of individual factors in a set of data (Levshina, 2015, pp. 253-275). This kind of technique allows researchers to accurately differentiate between nearly synonymous lexical items and constructions, or to compare the behaviour of various linguistic items in different environments, such as language varieties (Deshors and Gries, 2016) or learner and native usage (Deshors and Gries, 2014). Logistic regression relies on a close examination of a sample of actual tokens of usage, each of which is manually coded for a number of features. The method offers a fine-grained, multivariate approach, which can establish the relative importance of particular factors in complement choice.

Samples of each of the four constructions were downloaded from the corpus using the search strings: [fear].[*v] _v?g, [fear].[*v] to _v?i, [dread].[*v] _v?g, [dread].[*v] to _v?i. These queries produce tokens with the matrix verb directly followed by the complement, which excludes from the analysis different-subject constructions and same-subject constructions with intervening lexical material. The reason for leaving out different-subject constructions is that the choice of their complement form is motivated by other factors: the construction is unattested in the COCA corpus with *dread to-infinitive*, while with *fear to-infinitive* (e.g., *She **fears herself to be** a freak.* = She fears that she might be a freak.) it has a different function, expressing a judgment about the ontological status of the situation rather than an attitude towards it. In Egan's terminology (2008: 22-23), it is a mental process construction rather than an attitude construction. Same-subject constructions with intervening material (adverbs and negative particles) constitute approximately 2% of each of the constructions, which is not enough for this factor to produce meaningful results if included in the analysis as a potential complexity-related feature.

Since the COCA corpus yielded only 98 tokens of *dread to-infinitive*, the sample of this construction was supplemented with examples from the American section of the NOW corpus. Although the two corpora differ in their composition (NOW is based on online newspapers and magazines, while COCA is more balanced, covering a range of genres), the usage of the target construction seems to be similar: 7 out of the top 10 complement verbs are the same in the two corpora (i.e., practically all those that occur more than twice), and the most frequent collocate of the construction is the verb *think*, which constitutes 30% of its uses in COCA and 44% in NOW.

The process of constructing the samples for the analysis involved the manual removal of false positives, including examples which contained *fear/dread* as erroneously tagged nouns (2a), *-ing* forms used as participial modifiers (2b) and gerundive subjects of a finite clause complement (2c). Tokens with passives of the verb *fear* were also excluded (2d) since this construction does not have a corresponding *-ing* variant (* *Dozens are feared having died.*)

(2) a. They use **fear to make** people do what they want.
b. They **fear rising taxes** and the damage that is going to ensue from ObamaCare.
c. They **feared speaking up** would damage their careers.
d. Dozens **are feared to have died** in the repression of the protests.

Finally, 250 random tokens of each of the four constructions were coded for a set of 12 features: 3 related to the whole sentence, 3 to the matrix verb and 6 to the complement verb. The choice of those variables was informed by the previous research discussed in section 2.

Table 1. Initial set of regression variables

Scope	Variable	Levels
The sentence	Extraction	extraction / none
	Inference	triggered / absent
	Specificity	generic / specific
The matrix verb	Verb form	finite / infinitive / present participle
	Aspect	perfective / progressive / simple
	Polarity	negative / positive
The complement verb	Domain	action / cognition / communication / motion / perception / state
	Voice	active / passive
	Transitivity	intransitive / link verb / transitive
	Agency	agentive / non-agentive
	Telicity	atelic / telic
	Duration	punctual / durative

Extraction is a variable related to the immediate syntactic context of the complement construction. Tokens with extraction feature a clause element situated beyond the clause boundaries due to, e.g., a question, fronting, or a relative clause. Such extracted elements are underlined in the examples below: they can be either objects (3a, 3b) or adjuncts (3c). Even though such contexts have been found to favour the infinitive complement (Vosberg, 2003, p. 308), the examples show the usage of both constructions:

(3) a. She then found it hard to concentrate at work, resulting in the silly slipups she'**d feared making**.
b. There are certain subjects parents **dread talking** to their kids **about** – like sex, drugs, and money.
c. (…) move unashamedly into the world of the empirical, where theorists **fear to tread**.

