How Come We Can Say 'How Come'?
Verb Second and V-to-I Movement
in Present Day English and Early Modern English

Freja Hovgaard Knudsen
Aarhus Universitet
201804725@post.au.dk / 03maj97@gmail.com

Abstract

In Present Day English (PDE), verb second (V2) occurs only in a few types of sentences, one of which is the interrogative main clause containing a wh-element that is not the subject of the sentence. Even then, it is never the main verb that travels to the second position. Instead, an auxiliary verb will move from the I° position to the C° position, while the main verb stays at its V° base position. For these reasons, the occurrences of the construction how come [emb] *in PDE is surprising.*

Paying attention to the syntactic structures of PDE and Early Modern English (EModE), this paper investigates the ways in which the phrase how come [emb] is different from other PDE syntactic structures. It introduces and examines explanations for the syntax of the phrase, e.g., how come *being a single constituent or* come *being a rare auxiliary verb. However, these explanations are found to be unsatisfactory. This paper instead proposes that how come [emb] is an idiomatic expression with a V2 structure adhering to earlier syntactic conventions such as those found in EModE and earlier versions of English where main verbs were allowed to move out of their V° position.*

Keywords: verb movement, language change, verb second (V2), Present Day English (PDE), Early Modern English (EModE), syntax

1. Introduction

The syntactic rules of a given language determine how words can form constituents and clauses in said language. From a vast set of data, the job of linguists is to deduce and formulate the rules of syntax. However, sometimes a syntactic construction turns up, contradicting the rest of the data. The clause *how come [emb]* is a Present Day English (PDE) example of this. In PDE, interrogative main clauses with a *wh*-element are constructed with an auxiliary C°, but *how come [emb]* has no auxiliary. The question, then, is if PDE interrogative main clauses theory should be discarded since it does not fit all available data. However, PDE words, clauses, and rules have not sprung into existence suddenly and in a vacuum. They have evolved for centuries, and sometimes specific words or clauses linger. It is therefore worth examining if e.g., *how come [emb]* could be a syntactic relic from a former variant of the English language.

Through the generative syntactic model used by e.g., Haegeman and Guéron (1999; 2007) and Vikner (2020a-d) and in a historical approach, I will examine and compare the rules of verb movement in

PDE and Early Modern English (EModE) with the aim of reaching a better understanding of *how come [emb]*'s syntactic structure. Structure hypotheses other than the historical one will also briefly be examined. Lastly, I will raise arguments for why *how come [emb]* can be considered an idiomatic expression that cannot be structurally altered.

2. Verb Second, Inflection Phrases and Complementiser Phrases

Verb Second (V2) refers to clauses in which 'the finite verb occupies the second position in the clause, irrespective of which constituent occupies the first position' (Vikner 2020d, 368). PDE is not a V2 language. In PDE, the finite verb <u>may</u> be the second overt constituent in a clause, but it is not necessarily so:

1) She enjoys reading books in the spring
2) In the spring, she enjoys reading books

In (1), *enjoys* <u>seems</u> to be in the second position, but it is clearly not so in (2) where the Preposition Phrase (PP), 'In the spring', has been topicalized. In V2 languages, e.g., Danish, the verb will remain the second constituent even when a non-subject constituent is topicalized:

3) Hun nyder at læse bøger om foråret
   she enjoys to read books in spring-the
4) Om foråret nyder hun at læse bøger
   in spring-the enjoys she to read books

When a non-subject constituent is topicalized in V2 languages, the subject follows the finite verb rather than precedes it. The difference between V2 and non-V2 main clauses can be explained by their highest maximal projections, i.e., the top-most syntactic phrase. In V2 languages, 'the finite verb in V2 main clauses occupies the same position that the complementizer … occupies in an embedded clause, namely C°' (Vikner 2020c, 5) whereas PDE main clauses are generally Ips, i.e., the phrase type complementizers select in embedded clauses (Haegeman and Guéron (1999, 97).

