Nominalization of clauses: The clausal prolepsis strategy

Abstract

This paper explores the syntax of clausal prolepsis in Dutch, with a specific focus on object clausal prolepsis, the phenomenon where an object pronoun is linked to a CP that is situated at the right edge of the clause. The paper presents new evidence that distinguishes Dutch clausal prolepsis cases where the embedded CP conveys a familiar or a factive interpretation. Previous analyses of clausal prolepsis in other languages, such as German, have proposed two radically different syntactic structures to capture the meaning differences of these two cases. However, this paper proposes a more uniform syntactic structure that reconciles them. The proposed analysis considers clausal prolepsis to realize an underlying syntactic structure of a nominalized clause, similar to structures found in Greek, Persian, and Washo. This analysis captures the meaning differences expressed by clausal prolepsis in a straightforward manner, using a single lexical D-entry and independently attested structural components that can be merged with D. Furthermore, the proposed lexical entry can be extended to uses of the proleptic pronoun in different contexts, such as a definite D, thus avoiding accidental homophony. Additional advantages of the proposed analysis are that it can capture a new generalization describing with which predicates clausal prolepsis is possible in Dutch as well as other aspects of the distribution of clausal prolepsis such as that in Dutch, it is only allowed with verbs or infinitives, but not with nominals. From a theoretical standpoint, the proposed analysis teases apart factivity from familiarity and shows that familiarity does not arise through D but through nominal structure that can be merged with D. From an empirical perspective, the paper concludes that nominalization of a clause is more pervasive cross-linguistically than is usually assumed. Lastly, this study evaluates earlier accounts of clausal prolepsis and shows that those involving a CP base generated in an extraposed position, like the one in Bennis (1986), are not tenable. This is also true for other analyses taking the distribution of the prolepsed CP to be a consequence of phonological rules. Instead, the paper shows that the only necessary mechanism to explain the distribution of the prolepsed CP is Merge, in conjunction with standard assumptions regarding constituency structure.

1 Introduction

This paper examines a specific construction in Dutch, which involves an optional pronoun *het* occupying the object position of a verb and being linked to a CP. This construction is illustrated in (1) and is the focus of our investigation.[1]

(1) Ik hoop (het_i) [dat je wint]$_i$.
I hope it that you win
'I hope that you win.'

Looking at this construction, hereafter clausal prolepsis, this study aims to investigate the relationship between the pronoun *het* and the doubled CP in the underlying syntactic structure. Numerous analyses have been conducted on clausal prolepsis in various languages, including Dutch (Bennis 1986) and others (Gluckman 2021, Postal and Pullum, Rothstein 1995, Stroik 1996 i.a.). The consensus among these analyses has been that clausal prolepsis represents a uniform phenomenon with a unique syntactic structure across different verbs. However, recent work by Frey (2016) and Sudhoff (2016) has challenged this prevailing assumption by demonstrating that the properties of clausal prolepsis may vary depending on the matrix verb, as originally observed by Pütz (1986). As a result, a dual analysis has been proposed, which posits that clausal prolepsis can correspond to two distinct syntactic structures, rather

[1] The data presented in this paper were obtained from informant sessions conducted with four native speakers of Belgian Dutch. I explicitly indicate cases where speaker variation was observed.

than just one. The original observation made by Pütz (1986) shows that the presence of a proleptic pronoun such as *het* and the choice of matrix verb, *betreurt* 'regrets' vs *zegt* 'says,' can significantly affect the meanings expressed through clausal prolepsis. Specifically, in the context of an information-seeking question, such as *wat is er aan de hand?* 'what's the matter?,' the use of clausal prolepsis is possible after the verb *betreurt* 'regrets' whereas it is not allowed after the verb *zegt* 'says' (see Sudhoff 2016 and references therein), as shown in the pair below.

(2) Wat is er aan de hand? 'What's the matter?'

a. Pieter betreurt (het) dat Marie WEGgaat.
Pieter regrets it that Marie goes-away.
'Peter regrets that Mary goes away.'

b. Pieter zegt (* het) dat Marie WEGgaat.
Pieter says it that Marie goes-away
'Peter asserts that Marie goes away.' Sudhoff (2016, (46))

Sudhoff (2016) argues that the interpretation of the embedded clause is the crucial factor that distinguishes the acceptability of clausal prolepsis in these examples. Specifically, in the case of the verb *betreurt*, clausal prolepsis is felicitous, as shown in (2a), since the doubled CP can convey novel information and provide an appropriate answer to an information-seeking question. In contrast, the doubled CP after the verb *zegt* receives a familiar interpretation, ruling out clausal prolepsis as a possible response in (2b), since the doubled CP expresses information that is already known. Frey (2016) and Sudhoff (2016) investigate German clausal prolepsis, specifically focusing on cases where the familiar interpretation is absent, such as in (2a). Both studies propose a nominalization analysis, wherein the proleptic pronoun functions as a D-head taking the prolepsed CP as its complement, as illustrated in (3a). However, neither study extends the nominalization analysis to (2b) and instead proposes a different right dislocation structure. For instance, in Sudhoff, the proleptic pronoun is merged in the complement position of the verb and the prolepsed CP is adjoined to the matrix CP. A similar analysis was also proposed by Bennis (1986) for Dutch clausal prolepsis. In his proposed structure, the proleptic pronoun, *het* in Dutch, is analyzed as a true pronoun merged in the verb's complement position, and the prolepsed CP is base generated as a VP-adjunct, as shown in (3b).

(3) a. $[_{VP}\ V\ [_{DP}\ D\ [_{CP}\ ]]]$ b. $[_{VP}\ [_{VP}\ V\ het_{DP}]\ CP]$

I provide arguments against treating Dutch clausal prolepsis as proposed in Bennis (1986). Yet, my proposal is in line with previous research by demonstrating that Dutch clausal prolepsis should be divided into two cases based on the interpretation of the embedded clause. I present new findings supporting the observation in (2) that the proleptic clause can have a familiar interpretation, especially after non-factive verbs. However, after factive verbs like *betreurt* in (2a), familiarity is not always necessary. Unlike Frey (2016) and Sudhoff (2016), I propose that a nominalization analysis is always required. The novelty of my analysis lies in linking the interpretive distinctions of clausal prolepsis to the different syntactic structures of the proleptic pronoun. Specifically, I propose that the proleptic pronoun uniformly represents a definite D-head across different contexts. The prolepsed CP always forms a DP constituent with the proleptic pronoun, with various structural components possibly occurring between D and CP. These components are responsible for the two interpretations that a prolepsed clause can have: familiarity and factivity. Previous literature has proposed different accounts for familiarity. Some argue that it arises through the definition of D (Kastner 2015, Bogal-Allbritten and Moulton 2018 i.a.), while others suggest the presence of an additional structural component, such as idxP, roughly corresponding to a pronoun (Elbourne 2005, Simonenko 2014, Schwarz 2009, 2019, Hanink 2021 i.a.). I adopt the latter view and specifically the implementation proposed by Hanink (2021), where idxP is merged in D's complement position, (4a). This structure corresponds to the cases of clausal prolepsis after non-factive predicates where the prolepsed clause is interpreted as familiar. An immediate advantage of this analysis is that besides accounting for familiarity, it allows for a unification of the various usages of the Dutch proleptic pronoun *het* under a single lexical entry. Turning to the cases where a prolepsed clause has a factive interpretation, it is important to note that factivity has received different analyses, including a silent noun FACT (Kiparsky and Kiparsky 1968, Kayne 2005 i.a.), a silent head FACT (Elbourne 2013), or a distinct C denotation (Kratzer 2006) among others. These analyses are all compatible with the core of the proposal that *het* is a definite

D forming a DP constituent with the embedded CP. In my analysis, I assume that a silent noun FACT is merged in D's complement position, (4b), primarily because it reinforces a structural parallelism with (a) the structure in (4a), where the CP is not merged with D directly but is further embedded instead under idxP, and (b) the independently attested structure [$_{DP}$ het$_{D}$ [$_{DP}$ feit$_{NP}$ [$_{CP}$ dat$_{C}$ …], where a CP is combined with the noun *feit* 'fact' and *het* 'the.'

(4) a. [$_{DP}$ D [$_{idxP}$ idx [$_{CP}$]]] b. [$_{DP}$ D [$_{NP}$ FACT$_{NP}$ [$_{CP}$]]]

The analysis presented in (4) proposes that clausal prolepsis is essentially a form of CP-nominalization, which is similar to cases of nominalized clauses found in Washo (Bochnak and Hanink 2021) and other languages. This suggests that nominalized clauses are more prevalent cross-linguistically than previously thought. However, clausal prolepsis differs from nominalized clauses in that the CP is extraposed, causing it to be separated from the proleptic proform on the surface. This proposed analysis of clausal prolepsis has several advantages. It explains which predicates allow clausal prolepsis, and associates familiarity with a structural component, namely, idxP, which is independently attested, as it is morphologically realized in some languages, such as Washo. Furthermore, it allows for the unification of different usages of *het* as anaphoric or non-anaphoric D or as a pronoun via a single lexical D-entry. The proposed analysis offers an account of various properties of Dutch clausal prolepsis. It accounts for the ability of oblique arguments to bind into a prolepsed clause, and the specific environments in which clausal prolepsis is allowed, namely, with nominalized infinitives but not with plain nominalizations. Finally, this proposal is important for theories of clausal complementation because it suggests that factivity exists independently of familiarity (as in Bochnak and Hanink 2021, and *pace* Kastner 2015) and that familiarity is not contributed by D (*pace* Kastner 2015, Bogal-Allbritten and Moulton 2018).

The paper is organized as follows. Beginning with [Section 2], its purpose is to document the various usages of *het*, as well as to provide evidence that *het* of clausal prolepsis is a contentful element. [Section 3] proposes a novel analysis according to which *het* of clausal prolepsis is the same D-element across its different usages. In clausal prolepsis, the prolepsed CP enters the derivation DP-internally, where it can be combined either with idxP or a silent FACT. [Section 4] is concerned with the surface position of the embedded CP in clausal prolepsis, and argues that it is the result of an extraposition movement step that CPs are independently known to undergo in Dutch. In [Section 5], I discuss two additional advantages of the proposed analysis. First, by treating clausal prolepsis as nominalized clauses, we can establish a clear link to the nominalized clauses of Washo and Korean, which, just like clausal prolepsis, convey a familiar interpretation. Second, the proposed analysis helps us understand why clausal prolepsis is allowed in certain syntactic contexts like nominalized infinitives, but not in nominalizations. [Section 6] discusses how the proposed analysis could be extended to account for variation in the D-forms used in the formation of clausal prolepsis cross-linguistically. [Section 7] concludes.

2 *Het* is not an expletive

In this section I discuss two different usages of *het* as both a definite determiner (Section 2.1) and pronoun (Section 2.2), highlighting the different meanings it can express in each case. Building on this background, I present three key pieces of evidence to suggest that *het* of clausal prolepsis has semantic content. Firstly, I demonstrate that, like pronominal *het*, *het* of clausal prolepsis can bind a parasitic gap (Section 2.3). Secondly, I establish that *het* of clausal prolepsis can only occur in syntactic positions where pronominal *het* is also possible (Section 2.4), further emphasizing the parallelism between the two usages. The properties of *het* discussed in Section 2.3 and 2.4 will be shown to hold consistently with clausal prolepsis following either a factive or non-factive predicate, suggesting that it is always a contentful element in clausal prolepsis. This conclusion challenges previous works, such as Sudhoff's (2016), which distinguishes between two structures of clausal prolepsis after factive and non-factive predicates based on whether *het* functions as a correlate/placeholder or a genuine contentful pronoun. An additional piece of evidence supporting the idea that *het* is contentful in clausal prolepsis is that in its presence, a prolepsed clause after non-factive predicates obtains an obligatory familiar interpretation requiring a contextual antecedent (Section 2.5).

2.1 *Het* as a definite determiner

To start, let us consider the use of *het* as a definite determiner. In this usage, *het* can combine with an NP to form a DP. Extensive previous literature has distinguished two interpretations in this case: a non-anaphoric unique interpretation and an anaphoric interpretation. To illustrate these interpretations, let us consider two examples taken from Schwarz (2009, 40):

(5) a. In dit dorp ligt het kerkhof naast een tankstation.
in this village lies the cemetery next to a gas station
'In this village the cemetery lies next to a gas station.'

b. Hans heeft een zonnetje$_i$ en een maan getekend. Het zonnetje$_i$ was blauw.
Hans has a small sun and a moon drawn the small sun was blue
'Hans has drawn a small sun and a moon. The small sun was blue.'

In (5a), *het kerkhof* 'the cemetery' is a non-anaphoric unique definite, indicating that a given town has a unique cemetery. Conversely, in (5b), *het zonnetje* is anaphoric to an antecedent, namely, *een zonnetje* 'a small sun,' introduced in the first clause, which leads to inter-sentential anaphora.

2.2 *Het* as a pronoun

Two meanings can be distinguished when *het* is used as a pronoun, propositional and individual-denoting. The propositional interpretation of *het* can be observed with verbs that require propositional arguments, such as *hopen* 'hope,' as shown in (6b). A plain DP such as *het antwoord* 'the answer' cannot be used as an argument of this verb, as demonstrated in (6a). Interestingly, *het* can serve as an argument of this verb, but only when referring to a proposition from the discourse, as illustrated in (6c) with the embedded clause from (6b).[2] In contrast to verbs like *hopen*, there are verbs such as *kennen* 'know' that can take a plain DP as an argument, as seen in (7a), and reject a propositional one, as in (7b). *Het* can also function as an argument of this verb, but, as shown in (7c), it can only refer to an individual denoting expression from the discourse, such as *het antwoord* 'the answer' in (7a).

(6) a. * Hij **hoopt** [het antwoord]$_m$.
he hopes the answer
Intended: 'He hopes for the answer.'

b. Hij **hoopt** [dat jij er bent]$_j$.
he hopes that you there are
'He hopes that you are there.'

c. Hij **hoopt** het$_{*m/j}$.
he hopes it
'He hopes so.'

(7) a. Hij **kent** [het antwoord]$_m$.
he knows the answer
'He knows the answer.'

b. * Hij **kent** [dat jij er bent]$_j$.
he knows that you there are
Intended: 'He knows that you are there.'

c. Hij **kent** het$_{m/*j}$.
he knows it
'He knows it.'

I propose that the interpretation of *het* as propositional or individual-denoting is determined by the selectional properties of the predicate it occurs with. Propositional *het* occurs as an argument of predicates like *hopen* that c-select a DP and s-select a proposition, while individual-denoting *het* occurs as an argument of predicates like *kennen* that c-select a DP and s-select an individual denoting expression.[3] Moving on to *het* of clausal prolepsis, I argue that it shares two properties with propositional *het*. Firstly, it can bind a parasitic gap, and secondly, it can occur in the same contexts as propositional *het*.