The distinction labelled as "inference" was prompted by regularities observed in the data, as well as Smith's (2009, p. 380) notion of "implied necessity or obligation", i.e., the subjective construal of the complement event as likely or due to happen from the speaker's perspective. Its purpose is to investigate the pragmatic functions of the complement constructions, specifically whether one of them is more likely to give rise to the contextual inference of the occurrence or non-occurrence of the complement event. Inference is marked as "triggered" if the sentence with an affirmative matrix verb implies that the complement action did not take place, and its negative form implies that it did. Consider the following:

(4) a. There are men in the House of Commons who pride themselves that they **have not feared to tell** the Prime Minister what they think of him.
b. They just perish away like that because they **fear going** to the hospital.
c. I **dreaded going** to dances because I was forced to wear heels.
d. Why many spouses don't come clean about what they buy: 35% **dread getting** a lecture.
e. I **dread to imagine** what might have occurred had we not intervened.

In 4a, the negative matrix verb, *have not feared*, triggers the inference that the men did in fact tell the Prime Minister their opinion, while in 4b the people do not go to the hospital, because fear prevents them from seeking treatment. The effect is similar to that produced by polarity-reversing implicative verbs such as *fail* or *avoid*, but its mechanism is different: with implicatives the entailment is part of their lexical meaning and cannot be cancelled by context. By contrast, with *fear* and *dread* the inference is potentially triggered by the construction, but the context may or may not support it. Inference is marked as "absent" if the action is either performed despite the subject's feeling of apprehension (4c) or its occurrence remains potential (4d). *Dread to think* and related constructions, exemplified in 4e, are somewhat paradoxical in this respect as they involve both thinking and not thinking about the topic in question: first the subject brings to mind a particular topic, which causes them to experience the feeling of dread, so as a result they choose not to consider the topic in any more detail. Since the construction explicitly profiles the latter part of the scenario (the feeling of dread causing one not to contemplate a topic), it is classified as inference-triggering.

The exact mechanism of contextual inference with apprehension predicates requires further research. It seems to correlate with the volitionality of the complement verb, reinforced by contextual information. The inference proceeds from a degree of subject control over the complement process and the willingness or unwillingness to bring it about. Lexical idiosyncrasies of the matrix verbs also seem to play a part: note that 4b and 4c share the same complement verb and construction, but with different matrix verbs the sentences yield opposing inferences about the actual occurrence of the event.

The next distinction concerns the specificity of the events denoted by the constructions. They are divided into generic and specific. The former occur on a number of different occasions (5a), while the latter are either single events (5b) or events repeated (5c) on the same occasion or within a particular period of time (Egan, 2008, p. 191).

(5) a. Let us never negotiate out of fear, but let us never **fear to negotiate**.
b. "Dora!" I called, still **dreading to learn** the worst, "What happened?"
c. When Crookham first led trips to the riverbanks, kids **feared to put** their hands in the stinking water.

With regard to the form of the matrix verb, the tokens divide into infinitives and present participles (both as non-finite forms and elements of progressive tense constructions), with all the remaining forms summarily labelled as "finite". The reason behind this distinction is to examine the potential role of the *horror aequi* principle, that is whether the form of the matrix verb inhibits an identical form of the complement. So, it can be expected that the infinitive form of the matrix verb will favour the gerund (*to fear/dread -ing*), while the present participle will lean towards the infinitival complement (*fearing/dreading to inf*). As can be seen below, the sample contains examples which both conform to (6a, 6b) and go against (6c, 6d) the *horror aequi* principle, with both matrix verbs and both patterns of complementation:

(6) a. He clung to his office, **fearing to leave**.
    b. We don't want people **to dread coming** to our events
    c. He should not have any reason **to fear to have** the machine count checked by a hand count.
    d. I was **dreading introducing** you, with your very difficult family name.

The remaining distinctions related to the matrix verb are aspect and polarity, that is the presence or absence of negation in the matrix clause, which may be a complexity factor (Ruohonen and Rudanko, 2021, p. 75). *Fear* and *dread* as verbs of attitude are limited in their aspectual potential, and the matrix verbs other than the default simple form were fairly rare in the data. The distinction into simple, progressive and perfective forms is introduced in order to examine whether such atypical structures (7a, 7b) correlate with a specific complement construction.