The head of the Inflection Phrase (IP), the I°, contains either the finite verb or an empty category if the finite verb cannot move out of its base position. As 'the properties of a sentence ("finite"/"non-finite") are a function of its [I]', the I is believed to head a clause. The specifier position of an IP (IP-spec) is immediately to the left of the I-bar (I'). Together the two form the IP, and agreement between verb and subject is ensured (1999, 94-95).

The occurrence of Complementiser Phrases (CPs) varies between languages and clause types as explained below. Traditionally, 'complementizers are referred to as subordinate conjunctions' and a such complementizer (C°s) 'selects a clause, an IP, as its complement' (1999, 97). However, CPs are not necessarily headed by a subordinate conjunction as not all CPs are embedded clauses. In embedded clauses in standard PDE, the C° and the CP-spec is never filled at the same time (Vikner 2020c, 1-2). In main clauses, both the CP-spec and the C° will be filled, the latter by the finite verb

that has moved there from I° (having possibly also moved from V° to I°, depending on the type of verb). That is how verb second (V2) occurs as the positioning of C° in relation to CP-spec on its left and all other positions on its right ensures that the verb will always be the second constituent of the clause, following only the constituent in the CP-spec. The CP-spec position will always be filled in main clauses and may be so by a number of different types of constituents, including but not limited to the subject of the clause, a topicalized PP, or a *wh*-element, depending on the language (Vikner 2020c, 3).

The maximal projection of V2 languages is always the CP-level, whereas the maximal projection of non-V2 languages can be either the IP- or CP-level depending on the type of sentence (Vikner 2020c).

The CP- and IP-levels are also used to differentiate between types of embedded clauses. In PDE, a finite embedded clause is always a CP regardless of whether it contains an overt complementizer in C°. Infinitival embedded clauses are only CPs <u>if</u> they contain an overt complementiser in C°, otherwise they are IPs (Haegeman and Guéron (1999, 99-101).
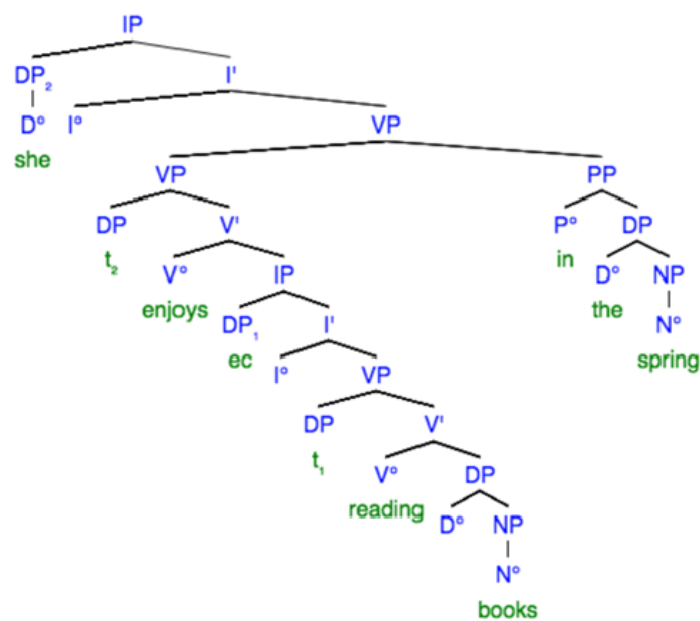