[2] For discussion on expressions like *het* or *something* in English, which can refer to propositions, see Elliott (2016). Elliott provides new arguments that these expressions are nominals (see also Moltmann 2013)

[3] The distribution of propositional *het* does not correlate with the semantic property of familiarity, which is often associated with verbs that presuppose the existence of their clausal complement as a proposition in the Common Ground, as discussed in Cattell (1978). For instance, the verb *hopen* 'hope' can take propositional *het* as an argument, even though its clausal complement does not necessarily have to be interpreted as familiar. This suggests that the distribution of *het* cannot be solely accounted for by appealing to a semantic property like familiarity.

2.3 Parasitic gaps

I present new data that show how propositional *het* and *het* of clausal prolepsis behave similarly in licensing a parasitic gap, based on Bennis (1986). I examine two verbs, *hopen* 'hope' and *betreuren* 'regret,' to provide a more complete picture of the distribution of *het* in this phenomenon than Bennis, who focused only on *betreuren*. To start, when merged as an argument of *hopen* or *betreuren*, propositional *het* undergoes scrambling, as do all unstressed pronouns. As shown in (8a), scrambled *het* precedes the adjunct clause containing the parasitic gap, and it can bind the gap. In this respect, propositional *het* behaves like *het* of clausal prolepsis with both *hopen* and *betreuren*, which also undergoes scrambling and can bind a parasitic gap, as shown in (8b) (= modified from Bennis 1986, (19a)).[4]

(8) a. Jan zei dat hij *(het$_i$) [na nogmaals e overwogen te hebben] toch t$_i$ hoopte/ betreurde.
John said that he it after again considered to have yet hoped regretted
'Jan said that he hoped/regretted it (after considering again).'

b. Jan zei dat hij *(het$_i$) [na nogmaals e overwogen te hebben] toch t$_i$ hoopte/ betreurde dat
John said that he it after again considered to have yet hoped that this
deze beslissing genomen was.
decision taken was
'Jan said that he hoped/regretted it, after considering again, that this decision had been made.'

To summarize, both *het* of clausal prolepsis and propositional *het* can bind parasitic gaps, and this ability is not affected by different verb classes. Since binding of a parasitic gap is expected of contentful elements, the common behavior of propositional *het* and *het* of clausal prolepsis illustrated in (8) suggests that they must both have semantic content.

2.4 The Prop-prolepsis generalization

I propose the following generalization: the syntactic contexts in which clausal prolepsis occurs are identical to those where propositional *het* can occur.

(9) **Prop-prolepsis generalization:** Clausal prolepsis can occur only in those contexts in which propositional *het* can occur.

This generalization is supported by the behavior of three types of predicates, distinguished based on whether they can take propositional *het* as an argument, whether they can take an embedded clause as an argument, and whether they can license clausal prolepsis. Type I predicates include verbs such as *hopen* 'hope,' a non-factive verb, and *beseffen* 'realize,' a factive verb. As shown in (10a) and (11a), both verbs can take *het* as an argument. *Het* is propositional here because it can have a proposition as its antecedent (see 6b). Additionally, both *hopen* and *beseffen* can take an embedded clause as an argument, as shown in (10b), (11b), and they can license clausal prolepsis, as illustrated in (10c) and (11c). This behavior is consistent with the Prop-prolepsis generalization, as clausal prolepsis is licensed in a syntactic context where propositional *het* is also allowed.

[4] It is worth noting that individual-denoting *het*, which is arguably a referential pronoun, can also bind a parasitic gap, just like propositional *het*. This is illustrated in the following example where *het* is an argument of *kennen* 'know' and refers to an individual denoting expression (see 7). As shown in the example, *het* can bind the gap from its scrambled position:

(i) Jan zei dat hij het$_i$ [na e grondig bestudeerd te hebben] wel t$_i$ kende.
Jan said that he it after thoroughly studied to have well knew
'John said that he knew it after having studied for long.'

(10) a. Ik hoop het.
I hope it
'I hope so.'

b. Ik hoop dat je wint.
I hope that you win
'I hope that you win.'

c. Ik hoop het dat je wint.
I hope it that you win
'I hope that you win.'

(11) a. Ik besef het.
I realize it
'I realize it.'

b. Ik besef dat ik te laat ben.
I realize that I too late am
'I realize that I am too late.'

c. Ik besef het dat ik te laat ben.
I realize it that I too late am
'I realize that I am too late.'

Type II consists of predicates such as *oordelen* 'judge' and *blij zijn* 'be happy.' They differ from *hopen* in that they cannot take propositional *het* as an argument, as shown in (12a) and (13a). However, they behave similarly to *hopen* in that they accept a clausal argument, illustrated in (12b) and (13b). Despite this, as demonstrated in (12c) and (13c), clausal prolepsis is not licensed with these verbs. The behavior of *oordelen* and *blij zijn* supports the Prop-prolepsis generalization because, as predicted by the generalization, clausal prolepsis is not allowed in a syntactic context where *het* is also blocked.

(12) a. * Hij oordeelde het.
he judged it
'He judged it.'

b. Hij oordeelde dat Jan slaapt.
he judged that John sleeps
'He judged that John sleeps.'

c. * Hij oordeelde het dat Jan slaapt.
he judged it that John sleeps
'He judged that John sleeps.'

(13) a. * Ik ben het blij.
I am it happy
Intended: 'I am happy about it.'

b. Ik ben blij dat Jan slaapt.
I am happy that John sleeps
'I am happy that John sleeps.'

c. * Ik ben het blij dat Jan slaapt.
I am it happy that John sleeps
'I am happy that John sleeps.'

The third type of predicate is exemplified by the non-factive verb *aandoen* 'do to' and the factive *beu zijn* 'be tired.' This type behaves like *hopen*, in that it can take propositional *het* as an argument, as illustrated in (14a) and (15a). Here, *het* is propositional and can have a proposition clause as its antecedent. However, *aandoen* and *beu zijn* cannot take a bare clausal argument, as shown in (14b) and (15b). Despite this, clausal prolepsis is permitted with this predicate type, as demonstrated in (14c) and (15c). The availability of clausal prolepsis in this syntactic context aligns with the Prop-prolepsis generalization, as it correlates with the availability of propositional *het*.

(14) a. Ik kan het$_i$ hem niet aandoen dus zal ik haar tegenhouden.
I can it him not do so will I her stop
'I cannot do it to him so I will stop her.'

b. ?* Ik kan hem niet aandoen dat ik hem nu in de steek laat.
I can him not do that I him now in the stab let
Intended 'I cannot do that to him, that is, to abandon him.'

c. Ik kan het hem niet aandoen dat ik hem nu in de steek laat.
I can it him not do that I him now in the stab let
'I cannot do that to him, that is, to abandon him.'

(15) a. Ik ben het$_i$ beu.
I am it tired
'I am tired of it.'

b. * Ik ben beu dat Jan slaapt.
I am tired that John sleeps
Intended: 'I am tired of the fact that John sleeps.'

c. Ik ben het beu dat Jan slaapt.
I am it tired that John sleeps
'I am tired of the fact that that John sleeps.'

The (un)availability of clausal prolepsis and propositional *het* after the three types of predicates examined is summarized in the table below:

	PropDP	*dat*-clause	Prolepsis
Type I: (hopen, beseffen)	✓	✓	✓
Type II: (oordelen, blij zijn)	✗	✓	✗
Type III: (aandoen, beu zijn)	✓	✗	✓

Table 1: The distribution of propositional *het* and clausal prolepsis.

Two conclusions can be drawn from the table above. First, clausal prolepsis is available in syntactic contexts where propositional *het*, is allowed to occur, as seen with Type I and Type III verbs. The fact that the distribution of *het* in clausal prolepsis resembles that of propositional *het*, which is arguably a contentful element reinforces the idea that the former carries semantic content. However, the table above also suggests that the licensing of a clausal argument alone is not enough to license clausal prolepsis, as observed with Type II verbs. This finding is particularly relevant for alternative analyses of clausal prolepsis, such as Stroik's (1996) analysis. In his analysis, *het* is an expletive pronoun that enters the derivation in Spec,CP of the embedded clause, and undergoes movement into the matrix clause for formal reasons, e.g. case (see also Gluckman 2021). Stroik's analysis predicts that any predicate that can embed a CP should also license clausal prolepsis. However, this prediction is not borne out, as demonstrated by the failure of predicates like *blij zijn* 'be happy' to license clausal prolepsis, despite being able to take a CP-complement. Moving on to the next section, we will demonstrate that a prolepsed CP doubled by *het* has to be interpreted anaphorically after certain predicates, just like propositional *het*, and requires a contextual antecedent. We take this to provide additional evidence supporting the idea that *het* must have semantic content.

2.5 The prolepsed clause and its meaning: Factivity and Familiarity

There are two distinct interpretations that a CP can have when combined with *het*: factivity and familiarity. The availability of these readings is determined by the matrix verb, as will be shown in this paper, and has been noted in previous works focusing on other languages (see Sudhoff 2016 and references therein for German). For non-factive predicates, a bare CP can be interpreted as either ±familiar, but when *het* is present, the doubled CP must be interpreted as familiar. Turning to factive predicates, a bare or prolepsed CP that occurs in their context is obligatorily interpreted as factive, and familiarity is only optional in such cases.

	CP	het+CP
Non-factive predicates	±Familiar,−Factive	+Familiar,−Factive
Factive predicates	±Familiar,+Factive	±Familiar,+Factive

Table 2: Familiarity and Factivity in clausal prolepsis.

Let us now move on to discussing the data. First, we will examine non-factive verbs, before turning our attention to factive verbs.

2.5.1 Non-factive verbs

In this section, I will present new data on non-factive verbs in (16-20) that show a difference between prolepsed embedded clauses and plain embedded clauses in terms of their interpretive properties. Specifically, I will show that prolepsed embedded clauses are obligatorily interpreted as familiar, while plain embedded clauses have an optional familiar interpretation. I will use new diagnostics/contexts from Bogal-Allbritten and Moulton (2018), which will allow us to arrive at a more precise understanding of the familiarity requirement that arises with clausal prolepsis

after non-factive predicates. To begin, I will present different contexts in (16) and (17) that demonstrate how both plain embedded clauses and prolepsed clauses can have a familiar interpretation. In (16), the context contains an assertion that is compatible with the proposition expressed by B in (16b) and (16c). For example, A asserts "I finished my homework," and in this discourse, both (16b) and (16c), illustrating a plain and a prolepsed clause respectively, are felicitous utterances of B.

(16) a. A: I finished my homework. Can I go outside and play?
B: No.
A: Don't you believe me?

b. B: Ja, ik geloof [dat je je huiswerk af hebt], maar het is etenstijd.
yes I believe that you your homework PRT have, but it is dinner.time
'Yes, I believe that you have done your homework, but it is dinner time.'

c. B: Ja, ik geloof het [dat je je huiswerk af hebt], maar het is etenstijd.
yes I believe it that you your homework PRT have, but it is dinner.time
'Yes, I believe it that you have done your homework, but it is dinner time.'

Similarly, (17) shows that a plain embedded clause or a prolepsed one can be licensed in a context where the proposition they express is not repeated from A's utterance but is entailed by it.

(17) a. *Context*: B has a rule that A must eat vegetables before having cake.
A: I ate peas. Can I have cake now?
B: No. A: Why? Don't you believe me?

b. B: Ja, ik geloof [dat je je groenten hebt gegeten], maar de cake is nog niet klaar.
yes I believe that you your vegetables have eaten but the cake is yet not ready
'I believe that you ate your vegetables, but the cake is not ready.'

c. B: Ja, ik geloof het [dat je je groenten hebt gegeten], maar de cake is nog niet klaar.
yes I believe it that you your vegetables have eaten but the cake is yet not ready
'I believe it that you ate your vegetables, but the cake is not ready.'

In what follows, I also show that the propositional content of the prolepsed clause must also be consistent with the content of a prior utterance introduced in the discourse. Specifically, in (18), the propositional content of the *dat*-clause is not consistent with the context created by the utterance in (18), where it falls within the scope of negation. In this case, clausal prolepsis cannot be employed, (18b). Plain embedded clauses, such as the one shown in (18a), do not have this consistency requirement and can be licensed in the same context.

(18) Pieter has certainly heard in his geography class that Toronto is not the capital of Canada…

a. Hoe dan ook, Pieter gelooft nog steeds dat Toronto de hoofdstad van Canada is.
even so Pieter believes still that Toronto the capital of Canada is
Even so, Pieter still believes that Toronto is the capital of Toronto.

b. # Hoe dan ook, Pieter gelooft het nog steeds dat Toronto de hoofdstad van Canada is.
even so Pieter believes it still that Toronto the capital of Canada is
Even so, Pieter still believes it that Toronto is the capital of Toronto.

When an embedded clause conveys new information, meaning that it is uttered in a discourse without an assertion that carries content comparable to the proposition expressed by B, a plain embedded clause is the only felicitous option, as shown in (19a). In contrast, note that a prolepsed clause is not allowed in this case, as demonstrated in (19b).

(19) Can Johny go outside and play?

a. Ja, ik geloof [dat hij zijn huiswerk af heeft].
yes I believe that he his homework PRT has
'Yes, I believe that he has done his homework.'

b. # Ja, ik geloof het [dat hij zijn huiswerk af heeft].
yes I believe it that he his homework PRT has
'Yes, I believe it that he has done his homework.'

Moving on to factivity, it is widely known that a bare CP after verbs like *geloven* 'believe' or *hopen* 'hope' is not interpreted as factive, which is why they are considered non-factive (see Kiparsky and Kiparsky 1968 i.a.). The presence of *het* does not change this, as illustrated in (20). Therefore, the speaker does not presuppose the truth of the embedded clause after *geloven* 'believe' whether it is bare, (20a), or doubled by *het*, (20b), as indicated by the fact that in both cases, its content can be negated through the continuation, *maar dat is niet waar* 'that is not true.'

(20) a. Jan geloofde dat de aarde gemaakt is van boerenkool, maar dat is niet waar.
John believed that the earth made is of kale but that is not true
'John believed that the earth is made of kale, but that it is not true.'

b. Jan geloofde het dat de aarde gemaakt is van boerenkool, maar dat is niet waar.
John believed it that the earth made is of kale but that is not true
'John believed that the earth is made of kale, but that is not true.'

To sum up, the presence of *het*, as demonstrated in (20), does not change the non-factive interpretation of a bare CP after predicates such as *geloven* 'believe' or *hopen* 'hope.' However, the fact that *het* obligates a familiar interpretation indicates that it must possess semantic content, consistent with the conclusions drawn in previous sections based on parasitic gaps and the Prop-prolepsis generalization.

2.5.2 Factive verbs

Factive verbs such as *betreuren* 'regret' behave similarly to non-factive ones, like *geloven* 'believe' in that when they are followed by a bare CP, this CP can be interpreted as familiar. In contrast to non-factive predicates, after a factive verb, a CP doubled by *het* can be interpreted as familiar but it does not have to. This is briefly demonstrated through examples like (21), although applying the same tests from the previous section would lead to the same conclusion. In this example, A's utterance in (21a) establishes the content of the embedded clause in the discourse. B's response shows in (21b) that the presence of *het* is only optional in this case, indicating that familiarity can be expressed either via a bare CP or clausal prolepsis, just like with non-factive verbs. However, clausal prolepsis with factive verbs is different because in this case, familiarity is only optional, unlike non-factive predicates where it is obligatory. Thus, as shown in (22), repeated from (2), a CP doubled by *het* can express new information and serve as an answer to an information-seeking question when it follows a factive verb.