(7) a. He **was dreading to arrive**, figured his place must have gone to rack and ruin.
    b. (…) a child of an alcoholic who **fears having been** imprinted on abusive relationships and not being able to have 'healthy' (happy?) ones.

Distinctive collexeme analysis (Podhorodecka, 2025) revealed significant associations of each of the two constructions with particular groups of semantically related complement verbs, so in order to account for such lexical motivation, complement verbs are assigned into 6 domains: action (8a), cognition (8b), communication (8c), motion (8d), perception (8e), state (8f, 8g). The last term, for the purposes of the present analysis, refers to the conceptual domain that the verb evokes rather than the stative/dynamic distinction: it covers verbs which denote being in a particular state (8f) or changing one's state (8g). They can be either stative or dynamic, and so can be verbs of cognition and perception.

(8) a. Most people **fear to take** risks.
    b. I **dread to think** what they taught their students.
    c. I **dread telling** my mother about Bill's moving out.

d. You **fear to go** into those mines.
e. She was telling me things every mother **dreads to hear**.
f. 81% of 10 Year Olds **Fear Being Fat**.
g. You never had to **fear becoming** pregnant and losing your job.

The variable of "voice" introduces the differentiation of the complement verbs into active (9a) and passive, including get-passives (9b), while "transitivity" divides them into intransitive verbs (9c), link verbs (9d) and transitive verbs (9a). The final term covers all the structures with at least one object, irrespective of their complexity (e.g., ditransitives or clausal objects).

(9) a. I have **feared to give** him your kind message.
b. New Yorkers **don't fear getting smacked** by falling coconuts (…)
c. One **dreads to wake up** and read the front page these days.
d. She **dreaded being** home alone with him.

Agentivity is a key notion in the analysis of apprehension predicates, yet operationalizing it is far from straightforward. It has been described in terms of a number of more specific concepts, such as volition, instigation, and control (Naess, 2007, pp. 32-36). For the purposes of this analysis, tokens are labelled as agentive if their complement verb describes an action which is purposefully instigated and controlled by the subject (10a, 10b). The reason for focusing on the complement verb was to avoid impressionistic judgments based on the semantic impact of the matrix verb and contextual factors. Adopting this more specific and thus arguably more objective criterion directly follows from the original formulation of the Choice Principle, which focuses on the agentive role of the implied subject of the complement verb (Rudanko, 2010, pp. 12-13). The procedure also rests on the assumption that the agentive or non-agentive meaning of the entire construction will be reflected in its choice of collocates. The tests adopted to establish the agentivity of the complement clause were the acceptability of the imperative (Rudanko, 2010, p.13) and modification with a purpose clause (*in order to*…) (Rickman and Rudanko, 2024, p. 9). Consider the following examples:

(10) a. I **fear to say** the dreaded phrase "climate change".
b. Most people **dread going** to the dentist's office.
c. (…) she still **feared being hungry** again.
d. Train yourself to let go of everything you **fear to lose**.
e. People who cheat tend **to fear being cheated on**.
f. His staffers **dread getting caught** in a traffic jam with him.
g. (…) as if by the least inadvertence they **feared to draw** down his despotic displeasure.
h. Chinese archaeologists **feared setting off** a political minefield.

Tokens classified as agentive feature an action purposefully performed by a volitional instigator (10a, 10b). Even if the action is potentially influenced by external necessity, as in (10b), it is still under the subject's control and could be logically paraphrased using the tests (*Go to go to the dentist!, going to the dentist in order to have a tooth extracted*). Tokens labelled as non-agentive contained either stative verbs (10c), involuntary action verbs (10d), or various passive and passive-like constructions (10e, 10f). Predictably, the sample contained some borderline cases: tokens describing unintended indirect results of the subjects' actions (approximately 6% of the *fear* sample, less than 1% for *dread*). They were classified as non-agentive only if explicitly described as accidental (10g), otherwise they were labelled as agentive, since technically they pass the agentivity tests (10h: *Set off a political minefield in order to confuse your opponents!*).

The final two variables, telicity and duration, are related to the lexical aspect of the complement verb. Events denoted by telic verbs have a natural endpoint (11a), while those described by atelic verbs do not (11b). Durative events extend in time (11b) and punctual events are instantaneous or conceptualized as such (11c) (Croft, 2012, pp. 34-36).