Figure 1   : Analysis of the syntactic structure of sentence (1).
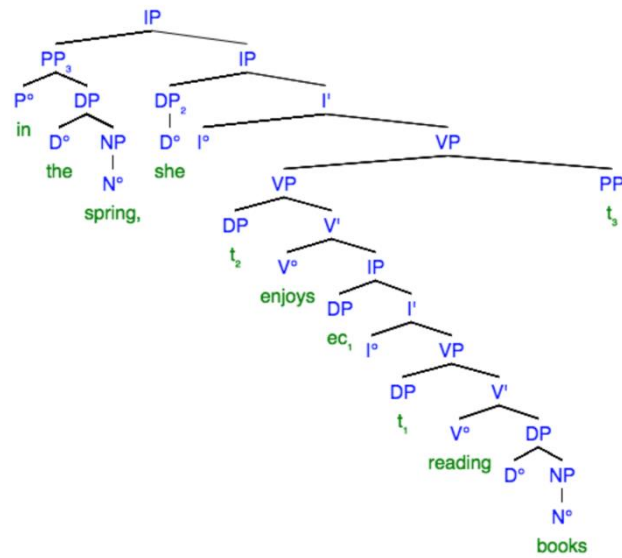
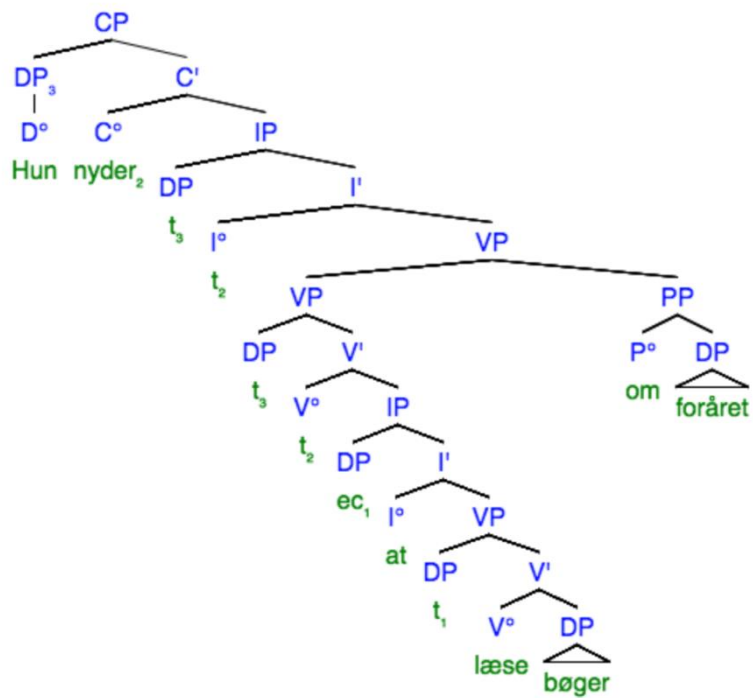Figure 2   : Analysis of the syntactic structure of sentence (2).

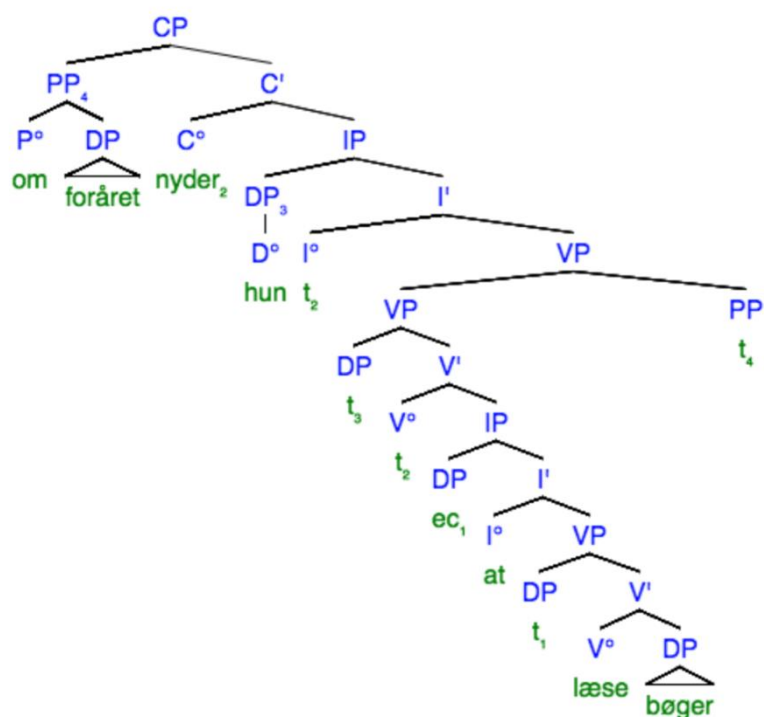Figure 3   : Analysis of the syntactic structure of sentence (3).