(21) a. A: Ik betreur dat ik mijn huiswerk niet heb gemaakt.
I regret that I my homework not have done
'I regret that I have not done my homework.'

b. B: Ik betreur (het) ook dat ik het mijne niet heb gemaakt.
I regret it too that I my homework not have done
'I regret it too that I have not done my homework.'

(22) Wat is er aan de hand? 'What's the matter?'

a. Pieter betreurt (het) dat Marie WEGgaat.
Pieter regrets it that Marie goes-away.
'Peter regrets it that Peter becomes famous.'

b. Pieter zegt/ gelooft (* het) dat Marie WEGgaat.
Pieter says believes it that Marie goes-away
'Peter asserts it that Peter becomes famous.' modifed from Sudhoff (2016, (46))

Note that in Dutch, clause embedding predicates such as *betreuren* are factive, presupposing the truth of their clausal complement, whether bare or proleptically introduced. This is illustrated by (23), where the continuation *maar dat is niet waar* 'but that is not true' is not accepted with either plain or prolepsed embedded clauses.

(23) a. Jan betreurt dat Els gaat emigreren, # maar dat is niet waar.
Jan regrets that Els goes emigrate but that is not true
'Jan regrets that Els is going to emigrate, but that isn't true.' Broekhuis and Corver (2019, (86b))

b. Jan betreurt het dat Els gaat emigreren, # maar dat is niet waar.
Jan regrets it that Els goes emigrate but that is not true
'Jan regrets that Els is going to emigrate, but that isn't true.'

To summarize, we have examined the different meanings expressed by *het* as both a definite determiner and a pronoun. We have identified a clear parallelism between propositional *het* and *het* of clausal prolepsis, both in terms of their distribution, as demonstrated by the Prop-prolepsis generalization, and in their ability to bind a parasitic gap. These similarities suggest that *het* is always a contentful element in clausal prolepsis, rather than a mere placeholder (*pace* Sudhoff 2016). This conclusion is reinforced by the obligatory familiar interpretation that arises in the presence of *het* after non-factive predicates. However, with factive predicates, a familiar interpretation is only optional, raising questions about whether *het* is the same item in clausal prolepsis with factive predicates. In the following section, I will propose a unified analysis of *het* that accounts for its various usages as a definite determiner and a pronoun, as well as the interpretations associated with it in clausal prolepsis, under a single lexical entry. Before moving forward, however, a short note is in order about the relationship between factivity and familiarity, which has been proposed by Kastner (2015). According to Kastner, factive clauses not only presuppose truth but also express familiarity. However, this claim has been criticized in Bochnak and Hanink (2021), who argue that factivity and familiarity are independent properties (see also Sudhoff 2016 i.a.), as demonstrated by Washo nominalized clauses. This point is also illustrated in (22a), where the embedded clause is factive but does not have a familiar interpretation, as it can serve as an answer to an information-seeking question. Thus, Dutch provides additional evidence against the assimilation of factivity to familiarity (*pace* Kastner 2015)

3 Analysis

In this section, I will present an analysis that reconciles the properties of *het* with the properties of clausal prolepsis discussed in the previous sections. I will start by demonstrating that the different uses of *het* as a determiner and as a pronoun can be accounted for by a single D-lexical entry for *het* (Section 3.1). Then, I will provide an analysis that can explain the Prop-prolepsis generalization (Section 3.2). Finally, I will show that this account is also able to explain the familiar or factive interpretation of prolepsed clauses (Section 3.3).

3.1 A single D entry for *het*

Recall from section 2.1 that when used as a definite D, *het* can have two interpretations, a non-anaphoric unique interpretation and an anaphoric interpretation. The contexts in which these interpretations arise are repeated below from (5). In what follows I discuss how these can be analyzed under a single lexical entry of *het*.

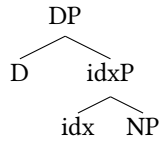
(24) a. In dit dorp ligt het kerkhof naast een tankstation.
in this village lies the cemetery next to a gas station
'In this village the cemetery lies next to a gas station.'

b. Hans heeft een zonnetje$_i$ en een maan getekend. Het zonnetje$_i$ was blauw.
Hans has a small sun and a moon drawn the small sun was blue
'Hans has drawn a small sun and a moon. The small sun was blue.'

Starting with the anaphoric interpretation, I follow Hanink (2021) in assuming that this arises due to a syntactically projected index, idx. This approach builds on previous proposals by Schwarz (2009, 2019) i.a. which suggest that anaphoric definites realize a distinct syntactic structure from non-anaphoric ones. Specifically, in Hanink's analysis, anaphoric definites are associated with the structure in (25), where idx heads a functional projection, idxP, within the extended projection of N (cf. also Elbourne 2005, Simonenko 2014). According to this view, idx is interpreted as a pronoun via the Traces and Pronouns Rule of Heim and Kratzer (1998), (27), which allows idx to be mapped

back to an antecedent via the assignment function. This results in the anaphoric interpretation illustrated in (24b). In contrast, non-anaphoric definites lack idx and realize a smaller syntactic structure in which D is merged directly with an NP complement, as shown in (26).

(25)



(26)

(27) **Traces and Pronouns Rule:**
If α is a pronoun or trace, g is variable assignment, and $i \in$ dom(g), then $[\![\alpha_i]\!] = g(i)$.

Hanink (2021) puts forward a unified lexical entry for D that encodes the meaning of an ι-operator, (28). Additionally, she assigns a property meaning to idx, as shown in (29), which denotes the property of being anaphoric. Specifically, idx has the meaning of an index that has undergone the IDENT type shift (Partee 1986). Thus, in (25), where idx takes the NP as its complement, idx functions as a modifier to the NP, and the two are semantically composed with, using Heim and Kratzer's (1998) Predicate Modification. After the application of Predicate Modification, idxP which is also predicate denoting is composed with D via Function Application.

(28) $[\![\text{D}]\!]$: $\lambda P_{<e,t>}{:}\exists!x(P(x)).\iota x_e[P(x)]$

(29) $[\![\text{idx}]\!]^g$: $\lambda y_e[y=g(i)]$

Returning to Dutch, I propose that both structures in (25) and (26) are available in the language. They are realized with *het* merging as a D-head. Under the non-anaphoric definite interpretation, *het* realizes the structure in (26), where it is a D-head and an NP, such as *kerkhof* in (24a), is merged in its complement position. On the other hand, when *het* functions as an anaphoric definite, it has the structure in (25), where the complement of D is the idxP which in turn takes an NP, such as *zonnetje* in (24b), as a complement. The question that arises next is how to reconcile the usage of *het* as a definite D with its pronominal usage where we saw that it can express two interpretations, individual-denoting and proposition-denoting. In order to address this, I discuss first the background I assume for pronouns. Building on previous work, I assume that pronouns are disguised DPs (e.g., Elbourne 2005, Postal 1969, Patel-Grosz and Grosz 2017 for German, Bi and Jenks 2019 for Mandarin and Hanink 2021 for Washo). In particular, I follow Hanink in assuming that pronouns are DPs that consist of an NP undergoing NP-deletion. Unlike Elbourne (2005), however, where the NP is merged in D's complement position, I assume as in Hanink, that the NP is merged as a complement of idx rather than D. Therefore, according to this view, pronouns realize the structure in (25). With this in mind, let us now return to Dutch. The individual-denoting *het*, as in (7), can be analyzed as a DP where D-head takes idxP as its complement, and idx takes in turn as its complement an NP that undergoes deletion. On the other hand, proposition-denoting *het*, as in (6), differs from individual-denoting *het* in that it can have a proposition or fact as its antecedent. In the case of a proposition antecedent, a CP is merged in idx's complement position, as in (30a), while in the case of a factive antecedent, the noun *feit* 'fact' and possibly a silent CP are merged in D's complement position, as shown in (30b). In both cases, the CP and the *feit* noun can remain silent, similar to the NP of the individual-denoting *het*.[5,6]

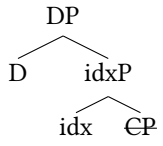
(30)

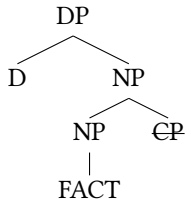
[5] See Collins (2015) for discussion of cases which involve CP-deletion, just like in the structure in (30a). Furthermore, as Elbourne (2013, 213) discusses, CP-deletion can be shown to be involved in instances such as the following ones:

(ii) Constantine discovered, and Diocletian regretted, that Christianity had become very popular among the Roman matrons.

(iii) It is not the suggestion that Europe is doomed that bothers me but the fact.

[6] The idea that the CP or the NP undergoes deletion opens up a number of possibilities. As shown, all options are attested except for options 3 and 7 below. I argue that 3 and 7 are not attested as a result of recoverability issues. Specifically, the D in these cases is non-anaphoric meaning that the content of the NP and the CP has not been introduced in the discourse. As a result of this, the NP and the CP in this case cannot be elided as a result of the fact that they do not have a contextual antecedent.

a. 

b. 

Elbourne (2013) has proposed a similar structure to (30a) for instances where the pronoun *it* refers to propositions. The only difference is that instead of an idxP, a silent noun *proposition* is merged in D's complement position in his analysis.[7] Leaving this difference aside, examples such as (31a) and (31b), which are constructed based on Elbourne's English data, offer supporting evidence for the existence of a CP within the syntactic structure realized by *het* in (30a).

(31) a. Domingo gelooft/ betreurt dat hij een genie is. Pavarotti gelooft/ betreurt het ook.
Domingo believes regrets that he a genius is Pavarotti believes regrets it too
'Domingo believes/regrets that he is a genius. Pavarotti believes/regrets it too.'

b. Domingo gelooft/ betreurt dat hij een genie is en alle andere beroemde tenoren geloven/ betreuren het ook.
Domingo believes regrets that he a genius is and every other famous tenor believe regret it too
'Domingo believes/regrets that he is a genius and every other famous tenor believes/regrets it too.'

The sentence in (31a) is ambiguous. Thus, it could mean either that Pavarotti believes that Domingo is a genius (strict interpretation) or that Pavarotti believes that he, Pavarotti, is a genius (sloppy interpretation). There are analyses in which the sloppy interpretation involves binding. Under these analyses, for binding to apply, there must be a hidden CP that is syntactically represented together with *het* in (31a). Yet, as Elbourne notes, there is a lot of controversy around the issue of whether sloppy readings require binding, so the most convincing case for the idea that there is a syntactically represented CP is the one in (31b). This example too is ambiguous: every other famous tenor could be believing that Domingo is a genius or that he himself is a genius. In this case, as Elbourne notes, '[…] it is pretty uncontroversial that binding must be involved for the second (i.e. sloppy) reading, since the antecedent is a quantifier phrase. So any theory of the syntax-semantics interface that postulates that binding relations are established in the syntax will have to say that the word *it* in his (78), (*het* in 31b), is the only audible reflex of some more extensive structure, some structure that includes a pronoun for every other famous tenor to bind.'

3.2 Clausal Prolepsis

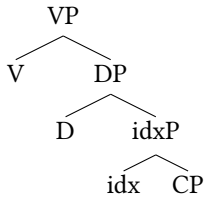
Moving on to clausal prolepsis, I propose that it can be realized through two distinct syntactic structures. The first structure, illustrated in (32a), involves merger of idxP with the CP, forming an idxP, which is then merged into the D's complement position. The second structure, (32b), is very similar, but instead of idxP, FACT is merged in D's complement position. Despite this difference, both structures have in common that the CP enters the derivation DP-internally.

(32)

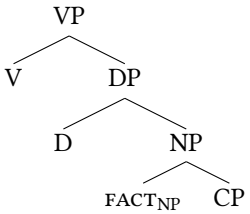
(iv)
1. $D_{non-anaphoric}$+overt NP=unique definite DPs
2. $D_{anaphoric}$+overt NP=anaphoric definite DPs
3. $D_{non-anaphoric}$+covert NP=?
4. $D_{anaphoric}$+covert NP=individual *het*
5. $D_{non-anaphoric}$+overt CP=clausal prolepsis with factive predicates
6. $D_{anaphoric}$+overt CP=clausal prolepsis with non-factive predicates
7. $D_{non-anaphoric}$+covert CP=?
8. $D_{anaphoric}$+covert CP=propositional *het*

[7] It is conceivable that, in place of idxP, Elbourne's *proposition* (see also Moulton 2020 on content nouns) may be used in (32a). This noun undergoes deletion, an operation that is only licensed in the presence of a suitable antecedent. Both analytical possibilities, idxP and *proposition*, are theoretically equivalent and consistent with the idea that clausal prolepsis involves nominalization of the prolepsed CP.

a.



b.



The two DPs in (32) are merged in the complement position of the verb in (32), and they both express a proposition. The key difference between the two DPs in (32a) and (32b) is that the CP in (32b) is factive, meaning that it expresses a true proposition due to the presence of FACT.[8] On the other hand, the CP in (32a) is interpreted as a plain proposition, which can be true or false. As both DPs in (32) express a proposition and have the categorical status of a DP, they correspond to propositional *het*. This implies that clausal prolepsis involves the use of propositional *het*. Consequently, the Prop-prolepsis generalization, which asserts that clausal prolepsis is permissible in all contexts where propositional *het* can be used, can be readily accounted for. Specifically, as argued in Section 2.2, propositional *het* is licensed with predicates that c-select a DP and s-select a proposition. The DPs in (32) satisfy these selectional requirements, as they are interpreted as propositions and have the categorical status of a DP. To further elaborate on how the Prop-prolepsis generalization can be accounted for, we can examine how the types of predicates that allow for clausal prolepsis can be mapped to one of the two syntactic structures in (32). They are repeated below from Table (1):

	PropDP	*dat*-clause	Prolepsis
Type I: (hopen, beseffen)	✓	✓	✓
Type II: (oordelen, blij zijn)	✗	✓	✗
Type III: (aandoen, beu zijn)	✓	✗	✓

Table 3: The distribution of propositional *het* and clausal prolepsis.