(11) a. She **dreaded going** inside the school.
b. But she **dreaded being** alone.
c. Most players **fear hitting** the ball too hard.

4. The results

The resulting data frames formed the basis for two separate logistic regression models (*fear -ing* vs *fear to-inf*; *dread -ing* vs *dread to-inf*) built in the R software environment (R Foundation for Statistical Computing 2022). In the process of building the models, the distinctions which did not contribute to their predictive power were gradually removed, eventually yielding the final set of 5 significant factors, which were very similar, but not identical for the two verbs: extraction, inference, verb form, lexical domain and agency.

Table 2 below contains a summary of the results obtained from the two logistic regression models. The full regression output, together with the model diagnostics and variance inflation factors, can be found in Appendix A (*fear* -ing vs *fear* to-infinitive) and Appendix B (*dread* -ing vs *dread* to-infinitive). The first column lists the regression variables, arranged according to their predicted impact on the complement choice. One of the levels of each factor is not listed in the table, since it serves as the reference level, to which other levels are compared (e.g., for the variable "verb form" the reference level is "finite", as opposed to the present participles and infinitives listed in the table; for "extraction" the reference is "no extraction", as it is a binary distinction). The coefficients describe the strength of effect of each individual factor. A negative coefficient indicates that a particular feature correlates with the gerund, whereas a positive figure shows its correlation with the to-infinitive. Finally, p-values measure the

statistical significance (i.e., the reliability) of the result. Lower values indicate more reliable results and the significance levels are coded with asterisks (*** 0.001, ** 0.01, * 0.05, · 0.1).

Table 2. Summary of logistic regression results for the verbs *fear* and *dread*

	FEAR		**DREAD**	
	Coefficient	**p-value**	**Coefficient**	**p-value**
Intercept	**-2.1457**	< 2e-16 ***	**-1.4680**	1.25e-08 ***
Verb form: present participle	**3.6496**	5.88e-08 ***	**0.7975**	0.023453 *
Verb form: infinitive	**-2.3603**	0.03085 *	-1.4344	0.172608
Extraction: extraction	**2.9353**	8.50e-09 ***	**1.8847**	8.08e-09 ***
Inference: triggered	**1.7934**	4.12e-07 ***	**1.3739**	0.000285 ***
Agency: agentive	**0.9173**	0.01949 *	--	--
Domain: cognition	**2.7793**	0.00628 **	**4.4153**	6.06e-08 ***
Domain: communication	0.2527	0.53727	-0.0407	0.928074
Domain: motion	0.1580	0.67500	**-0.6418**	0.081442 ·
Domain: perception	**1.6215**	0.00447 **	0.4790	0.214596
Domain: state	-0.9789	0.21915	-0.8392	0.129069

The form of the matrix verb emerges as the most powerful predictor for the complement of *fear*: the present participle has a high positive coefficient (3.6496) and an extremely low p-value (5.88e-08), which means that it strongly correlates with the infinitive complement (*fearing to-infinitive*), while the infinitive form of *fear* as the matrix verb attracts the gerundive complement (*to fear -ing*), which is indicated by its negative coefficient of -2.3603. The latter relationship is, however, slightly weaker, both in terms of its impact and its statistical significance. This shows that the *horror aequi* principle strongly influences the choice of the complement for the verb *fear*, particularly the choice of the infinitive complement after the present participle of the matrix verb. The impact of *horror aequi* is less pronounced with the verb *dread*, whose present participle form only weakly (0.7975) correlates with the infinitive complement, with the connection between the matrix infinitive and gerundive complement failing to achieve statistical significance. In fact, the verb *dread*, as opposed to *fear*, occurs several times in the sample in its progressive form followed by the -ing complement (as in 6d above: *was dreading introducing you*). The higher susceptibility of the -ing form to the *horror aequi* principle, visible with both *fear* and *dread*, is consistent with Ruohonen and Rudanko's findings (2021, p. 84).