Type I and Type III predicates, such as *beseffen* and *beu zijn* c-select a DP, as shown by the fact that they can take *het*, a prop DP, as their argument, and s-select a fact, as shown by the fact that the truth of the embedded clauses following these verbs is presupposed. Therefore, when clausal prolepsis is licensed with these predicates, they take as a complement the syntactic structure of the DP shown in (32b). Type I and Type III predicates, such as *hopen* and *aandoen*, c-select a DP and s-select a proposition, so clausal prolepsis is licensed with them under the structure in (32a). However, Type II predicates cannot realize any of the syntactic structures in (32), because even though they s-select a proposition, they do not select a DP, as shown by the fact that they cannot take a PropDP as their argument. Before proceeding, it is worth noting that the DP in (32b) has a very similar structure to that of a non-anaphoric definite DP, as seen in (26), in the sense that an NP is merged in D's complement position. This suggests that, just like with anaphoric definite DPs, idxP can be merged on top of NP, i.e., [$_{VP}$ V [$_{DP}$ D [$_{idxP}$ idx [$_{NP}$ FACT$_{NP}$ CP]]]], and give rise to a familiar interpretation. Additionally, since the NP in (32b) is FACT, the DP should convey both a factive interpretation and a familiar one. This prediction is borne out in cases like (21), where the embedded clause receives

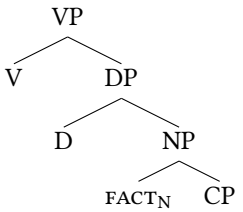
[8] The idea that a silent noun FACT exists in factive clauses was first explored by Kiparsky and Kiparsky (1968) and has since been incorporated into various syntactic (Aboh 2005, Collins 1994, Kayne 2005 i.a.) and semantic analyses of factivity (Elbourne 2013). I have chosen to adopt FACT because it allows for a parallel between the structure we propose for clausal prolepsis after non-factive predicates (see 32a), and an existing structure, i.e. *het feit dat …* 'the fact that…,' where FACT instead of idxP is merged in D's complement position. Within analyses like the current one taking *facts* to be concrete objects (see though Grimm and McNally 2022), they can be argued to have the meaning illustrated below, for instance, as has been suggested by Kratzer (2002, 660) within situation semantics (see also Moltmann 2013 on *facts* as modal objects):

(v) If *s* is a possible situation and *p* a proposition, then *s* is a fact exemplifying *p* iff for all *s'* such that *s'* ≤ s and *p* is not true in *s'*, there is an *s"* such that *s'* ≤ *s"* ≤ *s* and s" is a minimal situation in which *p* is true.

Besides FACT, various alternatives exist for factivity, including a distinct C denotation (see Kratzer 2006), or alternatives taking factivity to arise from constraints imposed by certain verbs on their arguments (Ozyildiz 2017, Bondarenko 2020, see also Djarv 2019). Fundamentally, the core tenet of the proposed analysis that clausal prolepsis involves nominalization of a clause can align with these alternative analyses.

a factive interpretation contributed by FACT and a familiar interpretation contributed by idxP. Before proceeding with the exact mechanism by which familiarity arises through idxP, a short note is in order about Elbourne's (2013) analysis of clausal prolepsis. It is discussed at this point because it bears a striking similarity to my proposed structure in (32b). His proposal is that clausal prolepsis realizes the syntactic structure illustrated in (33) where a head realized by the silent noun FACT takes the prolepsed CP as its complement:

(33)



A crucial distinction between my proposed structure and (33) has to do with the syntactic position of the prolepsed CP. In (33), the CP is a complement of FACT, while in (32b), it is an adjunct. However, (33) is untenable because, as was demonstrated in Moulton (2014), *fact* is not an argument structure noun, and thus, can never take any internal arguments. Thus, in constructions like *the fact that the Earth is flat*, the prevailing assumption, established since Stowell (1981), is that the CP is merged as an adjunct to the NP. This argument is reinforced by Kayne (2019), who posits that *fact* cannot take a clause or any XP as its complement, primarily due to the inherent inability of nouns to take complements. An additional difference between the proposed analysis and Elbourne's is that he does not distinguish an additional structure for clausal prolepsis. His idea is that clausal prolepsis in English is only licensed with factive predicates, thus, only one structure comprising a silent FACT noun is needed, as the one in (33). This idea is based on data from Kiparsky and Kiparsky (1968), who argue that clausal prolepsis is only allowed with factive predicates like *resent, regret* etc. whereas it is disallowed with non-factive (or semi-factive) predicates *know, believe, think, see* etc. However, Kiparsky and Kiparsky's observation is empirically incorrect, and thus, the proposal that clausal prolepsis realizes a single syntactic structure in English because it is only allowed with factive predicates cannot be maintained. This can be demonstrated by the many instances of clausal prolepsis in English discussed in previous literature showing that non-factive predicates can license clausal prolepsis. This is illustrated below with the verb *see*, or with counterfactual predicates, like *fake* (Gentens 2016). As noted in Gentens (2016), in (34a), *see* is used to express a personal opinion and its clausal complement does not have to be true. Likewise, clausal prolepsis after counterfactual predicates like *fake* entail the falsity of the clause they are followed, as in (34b), which as a result provides additional evidence that clausal prolepsis in English is not necessarily linked to factivity.

(34) a. Druidism or druidry is really a way of working with the natural world and connecting to one's roots in a particular way. We see it that part of the problems that humanity is facing is that it has become separated from nature and alienated from nature.

b. Gwen Stefani, 34, fibbed that she played the piccolo to get into her school band-but preferred sex games with her drummer lover instead. She said: 'I faked it that I played. I spent most of the time in the drum room making out with my boyfriend.' Gentens 2016, (31),(33)

3.3 Familiarity

To fully capture familiarity in clausal prolepsis, it is crucial to establish background assumptions about the interpretation of plain embedded clauses. To achieve this, I will be drawing upon previous works, such as Kratzer (2006), Moulton (2009, 2015), and Elliott (2016). These works argue that embedded clauses, particularly those that follow attitude predicates like *believe* and are introduced by the complementizer *that* in English, denote <e,<s,t>> and interpreted as predicates of individuals with propositional content, as shown in (35a). In particular, building on these previous works, I contend that Dutch embedded clauses introduced with *dat*, both plain and proleptic, are interpreted as predicates of individuals with propositional content. Under this view, an embedded proposition is linked to its content via a function CONT$_w$, which maps an individual x and a world w to sets of worlds that are compatible with the content of x, as illustrated in (35b) (Moulton 2015, 312). Moreover, a proposition is turned into properties of individuals via a functional head F$_{PROP}$, which has the denotation shown in (36a), and is introduced in

the left periphery. Specifically, F_{PROP} has the denotation of returning a predicate of individuals whose content is the proposition denoted by the clause, as expressed in (36b).

(35) a. $[\![\textit{that Bob is a fraud}]\!]^w = \lambda x.\text{CONT}_w(x) = \lambda w'.$Bob is a fraud in w'

b. $\text{CONT}_w(x) = \{w' : w' \text{ is compatible with the intentional content determined by x in } w\}$

(36) a. $[\![F_{PROP}]\!]^w = \lambda p_{<s,t>} \lambda x_e [\text{CONT}_w(x) = p]$ Moulton (2015), Elliott (2016)

b. $[\![F_{PROP}[CP]]\!]^w = \lambda x_e [\text{CONT}_w(x) = [\![CP]\!]]$

In Moulton (2015), the idea that clauses can function as predicates of individuals with propositional content is supported by examples like *the claim that the earth is round* where a content noun and a CP are combined. Moulton argues that content nouns lack argument structure and, therefore, cannot take a CP as an argument. Instead, he suggests that CPs stand in a modification relationship with content nouns, and the two are semantically composed via Heim and Kratzer's (1998) Predicate Modification. This operation is restricted to apply to two expressions that share the same semantic type. Moulton proposes that content nouns denote sets of individuals with propositional content. Such nouns can modify CPs since the latter also denote predicates of individuals with propositional content, enabling the combination of the two through Predicate Modification. Turning to Dutch, content nouns like *idee* 'idea' can be modified by *dat*-embedded clauses, indicating that these clauses are modifiers of content nouns.[9]

In clausal prolepsis, a *dat*-clause also denotes a predicate of individuals. However, due to the nominal structure merged on top of the CP, it is turned into type *e*. In (37), the prolepsed clause is semantically composed as shown in (38). The CP enters the derivation in the complement position of idx. The LF in (38) shows that idx and CP are semantically composed via Predicate Modification, just like idx and NP in anaphoric definites. idxP denotes a property, and so it is combined with D, realized by *het* in Dutch, via Function Application. The resulting DP is an individual-denoting expression that saturates the argument position of the matrix verb via Function Application. Following Bochnak and Hanink's analysis of familiarity in Washo nominalized clauses, I assume that the prolepsed clause in (38) is interpreted as familiar because idx maps the index (1, in 38) to the salient individual from the discourse whose content is compatible with the proposition expressed by the prolepsed clause.[10]

(37) Ik hoop het$_i$ [dat je wint]$_i$.
I hope it that you win
'I hope that you win.'

[9] Another indication supporting the proposal that Dutch embedded clauses denote denote <e,<s,t>> is found in copular constructions, such as (vib). Drawing from Potts (2002), Moulton argues that such constructions are equative in nature, meaning that what the DP subject refers to, the CP also describes. Given that content nouns denote predicates of individuals with propositional content, the fact that *dat*-CPs can be used in copular constructions like (vib) indicates that they also denote such predicates.

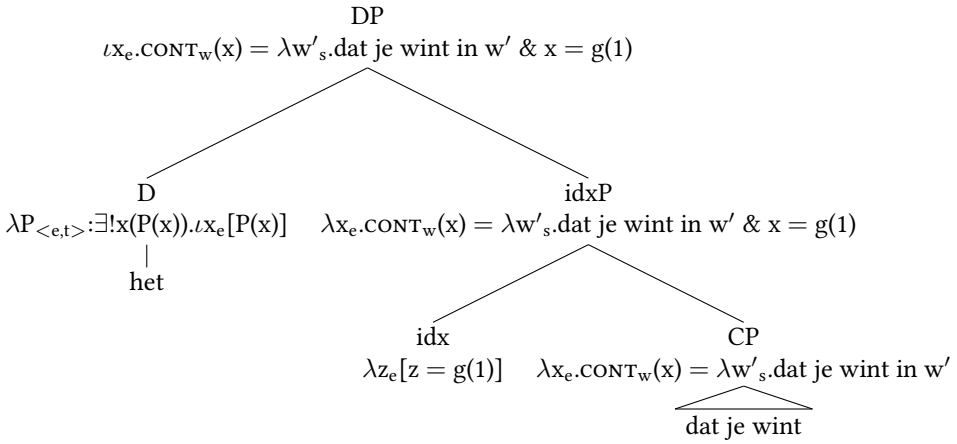
(vi) a. het idee dat zij gauw zou komen
the idea that she soon would come
'the idea that she would come soon'

b. het idee is dat zij gauw zal komen
the idea is that she soon would come
'the idea is that she would come soon'

Note, however, that as Keir Moulton (p.c.) pointed out to me, the data above are not strongly suggestive of the idea that embedded clauses are predicates, as the fact that the construction in (vib) is equative does not really prove that the CP is predicative.

[10] One question that arises is why it is not possible to directly combine D with the CP in (38) to create clausal prolepsis without familiarity. This may be related to case, assuming that Ds can only be combined with case-marked elements, and CPs lack case. Thus, only NP or idxP, which can possibly have a case feature can combine with Ds. An open question is whether this restriction that blocks merging D with CP is universal, or it only applies to Dutch (see Picallo 2002 and Moulton 2020 for cases in Spanish that have been analyzed as D+CP).

(38)

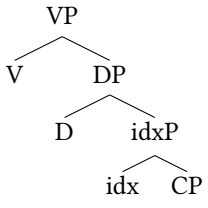


In summary, a DP-analysis of Dutch clausal prolepsis can account for several of its properties, including the fact that it distributes like propositional *het* and the resulting familiar interpretation.[11] Before proceeding, it is important to acknowledge that there are alternative views proposing that familiarity is instead conveyed through a definite D, rather than through idxP (Kastner 2015, Bogal-Allbritten and Moulton 2018 i.a.). This would suggest that familiarity is contributed by the presence of the definite D *het*. However, empirical evidence does not support this claim, as clausal prolepsis after factive verbs involves *het*, yet the prolepsed clause lacks familiarity (see Section 2.5.2). Therefore, it is more plausible to argue that familiarity is dissociated from the presence of D and instead arises through nominal components such as idxP embedded under D. This view has been independently supported, e.g. in Bochnak and Hanink (2021) and Schwarz (2009).

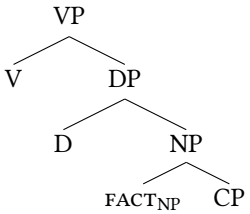
4 Extraposition

In the previous sections, I showed that a DP-analysis of clausal prolepsis according to which the prolepsed clause realizes a nominalized clause underlyingly, just as shown below, repeated from (32), can account for several properties of *het* and the prolepsed clause.

(39) a.



b.



[11] Familiarity can also be expressed in other ways in Dutch, such as through a plain CP. This was shown in (16), where a familiar CP can be expressed bare or with *het*. The fact that a bare CP can express familiarity is that it can follow predicates such as *agree*, which can only take familiar clausal complements (Kastner 2015 i.a.). The question that arises is how does familiarity arise with plain CPs. To better explain the distinction between plain and prolepsed CPs with familiar interpretation, I tentatively propose that both types of CPs involve merger of idxP, but in different positions. For plain CPs, idxP is merged in the left periphery of the clause, specifically in Spec,CP. This allows for the expression of familiarity in the absence of a DP. On the other hand, in clausal prolepsis, idx takes the CP as its complement, as previously discussed. The two possible syntactic positions in which idxP can be merged are illustrated in (vii). The advantage of the structure in (viib) is its ability to account for the distribution pattern of plain familiar clauses, which resemble CPs rather than DPs. This is evident from the fact that such clauses can e.g. never surface after prepositions, as is also noted by Jarvis (2021), who looks at English plain familiar CPs. This is in contrast to Kastner (2015), who considers all familiar clauses to be DPs.

(vii) a. $[_{DP}$ D $[_{idxP}$ idx $[_{CP}$]]]

b. $[_{CP}$ idxP $[_{C'}$ C]]

In the proposed syntactic structure, *het* and the prolepsed clause form a constituent. However, when we examine clausal prolepsis in an embedded context, as in (40), it is apparent that *het* and the prolepsed clause surface as separate units. This suggests that they do not form a constituent on the surface.

(40) Jan zei dat hij het$_i$ toch hoopte [dat deze beslissing genomen was]$_i$.
John said that he it yet hoped that this decision taken was
'Jan said that he hoped it that this decision had been made.'

Resolving the tension between the underlying structures in (39) and the surface order in (40), I propose that the CP in Dutch undergoes extraposition out of the DP.[12] This extraposition step is obligatory, as shown by the fact that central embedding of a CP is ruled out in clausal prolepsis, (41).[13]

(41) a. Jan heeft het altijd gehoopt dat Karl beroemd zou worden.
Jan has it always hoped that Karl famous would become
'Jan has always hoped that Karl would become famous.'

b. * Jan heeft het dat Karl beroemd zou worden altijd gehoopt.
Jan has it that Karl famous would become always hoped
'Jan has always hoped that Karl would become famous.'

I assume that the obligatory extraposition step of the CP in clausal prolepsis can be attributed to a commonly accepted principle in Dutch, according to which CPs must undergo extraposition (see Appendix for an explanation of why extraposition cannot be attributed to phonological rules). The most recent analyses of extraposition involve either leftward (Hinterholzl 1999, Kayne 2005, Koopman and Szabolcsi 2000 and Moulton 2015 i.a.) or rightward (Büring and Hartmann 1995, Bruening 2018) movement. I will not take a stand on the correct analysis of extraposition, but provide evidence that the CP moves to a higher position after starting out low in a DP-internal position.[14,15] The evidence comes from binding and distributional properties. I will start with the binding properties.

---

[12] Alternative analyses of extraposition according to which the prolepsed CP is never moved out of the DP, but instead is moved to a DP-adjunct position, and stay there, as in Elbourne (2013), cannot be maintained simply because the proleptic pronoun and the prolepsed CP do not form a surface constituent, as shown in (40).

[13] German permits the central placement of a CP linked to a pronoun, as demonstrated in (viiib). The structure in which the CP is placed at the end of the sentence is also acceptable, as seen in (viiia). The acceptability of (viiib) is explained by the observation that CPs in German are not always required to be extraposed, as brought to my attention by Martin Salzmann (p.c.) and supported by Frey (2016). Conversely, in Dutch, central embedding of the CP is independently not allowed, especially with non-factive verbs, thus, blocking (viiib), as shown in (41b).

(viii) a. Maria hat es stets behauptet dass Karl beruühmt ist.
Maria has it always claimed that Karl famous is
'Maria has always claimed that Karl is famous.'

b. Maria hat (es), dass Karl beruühmt ist, stets behauptet.
Maria has it that Karl famous is always claimed
'Maria has always claimed that Karl is famous.'

However, as noted by Frey (2016), (viiib) '[…] is only possible if the associated phrase (the prolepsed CP in our terms) is clearly marked prosodically as an apposition.' However, apposition is distinct from clausal prolepsis: (a) The intonation pattern in apposition differs from this in clausal prolepsis. In clausal prolepsis, the characteristic intonation break that separates the proleptic pronoun and the prolepsed CP in (viiib) is unnecessary. (b) The structures categorized as 'apposition' are usually assumed to be categorized into two distinct classes of phenomena, none of which are not related to clausal prolepsis: '*namely*'-constructions, which represent typical examples of apposition, and reduced relatives (see McCawley 1998). Consequently, the acceptability of (viiib) and its distinct properties, such as that it is incompatible with factive predicates, fall outside the scope of this paper, as they pertain to the syntax and semantics of apposition rather than clausal prolepsis. Nevertheless, the fact that (viiib) is only accepted as an apposition, not clausal prolepsis, is important: it suggests that just as in Dutch, extraposition of the prolepsed CP is obligatory in German clausal prolepsis too, and that when central embedding of a CP linked to a pronoun is possible, as in (viiib), it is only allowed under a different structure, such as apposition.

[14] See also Zwart (1993) for a non-movement analysis of CP-extraposition, and Moulton (2015) and fn.17 for criticism.

[15] CP movement does not affect the semantic composition of idx in (38) because the CP undergoes total reconstruction in its base position. Note that the CP cannot be interpreted in its moved position, as the lower copy would be turned into type *e* via Trace Conversion (Fox 2002, Takahashi 2010, Moulton 2015). This conflicts with the proposed analysis, as idx denotes a property of <e,t>, and if Trace Conversion applies, idxP will denote type *t* after being semantically composed with the CP-trace, making the LF not able to converge, because D cannot be semantically composed with type *t* expressions:

4.1 The prolepsed clause and its binding properties

By employing two diagnostics—pronominal binding and Condition C—I provide evidence that supports the proposed analysis, indicating that the prolepsed CP is interpreted like a plain argument clause, in the verb's argument position. Firstly, let us examine pronominal binding.

(42) a. Ik heb niet op een beleefde manier ook maar één student$_i$ verteld [dat hij$_i$ positief getest had].
I have not on a polite manner any student told that he positive tested had
'I have not told in a polite manner to any student that he had tested positive.'

b. Ik heb het niet op een beleefde manier ook maar één student$_i$ verteld [dat hij$_i$ positief getest had].
I have it not on a polite manner any student told that he positive tested had
'I have not told in a polite manner to any student that he had tested positive.'

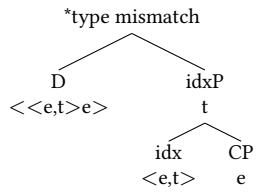
In (42a), it is shown that a pronoun found within the embedded clause, such as *hij* 'he' in this case, can be bound by the indirect object QP in the matrix clause, *ook maar één student* 'any student.' Two observations are worth noting about the QP: firstly, it occupies a low syntactic position within the vP, which is evidenced by its position after the manner adverbial PP *op een beleefde manier* 'in a polite manner'—a standardly accepted modifier of the vP. Secondly, the quantifier in this instance scopes lower than negation. The syntactic position of the indirect object QP, combined with its scope properties, suggest that it binds the pronoun in the embedded clause from a low position, possibly from the indirect argument position where the QP is introduced. (42b) illustrates that *ook maar één student* can also bind *hij* within a prolepsed clause. This implies that a prolepsed clause and a plain embedded clause occupy the same syntactic position at some point during the derivation, from which binding by the indirect object QP becomes possible. This is further supported by Condition C, as shown in (43a), which precludes coreference between a proper name, *Jan*, in the embedded clause and a pronoun, *hem* 'him,' in the matrix clause. Similarly, coreference is ruled out between a proper name in a prolepsed clause and a pronoun in the matrix clause, as seen in (43b).

(43) a. Ik heb hem$_{j/*i}$ verteld/ toevertrouwd [dat Jan$_i$ ziek is].
I have him told confided that Jan sick is
'I have told/confided him that Jan is sick.'

b. Ik heb het hem$_{j/*i}$ verteld/ toevertrouwd [dat Jan$_i$ ziek is].
I have het him told/ confided that Jan sick is
'I have told/confided it him that Jan is sick.'

Based on standard assumptions, the evidence from pronominal binding and Condition C[16] indicates that at some point in the derivation, the prolepsed clause must be within the c-command domain of the low position where the indirect object is introduced. So, where exactly is this position? According to prevailing literature, indirect objects enter the derivation within the specifier of an applicative head, which is situated higher than the VP, but lower than the External Merge position of the external argument, i.e., vP (cf. Pylkkänen 2008). Therefore, it is reasonable to assume that, just like plain embedded clauses, the prolepsed clause enters the derivation in the VP—in the internal argument position of V, which is lower than the applicative head introducing indirect objects. This is consistent with the proposed analysis according to which the prolepsed CP enters the derivation at a very low level, specifically, within the DP occupying the verb's argument position, exactly like it is suggested by the binding facts.

(ix)



[16] See though Bruening (2014) for discussion on why c-command might not be the relevant notion for Condition C, and Zwart (2015) for a critical response.

4.2 VP-fronting

I offer additional evidence to support the proposal that the prolepsed CP moves out of the DP, this time from VP-fronting in clausal prolepsis. To begin, I provide an illustration of a typical instance of VP-fronting in (44b), where the VP is moved from its underlying position in (44a) to a clause-initial position:

(44) a. Jan zal niet toegeven dat het probleem nu opgelost is.
Jan will not admit that the problem now solved is
'Jan will not admit that problem is now solved.'

b. [Toegeven] zal Jan niet dat het probleem nu opgelost is.
admit will Jan not that the problem now solved is
'Jan will not admit that problem is now solved.'

Although the term "VP-fronting" is used to describe the type of fronting shown in (44b), it is worth noting that it can target constituents that are larger than a VP. An example of this is shown in (45), where the fronted constituent includes not only a verb but also a scrambled object DP, *boeken* 'books.' The fact that the object has undergone scrambling is evidenced by its position before the adverb *meermaals* 'repeatedly.'

(45) [Boeken meermaals lezen] doet hij niet.
books repeatedly read does he not
'He does not repeatedly read books.'

Keeping this in mind, let us now investigate VP-fronting with clausal arguments, using the verb *beloven* 'promise' as an example. As shown in (46a), *beloven* can take an embedded clause as its argument. Furthermore, (46b) and (46c) demonstrate that *beloven* can undergo fronting either alone or together with the embedded clause.

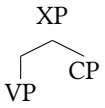
(46) a. Jan wil niet beloven [dat hij komt].
Jan wants not promise that he comes
'Jan doesn't want to promise that he will come.'

b. [Beloven] wil hij niet [dat hij komt].
promise wants he not that he comes
'Jan doesn't want to promise that he will come.'

c. [Beloven [dat hij komt]] wil hij niet .
promise that he comes wants he not
'Jan doesn't want to promise that he will come.'

Based on the analysis of Broekhuis and Corver (2019), I assume that the verb *beloven* undergoes fronting via VP-movement. Given this assumption, the fact that *beloven* can be fronted without the embedded clause, as illustrated in (46b), suggests that CPs move to a syntactic position outside the VP, as shown in (47).[17] This suggests that the VP, which becomes a constituent after the CP has been extraposed, can undergo fronting. Additionally, (46c) reveals that VP-fronting can target a larger constituent, shown as XP below, which contains the CP in its moved position.

(47)



The verbs *beloven* and *betreuren* can both license clausal prolepsis, as exemplified in (48a). Both verbs yield identical patterns of VP-fronting, confirming the proposal that clausal prolepsis after factive and non-factive verbs creates a uniform syntactic structure in which the CP is introduced embedded under a DP. In (48a), *het* occupies its typical

[17] This speaks against Zwart (1993) according to which embedded clauses stays in-situ, that is, in the complement position of the verb. If that were the case, VP-fronting should not be allowed to take place without the embedded clause, contrary to fact.

position above negation, and the verb can undergo fronting on its own, as in (46b). This suggests that both the prolepsed clause and *het* are situated outside the VP, permitting the VP to move, as in (48b). According to the proposed analysis, this implies that the CP enters the derivation inside the DP headed by *het* and is then moved to a position outside the VP. Conversely, (48c) demonstrates that the verb and the prolepsed clause cannot be fronted. The most plausible explanation for the ungrammaticality of (48c) is that the verb and the prolepsed clause do not constitute a single constituent.[18]

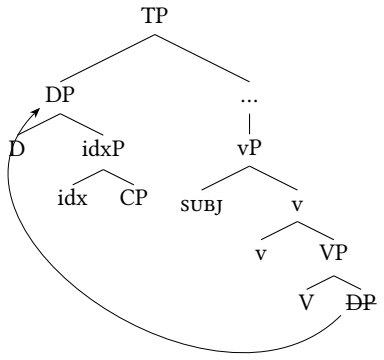
(48) a. Hij wil het niet beloven/ betreuren [dat hij komt].
he wants it not promise/ regret that he comes
'He doesn't want to promise/regret it that he will come.'

b. ? [Beloven/ betreuren] wil hij het niet [dat hij komt].
promise regret wants he it not that he comes
'Jan doesn't want to promise/regret it that he will come.'

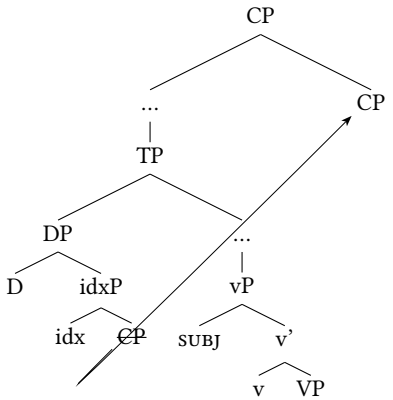
c. * [Beloven/ betreuren [dat hij komt]] wil hij het niet.
promise regret that he comes wants he it not
'Jan doesn't want to promise/regret it that he will come.'

The observed VP-fronting patterns in (48) can be explained by the proposed analysis, appealing to constituency. Initially, the CP enters the derivation in the verb's complement position embedded within a DP that is headed by *het*. In Dutch, unstressed pronouns like *het* undergo movement to a higher position. Assuming that Agree is necessary for movement (see Chomsky 2000), I propose that the DP headed by *het* must enter into an Agree relationship with a functional head such as T or v. Pronouns like *het* usually appear in the T-area, and so, they must first undergo Agree with T before moving to the T-area. Following Rackowski and Richards (2005), Van Urk and Richards (2015), Halpert (2019), I assume that Agree with a phase head is a necessary step for extraction out of that phase. I take Ds to be phase heads (Chomsky 2008 i.a.) so after D is agreed with T, the DP moves into the T-area, as shown in (49a). The CP subsequently Agrees with a higher probe in the C-area and can thus, be moved out of the DP-phase into a higher position, which corresponds to the CPs' final extraposed position. Now because the probe that is responsible for the movement of the CP is higher than the one that is responsible for *het*'s movement, the CP undergoes its final movement step after *het* has moved to its surface position, as shown in (49b). As a result, the CP occupies a higher position than *het*.[19,20]

(49) a.



b.



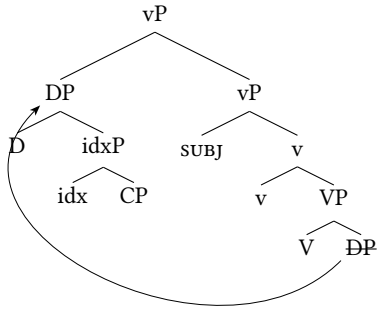
[18] In addition to *het*, the demonstrative *dat* 'that' can also be employed in Dutch clausal prolepsis, a phenomenon that has been observed in other languages as well, such as German (Sudhoff 2016 and references therein). Importantly, the exact same VP-fronting patterns observed with *het* in (48) and (51) are also replicated with *dat*. The fact that clausal prolepsis formed with *het* and *dat* leads to the same VP-fronting patterns suggests that they can be analyzed identically.

[19] Fischer (2018) shows that in German, the subject of a matrix clause can control into a prolepsed non-finite CP doubled by *es*. Assuming that control can only be licensed by a c-commanding controller, a structure in which the doubled CP is base generated in the CP-area, and the subject is in Spec,CP or lower is precluded because PRO inside the doubled CP is not c-commanded by the subject.

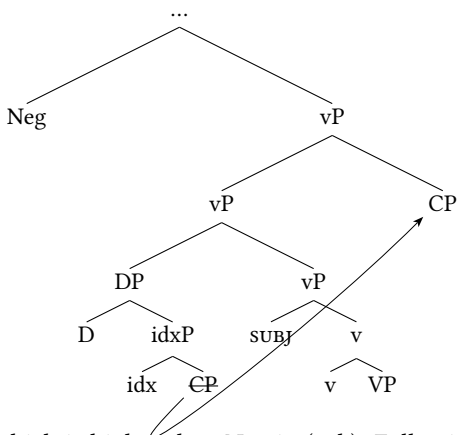
[20] Building on Sportiche (2016), we could hypothesize that the different components of a copy can be interpreted at distinct heights. Thus, the pronominal component of a DP, D and idxP, can only be interpreted in their moved position. This accounts for an independently known, even though not fully understood fact about pronouns, namely, that they cannot reconstruct. In contrast, CPs can freely reconstruct (see previous section), thus, it makes sense to assume that they can be interpreted either in the higher DP copy or the lower one.