Extraction is the second strongest predictor for the infinitive complement with both matrix verbs, but the strength of its effect is again much higher for the verb *fear* (2.9353) than for *dread* (1.8847), which shows that discontinuous syntactic contexts, requiring more processing

effort, indeed tend to prefer the more historically established infinitive construction. The results associated with the first two factors indicate that the verb *fear* is particularly susceptible to the influence of its immediate syntactic context, as the two strongest predictors in its complement choice (*horror aequi* and extraction) are syntactic. The pattern is similar for the verb *dread*, but the correlations are much weaker.

The next predictor relates to whether or not the construction gives rise to the contextual inference of the occurrence or non-occurrence of the complement event. Inference-triggering contexts are characteristic of the infinitive complement constructions in the case of both verbs, although again the correlation is stronger for *fear* (1.7934) than for *dread* (1.3739).

The agentivity of the action has been described as the main semantic motivation for the complement choice with apprehension predicates, yet, surprisingly, it did not prove to be the strongest predictor of complement choice in the examined sample. For the verb *fear*, it prompts the infinitive complement with moderate strength (0.9173), whereas for *dread* it was not included in the final model: it did not show any impact or statistical significance, and had to be removed from the analysis. This is most probably due to the fact that with this verb the non-volitionality is implied rather than overtly expressed, and many of its complement verbs describe actions that are technically agentive, yet practically unavoidable (as in 1d or 10b above).

In terms of semantic domains, the logistic regression model confirms the results of the distinctive collexeme analysis (Podhorodecka, 2025): cognition verbs prove to be the most characteristic verb type for the to-infinitive complement with both matrix verbs; however, this correlation is much stronger (4.4153) for the verb *dread*, whose collexemes form a more uniform group of mental predicates. There are two more correlations that achieved statistical significance: perception verbs prompt the infinitive complement fairly strongly with the verb *fear*, while for *dread* there is a marginal correlation of the gerund with motion verbs. This again is consistent with the results of the DCA in Podhorodecka (2025).

Finally, the intercept, listed at the top of the table, stands for the situation where all the predictors are at their reference levels. In this case, it signifies a combination of a finite form of the matrix verb with a non-agentive action verb as the complement, in a context with no extraction or inference. Since the coefficients are negative for both *fear* and *dread* (respectively -2.1457 and -1.4680), such a context is characteristic of the gerundive complement construction with both matrix verbs, which is exemplified in 12a and 12b below:

(12) a. I **feared losing** the respect of my parents.
b. He was your teacher. Students **dreaded getting** stuck with him.

5. Discussion

There is a considerable degree of overlap between the features that motivate complement choice for the two analysed verbs, but also some noticeable differences in the relative

importance of the individual factors: for the verb *fear* the syntactic context is decisive, while with *dread* the lexical motivation seems to prevail.

Table 3 below summarizes the factors which influence the complement choice with the verb *fear*, arranging them by the strength of their association with each complement form.

Table 3. Factors influencing the complement choice with the matrix verb *fear*

***FEAR* TO INFINITIVE**	***FEAR* -ING**
Present participle	No extraction
Extraction	Infinitive
Cognition verb	No inference
Perception verb	Non-agentive verb
Inference	
Agentive verb	

The model constructed for the verb *fear* has a larger number of statistically significant predictors and their impact is generally stronger, which indicates that the complement choice of this verb is motivated by a more complex set of factors. The immediate syntactic context emerges as the primary type of influence, as the *horror aequi* principle and extraction are the two highest-ranking predictors for both complements. The secondary motivation for the infinitival complement is lexical: the usage of the to-infinitive construction is prompted by cognition and perception verbs as complements, but no such correlation with a specific verb type was revealed for the gerund. Inference-triggering context and agentivity prompt the infinitive construction and inhibit the gerund. Thus, the features motivating the complement choice with the verb *fear* follow a cline from the immediate syntactic context of the construction, through verb type and pragmatic inference to agentivity.

Table 4 lists the factors correlating with the two non-finite complements of *dread*:

Table 4. Factors influencing the complement choice with the matrix verb *dread*

***DREAD* TO INFINITIVE**	***DREAD* -ING**
Cognition verb	No extraction
Extraction	No inference
Inference	Motion verb
Present participle	

The strongest predictor for the infinitival construction is a cognition verb as the complement, which results from the high frequency of the set phrase *dread to think* and its near-synonyms. For the gerundive complement, however, the lexical motivation is much less prominent, as motion verbs are its weakest predictor. This is consistent with the diachronic trend affecting the two complement forms: -ing complements are spreading into new contexts

at the expense of the infinitive (De Smet, 2013, p. 2), which is becoming increasingly restricted to lexically fixed combinations.