With these structures in mind, let us return to the facts in (48). Since the VP forms a constituent after the movement of the DP and CP outside of it, the grammaticality of (48b) follows. However, after the CP is moved to its final position above the TP, the verb and the CP no longer form a constituent, as shown in (49b), which explains why the two cannot be fronted together, (48c).[21] Now let us assume that v might also serve as a probe that can Agree with D, and thus, attract the DP realized by *het* in the v-area, as shown in (50a). Furthermore, as in (49b), let us assume that there is a higher probe that is responsible for the movement of the CP from inside the DP. This probe is situated higher than the vP, but lower than negation, as shown in (50b), and attracts the CP from within the DP in its moved position.

(50) a.



b.



In (50b), the subject is subsequently moved from Spec,vP to Spec,TP, which is higher than Neg in (50b). Following this movement step, two constituents are formed: (a) *het* and the verb form a constituent, (b) *het*, the verb and the prolepsed CP form a constituent to the exclusion of negation. These predictions are confirmed, indicating that the structure in (50) is allowed. This is evident as the verb and *het* can be fronted as seen in (51a), the verb, *het*, and the proleptic CP can also be fronted, as demonstrated in (51b). Notably, if *het* had moved to a position higher than negation, neither type of fronting in (51) would be permitted.

(51) a. [Het [beloven/ betreuren]] wil hij niet [dat hij komt].
it promise regret wants he not that he comes
'Jan doesn't want to promise/regret it that he will come.'

b. [Het [beloven/ betreuren [dat hij komt]]] wil hij niet.
it promise regret that he comes wants he not
'Jan doesn't want to promise/regret it that he will come.'

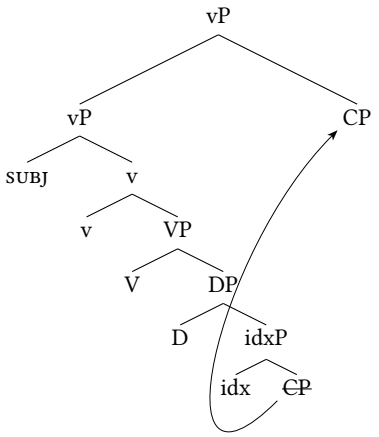
Essential to our analysis is the assumption that the prolepsed CP can only be extracted from the DP realized by *het* after the latter has established an Agree relation with a higher probe, resulting in scrambling to a higher position.

[21] Building on Abels (2007), I argue that movement of the CP out of the DP does not violate the so-called freezing constraint. Abels proposes that movement from moved constituents is subject to constraints according to a hierarchy of the types of movement involved. For instance, some A-bar moved constituents may block *wh*-movement, while A-moved constituents permit it (see also Müller 1996). Based on this, let us assume that the movement step the CP undergoes out of the scrambled DP in (49b) has the properties of topicalization. Given this, it is allowed to take place out of the DP because, as shown in Abels (2007), topicalization is allowed out of a scrambled object. On the other hand, note that the CP blocks *wh*-movement out of it, as demonstrated in (x). This follows from the idea that CP's movement has the properties of topicalization, as *wh*-movement is blocked out of Topicalized constituents (Abels 2007). On the other hand, plain CPs allow extraction out of them, as shown in (x), even though they are also extraposed because *wh*-movement is possible out of the type of movement they undergo, possibly A-movement (Moulton 2015), but not necessarily, given that CPs in Dutch or German obligatorily reconstruct, as pointed out to me by Martin Salzmann (p.c.).

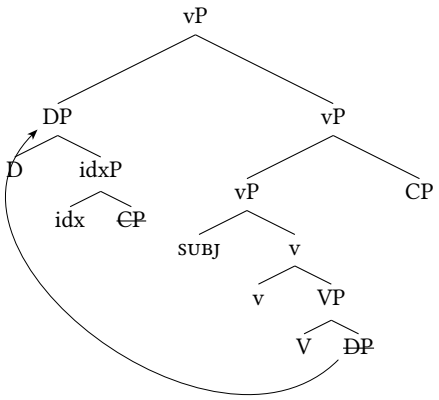
(x) Wat betreurde/ bevestigde/ zei jij (* het) dat hij gezegd had?
what regretted confirmed said you it that he said had
'What did you regret/confirm/say that he has said?'

This precludes an alternative structure, as illustrated in (52a). In this alternative structure, the CP initially moves from the DP to a vP adjunct position, before scrambling of *het*, and thus, before it is Agreed with by a higher probe, as shown in (52b). This derivation is almost the same as the structure proposed by Bennis (1986) for Dutch clausal prolepsis, with one key distinction: the CP is moved to the vP-adjunct position in (52b), whereas in Bennis's work, the CP is base generated in that position. The structure in (52) and that of Bennis are both precluded, as predicted, because *het* and the verb do not form a constituent in it, and thus, the type of fronting seen in (51a) is not possible to derive.[22]

(52) a.



b.



Before proceeding with the advantages of the proposed analysis, a short note is in order about a different analysis of clausal prolepsis.[23] Specifically, I would like to focus on the bi-clausal analysis of Ott and De Vries (2016), and show issues it faces in light of (51b). According to this analysis, the verb and the pronoun, shown as pro below, are contained in CP1 of the structure below. The embedded clause is hosted in a different CP, CP2, and undergoes movement into the left periphery where it is interpreted as a topic. The structure below the CP undergoes PF-deletion, resulting in the surface order. However, this analysis conflicts with (51b) because it predicts that the verb, *het* and the prolepsed CP, which is contained in a different CP can never form a constituent, which is not the case in the surface order, as was shown in (51b).

(53) a. $[_{CP1}$ . . .V pro . . .] $[_{CP2}$ CP_i $[. . . t_i . . .]]$ → PF-deletion

b. $[_{CP1}$ . . .V pro . . .] $[_{CP2}$ CP_i ~~$[. . . t_i . . .]$~~$]$

[22] In previous literature, it has been noted that topicalization of an embedded clause is not possible with *het*, (xib) (see Broekhuis and Corver 2019, Frey 2016 and references therein for German). Based on Frey (2016), who looks at the corresponding construction in German, I propose that the ungrammaticality of (xib) can be attributed to cyclic linearization (Fox and Pesetsky 2005). The CP enters the derivation inside the DP and is subsequently moved out of the DP, without, however, moving to the phase-edge (see Rackowski and Richards 2005 as to why extraction does not occur through a phase edge after a phase has been Agreed). As a result, the linear order within the DP is established as *het*>CP, with *het* linearly preceding the CP. This linear order is maintained in clausal prolepsis (xia), but not in topicalization (xib), resulting in ungrammaticality. A similar explanation can be offered for the fact that in N+CP constructions of the *the rumor that the earth is flat* type, not shown here, the CP cannot be moved leftwards stranding the DP, e.g. *the rumor*. The fact that clausal prolepsis exhibits the same constraint as N+CP constructions, where the D, noun, and CP form a constituent underlyingly, lends further support to the idea that clausal prolepsis is formed in a comparable manner, with the D and CP forming a constituent.

(xi) a. Jan vindt het$_i$ leuk [dat Marie morgen komt]$_{.i}$.
Jan considers it nice that Maria tomorrow comes
'Jan considers it nice that Marie will come tomorrow.'

b. [Dat Marie morgen komt]$_i$ vindt Jan (* het)$_i$ leuk.
that Maria tomorrow comes considers Jan het nice
'That Maria is sick, John considers it nice.'

[23] See Angelopoulos (2022) for a detailed discussion of alternative analyses of clausal prolepsis and how they fare with respect to the one proposed here.

In conclusion, based on the VP-fronting facts, it can be concluded that both the pronoun *het* and the embedded clause in clausal prolepsis occupy a position outside the VP, and that the embedded clause occupies a position higher than *het*. However, since the prolepsed clause enters the derivation in a lower position inside the matrix VP, the only possible way for it to occupy to a higher syntactic position is through movement. Therefore, the VP-fronting facts provide further evidence supporting the idea that the prolepsed CP undergoes movement outside the DP during the derivation.[24]

5 Additional advantages of the proposed analysis

The proposed analysis has the advantage of explaining two other properties of clausal prolepsis involving *het*, as I will demonstrate in this section. Firstly, by treating clausal prolepsis as nominalized clauses, a direct connection can be made to languages with nominalized clauses, such as Washo or Korean, which, as shown, also have a familiar interpretation. Secondly, the analysis can capture the fact that clausal prolepsis is permitted in nominalized infinitives but not in nominalizations. I will begin by discussing nominalized clauses.

5.1 Familiarity and nominalized clauses

The proposed analysis of clausal prolepsis suggests that it is an instance of CP-nominalization which is similar to the nominalized clauses found in Washo, for which the same structure has been proposed in Bochnak and Hanink (2021). Washo nominalized clauses consist of an embedded clause that is nominalized through the attachment of *-gi/-ge*. The distribution of this suffix in definite expressions, such as in (55a) where it functions as a personal pronoun, or in (55b) where it appears with a demonstrative, indicates its nominal character.

(54) [ø-haʔaš-ay-i-š-ge] di-hamup'ay-i
3-RAIN-INT.PAST-IND-DS-NM.ACC 1/3-forget-IND
'I forgot that it rained.'

(55) a. gí: pélew ʔíʔwi
3.PRO.NOM jackrabbit 3/3-eat.TR-IND
'He's eating the jackrabbit.'

b. hádi-gi pélew Mú:biʔ-i
DIST-gi jackrabbit 3.run-IND
'That jackrabbit ran.'

Bochnak and Hanink propose that Washo nominalized clauses are interpreted as familiar due to the nominal structure they project. They also argue that the correlation between the "nouny-ness" of the complement and the presuppositional requirement of familiarity is supported by Bogal-Allbritten and Moulton (2018), who build on Kim (2009) and propose a notion of familiarity implicated in nominalized clauses in Korean. In what follows, I focus on Korean nominalized clauses, highlighting the striking interpretive similarities they present with prolepsed clauses, specifically

[24] Even though this paper is about object clausal prolepsis, it is worth noting that subject clausal prolepsis also exists in Dutch. For example, in (xii), *het* occupies the subject position. These cases present a challenge, as it remains an open question where exactly *het* and the CP might enter the derivation in such constructions—whether in the verb's external or internal argument position, or whether *het* is assigned a theta-role (see Ruys 2010 for discussion). Assuming that *het* in (xii) is not an expletive, and is, thus, selected by *schijnen* 'seem,' the proposed analysis can be extended to account for subject clausal prolepsis. Specifically, under the proposed analysis, the CP enters the derivation inside the DP headed by *het*. In subject clausal prolepsis, the position of the DP realized by *het* is always a derived one, as *het* is moved to Spec,TP. This movement step is prefigured by an Agree relationship into which the DP enters with T. This Agree relation unlocks the phase, thus, permitting movement of the CP out of the DP into its extraposed position, just like in object clausal prolepsis.

(xii) Het schijnt dat dieven er tegenaan lopen.
it seems that thieves it into run
'It seems that thieves run into it.'

after non-factive predicates. To start with, to illustrate the contexts in which nominalized clauses are licensed, Bogal-Allbritten and Moulton (2018) conduct a systematic comparison between nominalized clauses and plain embedded clauses in Korean. The latter are formed with the complementizer *ko* whereas the first are formed with *kes*. The first context examined by Bogal-Allbritten and Moulton is presented in (56), and it served as the basis for constructing the corresponding context in Dutch. In this context, A makes an assertion that conveys the proposition expressed by B in (56b) and (56c), which is that A finished their homework. Bogal-Allbritten and Moulton demonstrate that both a plain embedded clause, as in (56b), and a nominalized clause, as in (56c), are permissible in B's utterances. This is similar to the case in Dutch, where a plain embedded clause or a proleptic clause can be used in the same context, as was shown in (16).

(56) a. A: I finished my homework. Can I go outside and play?
B: No.
A: Don't you believe me?

b. B: Um. Na-nun [ney-ka swukecey-lul ta ha-yess-ta-**ko**] mit-e. Haciman cikum-un cenyek
Yes I-TOP you-NOM homework-ACC all do-PST-DEC-ko believe-DEC but now-TOP evening
siksa sikan-i-ya.
meal time-COP.DEC
'Yes, I believe that you finished your homework, but it is dinner time.'

c. B: Um. Na-nun [ney-ka swukecey-lul ta ha-yess-ta-nun **kes**-ul] mit-e. Haciman cikum-un
Yes I-TOP you-NOM homework-ACC all do-PST-DEC-ADN kes-ACC believe-DEC but now-TOP
cenyek siksa sikan-i-ya.
evening meal time-COP.DEC
'Yes, I believe (the claim) that you finished your homework, but it is dinner time.'

Similarly, Bogal-Allbritten and Moulton note that both plain embedded and nominalized clauses are licensed in a context like (57), where the proposition expressed by these clauses is entailed by A's utterance. As we saw in (17), in this context, plain embedded and prolepsed clauses are also allowed in Dutch.

(57) a. *Context*: B has a rule that A must eat vegetables before having cake.
A: I ate peas. Can I have cake now?
B: No. A: Why? Don't you believe me?

b. Na-nun [ney-ka yachae-lul mek-ess-ta-ko] mit-e…
I-TOP you-NOM vegetable-ACC eat-PST-DEC-ko believe-DEC
'I believe that you ate vegetables (…but the cake's not ready).'

c. Na-nun [ney-ka yachae-lul mek-ess-ta-nun kes-ul] mit-e…
I-TOP you-NOM vegetable-ACC eat-PST-DEC-ADN kes-ACC believe-DEC
'I believe (the claim) that you ate vegetables (…but the cake's not ready).'

On the other hand, Bogal-Allbritten and Moulton show that when an embedded clause conveys new information, only a plain embedded clause formed with *ko* can be used, as in (58a), while a nominalized clause cannot, (58b). Proleptic clauses in Dutch pattern similarly to Korean nominalized clauses in that they cannot convey new information.

(58) Can Johny go outside and play?

a. Um. Na-nun [kay-ka swukecey-lul ta ha-yess-ta-ko] mit-e.
Yes I-TOP he-NOM homework-ACC all do-PST-DEC-ko believe-DEC
'Yes, I believe that he finished his homework.'

b. # Um. Na-nun [kay-ka swukecey-lul ta ha-yess-ta-nun kes-ul] mit-e.
Yes I-TOP he-NOM homework-ACC all do-PST-DEC-ADN kes-ACC believe-DEC
'Yes, I believe (the claim) that he finished his homework.'

In summary, there is an intriguing correlation between Korean nominalized clauses and prolepsed clauses occurring after non-factive predicates, as they are permitted in identical contexts. This provides further evidence for the

hypothesis that prolepsed clauses are initially formed as nominalized clauses, and that the movement during the derivation separates the embedded clause from *het*.[25]

5.2 Clausal prolepsis with nominalizations

Under the proposed analysis, the prolepsed CP needs to move outside the DP that is headed by *het*. However, as noted, this CP-movement step can only happen if there is a prior Agree relationship between D and a higher probe, which unlocks the DP phase. Based on this analysis, we predict that if a construction does not allow scrambling, then clausal prolepsis should not be possible because D is not Agreed with a probe, therefore, the DP is not transparent for movement (of the prolepsed CP) out of it. This prediction is confirmed by nominalizations, which do not permit scrambling, suggesting that there is no higher probe available for D to Agree with. On the other hand, nominalized infinitives, which have more structure and allow for scrambling, do permit clausal prolepsis, as predicted. Before we continue, it is important to note that the restriction on clausal prolepsis in nominalizations is not unique to Dutch; it is also found in other languages. In particular, Postal and Pullum (1988) were the first to show that English nominalizations behave the same, as demonstrated in (59). Given this, the proposed analysis can be extended to account for the English facts as well.

(59) a. her resentment of it (*that he won)
b. your demonstration of it to him (*that she was sick) Postal and Pullum (1988, (6c-d))

5.2.1 Nominalizations: Clausal prolepsis and scrambling

As previously discussed, I suggest that the reason why nominalizations do not permit clausal prolepsis is that they lack a probe that could Agree with the D-head realized by *het*, and thereby unlock the phase to license CP-movement from within the DP. In the following, I will provide an example to illustrate this point using (a) the nominalization *ontkenning* 'denial', which consists of the root *ontken* and the nominalizer *-ing*, and takes PPs as arguments, and (b) the nominal *hoop* 'the hope,' which like *ontkenning* takes PPs as arguments. One possible PP may contain *het*, which

[25] Note that the act of assertion is required for the felicity of a nominalized clause. This is meant to capture the behavior of nominalized clauses in contexts like (xiii). In this context, it is shown that a polar question is not sufficient to license B's utterance of a nominalized clause, despite the fact that the proposition expressed by the nominalized clause is string identical to the proposition contained in A's polar question:

(xiii) Has John finished his homework?
a. Um. Na-nun [kay-ka swukecey-lul ta ha-yess-ta-ko] mit-e.
Yes I-TOP he-NOM homework-ACC all do-PST-DEC-ko believe-DEC
'Yes, I believe that he finished his homework.'
b. # Um. Na-nun [kay-ka swukecey-lul ta ha-yess-ta-nun kes-ul] mit-e.
Yes I-TOP he-NOM homework-ACC all do-PST-DEC-ADN kes-ACC believe-DEC
'Yes, I believe (the claim) that he finished his homework.'

An important difference between clausal prolepsis in Dutch and German (Keir Moulton p.c. pointed out to me that German speakers present conflicting judgments for examples xiv) is that the act of assertion is not required (cf. Schwabe et al. 2016). It remains unclear to me at this point where this difference comes from.

(xiv) Has John finished his homework?
a. Ja, ik geloof [dat hij zijn huiswerk af heeft].
yes I believe that he his homework PRT has
'Yes, I believe that he has done his homework.'
b. Ja, ik geloof het [dat hij zijn huiswerk af heeft].
yes I believe it that he his homework PRT has
'Yes, I believe it that he has done his homework.'

is realized as *er* next to a P. However, as shown in (60a) and (60b), the proleptic *er* cannot be linked to an embedded clause, indicating that nominals like *ontkenning* or *hoop* do not permit clausal prolepsis.[26]

(60) a. De ontkenning er$_i$van (* [dat Jan ziek is]$_i$).
the denial it-of that Jan sick is
Intended: 'The denial of it that Jan is sick.'

b. De hoop erop (* [dat de oorlog zal stoppen]).
the hope it-for that the war will stop
Intended: 'The hope for it that the war will stop.'

With this in mind, I will proceed with scrambling and show that it is not allowed in nominalizations, in contrast to verbs. Consider the verb *vragen* 'ask', which can take a PP as an argument that may contain the pronoun *er*. As illustrated in (61), *er* can either remain in-situ within the PP, (61a), or undergo scrambling to a higher position, (61b).

(61) a. ? Jan zei dat ze nooit er-naar vraagt.
John said that she never it-for asks
'John said that she never asks for it.'

b. Jan zei dat ze er nooit naar vraagt.
John said that she it never for asks
'John said that she never asks for it.'

In contrast, when *er* is hosted in a PP that serves as an argument of the nominal like *ontkenning* or *hoop*, it is unable to undergo scrambling. This is demonstrated by the contrast between (62a), (63a) and (62b), (63b). While *er* can remain inside the PP in (62a), (63a), it cannot be scrambled to a prenominal position as in (62b), (63b).

(62) a. De ontkenning er-van.
the denial it-of
'The denial of it.'

b. * De er ontkenning van.
the it denial of
'The denial of it.'

(63) a. De hoop er-op.
the denial it-of
'The denial of it.'

b. * De er hoop op.
the it denial of
'The denial of it.'

Given this observation, it is reasonable to assume that nominalizations disallow clausal prolepsis because lacking a probe that can Agree with D, license scrambling of *het* and unlock the DP, the prolepsed CP cannot be extracted out of the DP where it enters the derivation. In contrast, as we will see in the next section, nominalized infinitives behave differently from nominalizations as they can license scrambling, and crucially, clausal prolepsis.[27]

[26] Note that the extraction site of the extraposed clause might also be inside a DP which is further embedded under a P, just as in the example below. This suggests that the ungrammaticality of (60a) is not because the PP is an island configuration blocking movement of the CP.

(xv) Omdat Jan over zijn hoop sprak om Marie ooit terug te zien.
because John of his hope talked for Marie ever back to see
'because Jan talked about his hope to see Marie again.'

[27] Unlike nominals, a proleptic pronoun, e.g. *er*, hosted within a PP functioning as a verb's argument, has the capacity to license clausal prolepsis, as demonstrated below:

(xvi) Abraham vertrouwde er-op dat God hem zou brengen in de stad.
Abraham trusted it-for that God him would bring in the city
Intended: 'Abraham trusted that God would bring him in the city.'

In (xvi), the prolepsed CP enters the derivation within the DP, following the proposed analysis. It is later extracted from the DP after *er* establishes an Agree relationship with a higher probe. Notably, in (xvi), even though *er* has been Agreed with by a higher probe, making the DP transparent to CP movement, *er* remains unscrambled. This aligns with the observation that when a proleptic pronoun is inside a PP that acts as a verb argument, the pronoun only optionally undergoes scrambling, as demonstrated in (61).

5.2.2 Nominalized infinitives: Clausal prolepsis and scrambling

I will start by providing some background on the formation of Dutch infinitives. They are typically formed with the suffix *-en* and exhibit nominal properties, as evidenced by their ability to combine with an overt determiner (see Looyenga 1992 and Reuland 2011). An example of such combination is shown in (64) below, where the demonstrative pronoun *dat* combines with an infinitive. Moreover, unlike nominalizations such as *ontkenning*, nominalized infinitives are structurally richer, as evidenced by their ability to permit scrambling. In example (64), scrambling of *het* is responsible for its position before *compleet*.[28]

(64) Dat vervelende (het) compleet (* het) ontkennen.
that annoying it completely it deny
'That annoying and completely denying of it.'

In contrast to nominalizations, nominalized infinitives permit clausal prolepsis, as shown in (65a). This can be accounted for by the proposed analysis, according to which scrambling is allowed because the DP Agrees with a higher probe, which also unlocks the D-phase, permitting CP movement. The fact that *het* can undergo scrambling in nominalized infinitives is shown in (65b) where *het* obligatorily surfaces before the adverb *compleet* 'completely.'

(65) a. Dat eeuwige (het) ontkennen [dat Jan schuldig is].
that eternal it deny that John guilty is
'that eternal denying that John is guilty.'

b. ? Dat vervelende (het) compleet (* het) ontkennen [dat Jan schuldig is].
that annoying it completely it deny that John guilty is
'the annoying denying completely that John is guilty.'

In conclusion, the analysis presented suggests that the availability of clausal prolepsis is contingent upon scrambling of *het*. Nominalized infinitives allow for prolepsis because they permit scrambling of *het*, which in turn allows for movement of the CP out of the CP where it enters the derivation. In contrast, nominalizations do not permit scrambling of *het*, and as a result, the CP cannot move out of the DP. Therefore, clausal prolepsis is not available in nominalizations of Dutch.[29]

6 Cross-linguistic variation: Pro-form choice

In Dutch, we observed that clausal prolepsis involves the use of *het*. However, in other languages that allow for clausal prolepsis, different elements are used to be linked to a clause. For example, in German, *es* (66a) is used, which has a pronominal usage but is not used as a definite D. In Hungarian, *azt*, (66b) is used, which also has a demonstrative usage (Brandtler and Molnár 2016). This cross-linguistic variation in the use of different elements for

[28] In current approaches to morphology (cf. Alexiadou and Borer 2020 for an overview), this contrast between nominalizations like *ontkenning* and nominalized infinitives like *dat ontkennen* follows as a result of differences in the height of attachment of *-ing* and *-en*. The first only attaches to a low position, which is hierarchically lower than, and thus, excludes the higher syntactic projections licensing *compleet*. On the other hand, infinitives are structurally richer, and thus, allow modification by *compleet* because the morpheme *-en* of infinitives attaches to a syntactic position which is structurally higher than the one licensing *compleet*.

[29] As pointed out to me by Keir Moulton (p.c.) and Martin Salzmann (p.c.), there appears to be a distinction between Dutch and English on the one hand and German on the other. Concretely, whereas there is a restriction, as we saw, in the first two blocking clausal prolepsis in nominalizations, German allows such instances (see Müller 2023), as shown below:

(xvii) die Zustimmung da-zu/ daf-ür$_i$), [dass wir fertig sind]$_i$.
the agreement there-to/ there.for that we finished are
'the agreement that we are finished' Müller (2023, (50b))

However, as noted in fn.13, German differs from Dutch and English in that it permits central embedding of a prolepsed CP in which case the CP is not extraposed. Based on this, I assume that (xvii) realizes the structure in (viiib) of fn.13. This structure realizes apposition, as discussed, and so it is allowed in nominalizations because, in contrast to the structure realizing clausal prolepsis, the CP does need not be extraposed, and thus, no higher probe is required.

clausal prolepsis is similar to what we observe in nominalized clauses. For instance, in Greek, nominalized clauses are formed with a definite determiner (Roussou 1991, Faure 2022), as shown in (67a), just like Dutch clausal prolepsis. In contrast, nominalized clauses in Persian use a demonstrative (Farudi 2007) (67b), similar to clausal prolepsis in Hungarian.

(66) a. Max bedauert es, dass Lea krank ist.
Max regrets it that Lea ill is
'Max regrets it that Lea is ill.'

b. Péter (azt) mondta, hogy gyakran találkoznak nunka után.
Peter that said that often gather work after
'Peter said that they often meet up after work.' Brandtler and Molnár (2016, (1))

(67) a. To oti efighe apodhiknii tin enohi tis.
the that left proves the guilt her
'The fact that she left proves her guilt.'

b. In ke Maryam raft ma'alum e.
this that Maryam left clear is
'The fact that Mariam left is clear.'

Under the proposed analysis, the observed variation could be reduced to the relation of D with CP at the underlying syntactic derivation and morphological restrictions on Ds. For instance, to account for the possible morphological restrictions, we can assume that CPs are inherently different from NPs, lacking certain features like person or ϕ-features in general (see Iatridou and Embick 1997). Given this, it could be that German *es* is an allomorph of the definite determiners *der, die, das*, which are used when the D is followed by NPs with different ϕ-features. In contrast, *es* could serve as the elsewhere realization or the realization of a D-head when it is followed by XPs, such as CPs, that lack ϕ-features. In Dutch, there are two definite articles *de* and *het*, but clausal prolepsis is only possible with the latter, as illustrated below:

(68) a. Ik hoop het$_i$ [dat je wint]$_i$.
I hope it that you win
'I hope that you win.'

b. * Ik hoop de$_i$ [dat je wint]$_i$.
I hope the that you win
'I hope that you win.'

De is used with feminine and masculine nouns, while *het* is used with neuter nouns (Broekhuis and den Dikken 2012 i.a.). Since neuter is considered the absence of gender (Bennis 1995, among others), and complement clauses lack gender, it is expected that prolepsis in Dutch is formed with *het* rather than *de*. In current morphological theories with a realizational component, such as Distributed Morphology (Halle and Marantz 1993), the realization of an element relative to its context is sensitive to syntactic locality (Embick 2003). This means that only local elements can condition the morphological realization of another element. Given this, an advantage of the proposed analysis is that the determiner and the complement clause are local to each other in the underlying syntactic derivation. In other words, the complement clause can act as the licensing environment for the realization of the determiner, which in Dutch is realized as *het* rather than *de*, because the two are syntactically local to each other in the underlying derivation, as shown in (69), repeated from (32). It is worth noting that since IDXP or FACT are silent, they do not count as interveners (following Embick's 2003 idea of Pruning).

(69) a. [$_{DP}$ D [$_{idxP}$ idx [$_{CP}$]]] b. [$_{DP}$ D [$_{NP}$ FACT$_{NP}$ [$_{CP}$]]]

While more research is necessary to investigate the morphological properties of a language's D-system and how this system relates to the form of D in prolepsis or nominalized clauses, the proposed analysis offers a promising way to account for the variation in the form that D can take when it is linked to a CP. This analysis respects locality in morphology, which is an important factor to consider when explaining how elements are realized relative to their syntactic context.

7 Conclusion

This paper presented novel findings suggesting that all instances of clausal prolepsis are formed using the same D-head. The prolepsed CP is introduced DP-internally, and the different interpretations of the clause arise from various intervening structural components, such as idxP or FACT, between D and the CP. During the derivation, the CP is extraposed or moved from its base position to a higher position than that of the proleptic pro-form. This unified analysis of clausal prolepsis avoids accidental homophony and also accounts for different VP-fronting patterns and properties of the construction, such as that the prolepsed CP receives a familiar interpretation after non-factive predicates, or that clausal prolepsis cannot occur with nominalizations, but is allowed with nominalized infinitives. Several theoretical conclusions were shown to follow from the proposed analysis such as that factivity must be distinguished from familiarity, or that familiarity arises through nominal structure rather than D. Empirically, the proposed analysis suggests that nominalization of a clause is more common cross-linguistically than previously thought. Lastly, by assuming a nominalization structure, the proposed analysis showed the potential to account for cross-linguistic variation in the choice of pro-forms used in clausal prolepsis.