Extraction, the second predictor for the to-infinitive, has much more impact than the *horror aequi* principle, as the present participle of the matrix verb is the least important predictor for the infinitive construction (*dreading to-inf*), while the combination of the infinitive of the matrix verb with gerunds (*to dread -ing*) did not prove statistically significant. Finally, contextual inference, or its absence, prompt with moderate strength for, respectively, the to-infinitive and the gerund. *Dread* is thus more lexically conditioned, and less susceptible to the influence of the *horror aequi* principle than the verb *fear*, in whose complementation the syntactic factors play the key role. The reason may be that the difference in pragmatic function between the complements of *dread* is slightly more pronounced, as *dread to-infinitive* is effectively limited to mental verbs (Podhorodecka, 2025) and partially lexicalized as an expression of emphatic refusal (Duffley and Fisher, 2021, p. 93). The tentative conclusion is that the less distinct the two forms are in their lexical preferences and pragmatic functions, the more decisive the strictly formal factors prove to be.

The limited role of agentivity in comparison with the immediate syntactic context may to an extent result from the criteria adopted for the purposes of this study, with the focus on the type of action denoted by the complement verb, rather than a holistic interpretation of the whole sentence. On the other hand, it may be a symptom of a more general regularity: in the case of complement constructions with no stark, categorial difference in meaning and function, the influence of the immediate syntactic context seems to far override any semantic considerations. Further evidence for this is the fact that many other semantic distinctions did not prove significant and had to be discarded at the early stages of the model building (e.g., specific/generic, durative/punctual and telic/atelic).

For both verbs, the -ing complement seems to be the default, unmarked version: it is more frequent and associated with more basic contexts (no inference, no extraction). By contrast, the to-infinitive requires more specific syntactic (extraction, *horror aequi*) and contextual motivation. This shows the importance of the diachronic development path for the present state of the complementation system, with the more conservative construction becoming more specialized and the developing construction becoming more functionally neutral as its usage expands.

6. Conclusions

The results of the analysis show that semantically similar matrix predicates can be expected to rely on similar factors motivating their complement constructions, but the importance of particular features may vary depending on the characteristics of the main verb. The analysis also suggests that, even taking into consideration particular verb classes and different constructional sub-schemas, verb complementation cannot be accounted for solely in semantic

terms but is instead conditioned by a multi-layered complex of probabilistic factors. Semantic features, such as agentivity, constitute only one aspect, while strictly formal mechanisms connected with the immediate syntactic context, such as *horror aequi* and extraction, may in fact prove to be decisive in the actual usage of functionally similar complement constructions. Additionally, the role of the lexical preferences of the matrix verbs and the associated pragmatic functions cannot be overlooked.

A multivariate statistical analysis of the corpus material has proved to be a method effective in teasing apart various aspects of the complex motivations involved in complement choice. Obviously, the analysis is not without its drawbacks. Operationalizing the notion of agentivity proved to be a challenge and the results might be clearer if the distinctions were to be refined (e.g., broken down into elements such as intention, instigation and control) or if the concept of agentivity was delineated relative to the elements of the transitive schema, as defined by Hopper and Thomson (1980, pp. 251-252), which was also suggested by Ruohonen and Rudanko (2019). The factors examined in this study certainly do not constitute a complete list; for instance, prosodic features and phonological weight may prove to be another type of extra-semantic variables which may affect the complement choice (De Smet 2013, p.30). However, both the method used and the set of variables established in this study could prove useful in examining other apprehension predicates (e.g., adjectives *afraid*, *scared*, *frightened*) and their remaining complementation patterns (e.g., finite clauses or different-subject constructions).