A CP-extraposition

In Section 4, I argued that CP-extraposition is triggered in clausal prolepsis due to an independent syntactic rule that is responsible for CP-extraposition in Dutch. Here, I examine and refute Sudhoff's alternative proposal, which attributes CP-extraposition in clausal prolepsis to a phonological rule. To start with, as discussed previously, in Sudhoff (2016), clausal prolepsis following factive predicates is taken to realize a syntactic structure where the proleptic pronoun is a D-head, taking a CP complement (see 3a). Sudhoff posits that the proleptic pronoun *es* can serve as a syntactic head but not a phonological head, meaning it cannot be stressed. An underlying assumption in his analysis is that the D-head of a DP containing a clause, such as the one realized by clausal prolepsis in (3a), must be a phonological head and therefore stressed. Consequently, CP-extraposition becomes a requirement in clausal prolepsis with factive predicates to avoid stressing the proleptic pronoun, *es*.

(70) * Peter hat es, dass Marie stur bleibt, bedauert.
Peter has it that Marie persistent stays regretted
'Peter regretted that Marie stays persistent.' Sudhoff (2016, (20b))

On the other hand, CP-extraposition is not required in (71) where clausal prolepsis is licensed with a non-factive predicate. In this case, Sudhoff's analysis is that clausal prolepsis realizes a distinct structure: *es* is merged as the verb's argument, and the CP is adjoined to the matrix CP (see also Frey 2016 for alternatives). Since the CP is not merged DP-internally in this case, CP-extraposition is not required because the D-head realized by the pronoun need not be stressed in this case.

(71) Maria hat es, dass Karl beruühmt ist, stets behauptet.
Maria has it that Karl famous is always claimed
'Maria has always claimed that Karl is famous.' Frey (2016, (24a))

While Sudhoff's analysis of CP-extraposition provides an explanation for the contrast between (70) and (71), his stress requirement faces certain challenges, as it lacks a solid foundation. Specifically, there is no established phonological principle requiring that the syntactic head of a DP must also serve as a phonological head. This can be illustrated by examining the case of Dutch *het*. Like *es*, it functions as a proleptic pronoun in clausal prolepsis, and, crucially, it cannot be stressed. Unlike *es*, it can also be used as a definite determiner combined with an NP (see Section 2.1). However, even though *het* is the syntactic head of the DP, there is no context in which it would trigger NP extraposition. This suggests that a syntactic head does not necessarily need to be a phonological head, contrary to the prediction of Sudhoff's street assignment rule. Note that in examples like (72), there is no context in which NP extaposition is allowed despite the fact that the NP complement of *het* is as large as a clause. This rules out the possibility that the size of D's complement determines the application of Sudhoff's stress assignment rule.

(72) het zeer smakelijke en goed gegaarde stuk van dit bijzondere rundvlees uit het zuidwesten van de
the very tasty and well cooked piece of this particular beef from the southwest of the
Verenigde Staten van Amerika
United States of America
'the very tasty and well-cooked piece of this particular kind of beef from the southwest of the United States of America'

From the perspective of Sudhoff's stress assignment rule, (72) suggests that the head of a DP can only be phonological head when it contains a clause, not an NP. Since the size of D's complement is not a factor, as discussed previously, this suggests that Sudhoff's stress assignment rule is sensitive to the grammatical distinction between clauses and NPs. This is problematic because phonological rules are not sensitive to grammatical distinctions. Similar issues arise with Sudhoff's stress requirement in different constructions in Dutch or beyond. For example, it fails to explain why extraposition is not obligatory in constructions where a CP is within a DP, like relative clauses and Dutch sentences such as *het feit dat...* 'the fact that ...' Likewise, in languages like Greek, where nominalized clauses are allowed, they can be introduced by a definite article, *to* 'the' which, like *es* in German, cannot be stressed. Even though *to* attaches directly to the embedded clause, forming a $[_{DP}$ D ... CP] structure identical to Sudhoff's proposal for German clausal prolepsis, CP-extraposition is not possible. This shows that the stress assignment rule lacks independent support.

(73) To oti ehis polus filus lei pola.
the that have.2SG many friends says a lot
'That you have many friends says a lot.'

Given multiple reasons undermining the assumed phonological requirement, an alternative explanation is necessary for the contrast in extraposition between factive and non-factive predicates. In this respect, the fact that (71) is accepted only as an apposition is important, as it might provide the key to understanding when central embedding is allowed. However, as noted before, since the acceptability of (71) and its distinct properties pertain to apposition, rather than clausal prolepsis, it falls outside the scope of this paper, and thus, I leave it for future research.

References

Abels, K. (2007). Towards a restrictive theory of (remnant) movement! *Linguistic variation yearbook 7*(1), 53–120.

Aboh, E. (2005). Deriving relative and factive clauses. In *Brugè, Laura et al.(eds.), Contributions to the thirtieth Incontro di Grammatica Generativa, Venezia: Libreria Editrice Cafoscarina, 2005, pp. 265-285.* Venezia, Libreria Editrice Cafoscarina.

Alexiadou, A. and H. Borer (2020). *Nominalization: 50 Years on from Chomsky's Remarks*, Volume 76. Oxford University Press.

Angelopoulos, N. (2022). Clausal prolepsis: an overview. *Proceedings of NELS 52 1*, 29–42.

Bennis, H. (1986). *Gaps and dummies*. Amsterdam University Press.

Bennis, H. (1995). The meaning of structure: The wat voor construction revisited. *Linguistics in the Netherlands 12*(1), 25–36.

Bi, R. A. and P. Jenks (2019). Pronouns, null arguments, and ellipsis in Mandarin Chinese. In *Proceedings of Sinn und Bedeutung*, Volume 23, pp. 127–142.

Bochnak, M. R. and E. A. Hanink (2021). Clausal embedding in washo: Complementation vs. modification. *Natural Language & Linguistic Theory*, 1–44.

Bogal-Allbritten, E. and K. Moulton (2018). Nominalized clauses and reference to propositional content. In *Proceedings of Sinn und Bedeutung*, Volume 21, pp. 215–232.

Bondarenko, T. (2020). Factivity from pre-existence: Evidence from Barguzin Buryat. *Glossa: a journal of general linguistics 5*(1).

Brandtler, J. and V. Molnár (2016). Rethinking clausal asymmetries. *Inner-sentential Propositional Proforms*, 241.

Broekhuis, H. and N. Corver (2019). *Syntax of Dutch: Verbs and Verb Phrases.*, Volume 3. Amsterdam University Press.

Broekhuis, H. and M. den Dikken (2012). *Syntax of Dutch: Nouns and noun phrases*, Volume 3. Amsterdam University Press.

Bruening, B. (2014). Precede-and-command revisited. *Language*, 342–388.

Bruening, B. (2018). CPs move rightward, not leftward. *Syntax 21*(4), 362–401.

Büring, D. and K. Hartmann (1995). All right! *On extraction and extraposition in German 11*, 179.

Cattell, R. (1978). On the source of interrogative adverbs. *Language*, 61–77.

Chomsky, N. (2000). Minimalist inquiries: The framework. In R. Martin, D. Michaels, and J. Uriagereka (Eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, pp. 89–155. MIT press.

Chomsky, N. (2008). On phases. *Current Studies in Linguistics Series 45*, 133.

Collins, C. (1994). The factive construction in Kwa.

Collins, C. (2015). Relative clause deletion. *Ángel J. Gallego & Dennis Ott (eds.) 50*, 57–70.

Djarv, K. (2019). *Factive and assertive attitude reports*. Ph. D. thesis, University of Pennsylvania.

Elbourne, P. (2013). *Definite descriptions*, Volume 1. Oxford University Press.

Elbourne, P. D. (2005). *Situations and individuals*, Volume 90. MIT press Cambridge, MA.

Elliott, P. D. (2016). Explaining DPs vs. CPs without syntax. *Proceedings of CLS 52*, 171–185.

Embick, D. (2003). Locality, listedness, and morphological identity. *Studia linguistica 57*(3), 143–169.

Farudi, A. (2007). An antisymmetric approach to Persian clausal complements.

Faure, R. (2022). (H)óti-clauses from DP to NPhood. The life of a Greek nouny clause. *Linguistische Berichte*.

Fischer, S. (2018). Locality, control, and non-adjoined islands.

Fox, D. (2002). Antecedent-contained deletion and the copy theory of movement. *Linguistic Inquiry 33*(1), 63–96.

Fox, D. and D. Pesetsky (2005). Cyclic linearization of syntactic structure. *Theoretical Linguistics 31*(1-2), 1–45.

Frey, W. (2016). On properties differentiating constructions with inner-sentential pro-forms for clauses. *Inner-sentential propositional proforms: Syntactic properties and interpretative effects 23*, 73–103.

Gentens, C. (2016). The discursive status of extraposed object clauses. *Journal of Pragmatics 96*, 15–31.

Gluckman, J. (2021). Null expletives and embedded clauses in logoori. *Syntax 24*(3), 334–375.

Grimm, S. and L. McNally (2022). Nominalization and natural language ontology. *Annual Review of Linguistics 8*, 257–277.

Halle, M. and A. Marantz (1993). Distributed morphology and the pieces of inflection. *1993*, 111–176.

Halpert, C. (2019). Raising, unphased. *Natural Language & Linguistic Theory 37*(1), 123–165.

Hanink, E. A. (2021). DP structure and internally headed relatives in Washo. *Natural Language & Linguistic Theory 39*(2), 505–554.

Heim, I. and A. Kratzer (1998). *Semantics in generative grammar*, Volume 1185. Blackwell Oxford.

Hinterholzl, R. (1999). *Restructuring infinitives and the theory of complementation*. University of Southern California.

Iatridou, S. and D. Embick (1997). Apropos pro. *Language*, 58–78.

Jarvis, R. (2021). Presuppositionality and syntactic nominalization infinite clausal complements. In *West Coast Conference on Formal Linguistics*, Volume 39.

Kastner, I. (2015). Factivity mirrors interpretation: The selectional requirements of presuppositional verbs. *Lingua 164*, 156–188.

Kayne, R. S. (2005). *Movement and silence*. Oxford University Press.

Kayne, R. S. (2019). *Questions of syntax*. Oxford University Press.

Kim, M.-J. (2009). E-type anaphora and three types of kes-construction in korean. *Natural Language & Linguistic Theory 27*(2), 345–377.

Kiparsky, P. and C. Kiparsky (1968). *Fact.* Linguistics Club, Indiana University.

Koopman, H. J. and A. Szabolcsi (2000). *Verbal complexes*, Volume 34. MIT Press.

Kratzer, A. (2002). Facts: Particulars or information units? *Linguistics and philosophy 25*(5/6), 655–670.

Kratzer, A. (2006). Decomposing attitude verbs. *Talk given at the Hebrew University of Jerusalem.*

Looyenga, S. (1992). A syntactic analysis of Dutch nominal infinitives. *Linguistics in the Netherlands 9*(1), 173–184.

McCawley, J. D. (1998). *The syntactic phenomena of English.* University of Chicago Press.

Moltmann, F. (2013). The semantics of special quantifiers in predicate position.

Moulton, K. (2009). *Natural selection and the syntax of clausal complementation.* University of Massachusetts Amherst.

Moulton, K. (2014). Simple event nominalizations. *Cross-linguistic investigations of nominalization patterns [Linguistik Aktuel/Linguistics Today 210]*, 119–144.

Moulton, K. (2015). CPs: Copies and compositionality. *Linguistic Inquiry 46*(2), 305–342.

Moulton, K. (2020). Remarks on propositional nominalization. *Nominalizations 50*, 255–276.

Müller, G. (1996). A constraint on remnant movement. *Natural Language & Linguistic Theory 14*(2), 355–407.

Müller, K. (2023). On noun-related complementizer clauses. In *On the Nominal Nature of Propositional Arguments*, Volume Linguistische Berichte Sonderheft 33.

Ott, D. and M. De Vries (2016). Right-dislocation as deletion. *Natural Language & Linguistic Theory 34*(2), 641–690.

Ozyildiz, D. (2017). Attitude reports with and without true belief. In *Semantics and Linguistic Theory*, Volume 27, pp. 397–417.

Partee, B. (1986). Noun phrase interpretation and type-shifting principles. *Formal semantics: The essential readings*, 357–381.

Patel-Grosz, P. and P. G. Grosz (2017). Revisiting pronominal typology. *Linguistic Inquiry 48*(2), 259–297.

Picallo, M. C. (2002). Abstract agreement and clausal arguments. *Syntax 5*(2), 116–147.

Postal, P. (1969). On so-called pronouns in english. *Readings in English transformational grammar 5*, 12–25.

Postal, P. M. and G. K. Pullum (1988). Expletive noun phrases in subcategorized positions. *Linguistic Inquiry 19*(4), 635–670.

Potts, C. (2002). The syntax and semantics of as-parentheticals. *Natural Language & Linguistic Theory 20*, 623–689.

Pütz, H. (1986). *Über die Syntax der Pronominalform es im modernen Deutsch.* Tübingen.Narr.

Pylkkänen, L. (2008). *Introducing arguments*, Volume 49. MIT press.

Rackowski, A. and N. Richards (2005). Phase edge and extraction: A Tagalog case study. *Linguistic Inquiry 36*(4), 565–599.

Reuland, E. (2011). What's nominal in nominalizations? *Lingua 121*(7), 1283–1296.

Rothstein, S. D. (1995). Pleonastics and the interpretation of pronouns. *Linguistic inquiry*, 499–529.

Roussou, A. (1991). Nominalized clauses in the syntax of Modern Greek. *UCLworking papers in linguistics 3*, 77–100.

Ruys, E. G. (2010). Expletive selection and CP arguments in Dutch. *The Journal of Comparative Germanic Linguistics 13*, 141–178.

Schwabe, K., W. Frey, and A. Meinunger (2016). Sentential proforms. *Inner-sentential Propositional Proforms: Syntactic properties and interpretative effects 232*, 1.

Schwarz, F. (2009). *Two types of definites in natural language*. University of Massachusetts Amherst.

Schwarz, F. (2019). Weak vs. strong definite articles: Meaning and form across languages. In A. Aguilar-Guevara, J. Pozas Loyo, and V. Vázquez-Rojas Maldonado (Eds.), *Studies in Diversity Linguistics*, Volume 25, pp. 1–37. Language Science Press.

Simonenko, A. (2014). *Grammatical ingredients of definiteness*. Ph. D. thesis, McGill University, Montreal.

Sportiche, D. (2016). Neglect.

Stowell, T. A. (1981). *Origins of phrase structure*. Ph. D. thesis, Massachusetts Institute of Technology.

Stroik, T. S. (1996). Extraposition and expletive-movement: A minimalist account. *Lingua 99*(4), 237–251.

Sudhoff, S. (2016). Correlates of object clauses in German and Dutch. *Inner-sentential propositional proforms: Syntactic properties and interpretative effects 23*, 48.

Takahashi, S. (2010). The hidden side of clausal complements. *Natural Language & Linguistic Theory 28*(2), 343–380.

Van Urk, C. and N. Richards (2015). Two components of long-distance extraction: Successive cyclicity in dinka. *Linguistic Inquiry 46*(1), 113–155.

Zwart, C. J. W. (1993). *Dutch syntax: A minimalist approach*. Ph. D. thesis, Rijksuniversiteit Groningen.

Zwart, J.-W. (2015). Precede-and-command revisited revisited. *Language 91*(3), 169–178.