Appendix A

Logistic regression output: *fear -ing vs fear to*-infinitive

Coefficients:

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-2.1457	0.2590	-8.285	< 2e-16 ***
verb_formppart	3.6496	0.6731	5.422	5.88e-08 ***
verb_forminf	-2.3603	1.0932	-2.159	0.03085 *
extractionextr	2.9353	0.5098	5.758	8.50e-09 ***
inferencetriggered	1.7934	0.3542	5.064	4.12e-07 ***
agencyagentive	0.9173	0.3926	2.336	0.01949 *
domaincog	2.7793	1.0171	2.733	0.00628 **
domaincom	0.2527	0.4096	0.617	0.53727
domainmotion	0.1580	0.3767	0.419	0.67500
domainperc	1.6215	0.5704	2.843	0.00447 **
domainstate	-0.9789	0.7967	-1.229 0.21915	-0.9789

Model statistics:

		Model Likelihood Ratio Test		Discrimination Indexes		Rank Discrim. Indexes	
Obs	500	LR chi2	296.30	**R2**	**0.596**	**C**	**0.893**
ING	250	d.f.	10	R2(10,500)	0.436	Dxy	0.786
TO-INF	250	Pr(> chi2)	<0.0001	R2(10,375)	0.534	gamma	0.833
max |deriv|	1e-10			Brier	0.128	tau-a	0.394

Variance inflation factors:

Verb form: infinitive	1.024328
Verb form: present participle	1.076380
Extraction: extraction	1.047882
Inference: triggered	1.966144
Domain: cognition	1.038524
Domain: communication	1.150664
Domain: motion	1.200786
Domain: perception	1.145081
Domain: state	1.048356
Agency: agentive	2.143504

Appendix B

Logistic regression output: *dread -ing* vs *dread to*-infinitive

Coefficients:

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-1.46802	0.25788	-5.693	1.25e-08 ***
verb_forminf	-1.43441	1.05172	-1.364	0.172608
verb_formppart	0.79751	0.35195	2.266	0.023453 *
extractionextr	1.88465	0.32681	5.767	8.08e-09 ***
inferencetriggered	1.37393	0.37862	3.629	0.000285 ***
domaincog	4.41534	0.81509	5.417	6.06e-08 ***
domaincom	-0.04072	0.45106	-0.090	0.928074
domainmotion	-0.64180	0.36835	-1.742	0.081442 ·
domainperc	0.47896	0.38594	1.241	0.214596
domainstate	-0.83917	0.55289	-1.518	0.129069

Model statistics:

		Model Likelihood Ratio Test		Discrimination Indexes		Rank Discrim. Indexes	
Obs	500	LR chi2	338.99	**R2**	**0.656**	**C**	**0.909**
ING	250	d.f.	9	R2(9,500)	0.483	Dxy	0.818
TO-INF	250	Pr(> chi2)	<0.0001	R2(9,375)	0.585	gamma	0.842
max |deriv|	2e-09			Brier	0.112	tau-a	0.410

Variance inflation factors:

Verb form: infinitive	1.025168
Verb form: present participle	1.057111
Extraction: extraction	1.046707
Inference: triggered	1.250626
Domain: cognition	1.262551
Domain: communication	1.196602
Domain: motion	1.369619
Domain: perception	1.257573
Domain: state	1.130693

References

Croft, W. (2012). *Verbs: Aspect and causal structure*. Oxford University Press.

Cuyckens, H., D'Hoedt, F., & Szmrecsanyi, B. (2014). Variability in verb complementation in Late Modern English: Finite vs. non-finite patterns. In M. Hundt (Ed.), *Late Modern English syntax* (pp. 182–203). Cambridge University Press.

Davies, M. (2008–). *The Corpus of Contemporary American English (COCA): 410+ million words, 1990–present*. http://www.americancorpus.org

Davies, M. (2008–). *The NOW Corpus (News on the Web): 10.3 billion words, 2010–present*. https://www.english-corpora.org/now

Deshors, S., & Gries, S. Th. (2014). A case for the multifactorial assessment of learner language: The uses of *may* and *can* in French-English interlanguage. In D. Glynn & J. A. Robinson (Eds.), *Corpus methods for semantics: Quantitative studies in polysemy and synonymy* (pp. 179–204). John Benjamins.

Deshors, S., & Gries, S. Th. (2016). Profiling verb complementation constructions across New Englishes: A two-step random forests analysis of *-ing* vs. *to* complements. *International Journal of Corpus Linguistics, 21*(2), 192–218.

De Smet, H. (2013). *Spreading patterns: Diffusional change in the English system of complementation*. Oxford University Press.

De Smet, H., & Cuyckens, H. (2007). Diachronic aspects of complementation: Constructions, entrenchment and the matching problem. In C. M. Cain & G. Russom (Eds.), *Studies in the history of the English language III: Managing chaos: Strategies for identifying change in English* (pp. 1–37). Mouton de Gruyter.

Diessel, H. (2019). *The grammar network: How linguistic structure is shaped by language use*. Cambridge University Press.

Duffley, P., & Fisher, R. (2021). To-infinitive and gerund-participle clauses with the verbs *dread* and *fear*. *Studia Linguistica, 75*(1), 72–96.

Egan, T. (2008). *Non-finite complementation: A usage-based study of infinitive and –ing clauses in English*. Rodopi.

Fried, M., & Nikiforidou, K. (Eds.). (2025). *The Cambridge handbook of construction grammar*. Cambridge University Press.

Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. University of Chicago Press.

Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford University Press.

Hoffmann, Th., & Trousdale, G. (Eds.). (2013). *The Oxford handbook of construction grammar*. Oxford University Press.

Hopper, P. J., & Thompson, S. A. (1980). Transitivity in grammar and discourse. *Language, 56*, 251–299.

Huddleston, R., & Pullum, G. K. (2002). *The Cambridge grammar of the English language*. Cambridge University Press.

Kaleta, A. (2014). *English sentential complementation: A usage-based approach*. Naukowe Wydawnictwo Piotrkowskie.

Levshina, N. (2015). *How to do linguistics with R: Data exploration and statistical analysis*. John Benjamins.

Næss, Å. (2007). *Prototypical transitivity*. John Benjamins.

Perek, F. (2015). *Argument structure in usage-based construction grammar: Experimental and corpus-based perspectives*. John Benjamins.

Perek, F. (2023). Construction grammar and usage-based theory. In M. Díaz-Campos & S. Balasch (Eds.), *The handbook of usage-based linguistics* (pp. 215–231). Wiley Blackwell.

Podhorodecka, J. (2025). Non-finite complement constructions with the verbs *fear* and *dread*: A usage-based study. *Forthcoming*.

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rickman, P., & Rudanko, J. (2018). *Corpus-based studies on non-finite complements in recent English*. Palgrave Macmillan.

Rickman, P., & Rudanko, J. (2024). The Choice Principle and the matrix verb *fear* in recent New Zealand English. *Studia Neophilologica*. https://doi.org/10.1080/00393274.2023.2292025

Rohdenburg, G. (2003). Cognitive complexity and *horror aequi* as factors determining the use of interrogative clause linkers in English. In G. Rohdenburg & B. Mondorf (Eds.), *Determinants of grammatical variation in English* (pp. 205–250). De Gruyter.

Rohdenburg, G. (2006). The role of functional constraints in the evolution of the English complementation system. In C. Dalton-Puffer, D. Kastovsky, N. Ritt, & H. Schendl (Eds.), *Syntax, style and grammatical norms: English from 1500–2000* (pp. 143–166). Peter Lang.

Rudanko, J. (2010). Explaining grammatical variation and change: A case study of complementation in American English over three decades. *Journal of English Linguistics, 38*(1), 4–24.

Ruohonen, J., & Rudanko, J. (2019). Comparing explanatory principles of complement selection statistically: A case study based on Canadian English. *Studia Neophilologica, 91*(3), 296–313. https://doi.org/10.1080/00393274.2019.1616215

Ruohonen, J., & Rudanko, J. (2021). *Infinitival vs gerundial complementation with* afraid*,* accustomed*, and* prone*: Multivariate corpus studies*. Palgrave Macmillan.

Smith, M. B. (2009). The semantics of complementation in English: A cognitive semantic account of two English complement constructions. *Language Sciences, 31*, 360–388.

Vosberg, U. (2003). The role of extractions and horror aequi in the evolution of -ing complements in Modern English. In G. Rohdenburg & B. Mondorf (Eds.), *Determinants of grammatical variation in English* (pp. 305–327). De Gruyter.