Figure 4   : Analysis of the syntactic structure of sentence (4).

In sentence (1), the maximal projection is an IP, and the Preposition Phrase (PP) remains at its base position. In (2) , the PP is topicalized and thus moves from its base position to the front of the clause. The remaining order of constituents in (2) is identical to (1), and the maximal projection is still an IP. For the Danish sentences, however, the maximal projections are CPs, meaning that the verb has moved to the second position at C°. In sentence (3), the subject DP has moved from IP-Spec to the CP-spec while the PP remains at its base position. In sentence (4), it is the PP that has moved from its base position to the CP-spec while the subject DP remains at the IP-Spec position.

In Figure 1 and Figure 3, the PP modifies the matrix clause. Another interpretation has the PP modifying the embedded clause[1], but then the PP cannot move out of the IP of the embedded clause and reach the CP-spec position in the main clause. This movement is shown in Figure 4 (Vikner 2020b, 2).

Thus, when the maximal projection in a main clause is a CP rather than an IP, the clause has V2 as the finite verb moves to the second position, C°. Even though PDE is not a V2 language, certain clause types in PDE do have a CP-level and thus have V2. One of those is the interrogative main clause. However, even then, there are restrictions as to the type of verbs that can make the move to the V2-position. It is when considering these constraints, the interrogative main clause *how come*

---

[1] Meaning that instead of doing the enjoying in spring, it is the reading that takes place in spring.

[emb] becomes particularly interesting as its structure seemingly is not determined by these restrictions.

3. V-to-I Movement in Present Day English

As displayed above, V2 entails verb movement from V° to I° to C°. The PDE verbs in (1) and (2), however, remain in V°. While PDE main verbs[2] always remain in V°, auxiliaries move to I° when finite, even though they do not continue the movement to C° in declarative main clauses (Vikner 2020d, 375). Modal auxiliaries are born in I° rather than in V°, but this distinction is not of importance here; V-to-I movement and modal auxiliaries' base position in I° is only visible in clauses containing a constituent between the I° position and the V° position, e.g., a medial adverbial (374). Consider these clauses without an adverbial:

    5)   I loved you
    6)   I will love you

Written out thus, it is not possible to see if *loved* in (5) and *will* in (6) have the same syntactic position. However, when an adverbial is added, the surface position of the finite verbs become observable, as seen in the table below:

		I°	Adverbial	V°	
7)	I		never	loved	you
8)	*I	loved	never		you
9)	I	will	never	love	you
10)	*I		never	will love	you

loved in (7) follows the adverbial, while *will* in (9) precedes it. If *loved* is made to precede the adverbial, or *will* is made to follow it, the clauses become ungrammatical. As we will see in section 7, the general invisibility of the verb's position without medial adverbials makes the syntactic structure of *how come [emb]* difficult to determine. Therefore, analyses of more typical and transparent structures in both PDE and Early Modern English (EModE) are needed to prepare a foundation for the study of the structure of *how come [emb]*.

Most adverbs, as e.g., *never*, do not require I° to be filled by an overt element as can be seen in (7). The negation adverbial *not* differs from the rest and <u>does</u> require I° to be filled (Vikner 2020a). Since main verbs cannot move to I°, in order to form a negated clause, the position must be filled either by a modal auxiliary, *have*, *be*, or by *do*-insertion if the clause does not contain another auxiliary verb:

---

[2] Save for copula main verb *be* which can move to both I° and C°. This, however, does not carry significance for this paper. Simply know that main verb *be* behaves like an auxiliary when it comes to verb movement (Haegeman and Guéron (1999, 322)

		I°	Adverbial	V°	
11)	*I		not	loved	you
12)	*I	loved	not		you
13)	I	did	not	love	you
14)	I	will	not	love	you

The ungrammaticality of (11) and (12) versus the grammaticality of (13) shows the need for *do*-insertion in clauses with negation when there is no auxiliary in the clause.PDE has V-to-I movement, but only for auxiliaries. Main verbs' inability to move to I° is clearly portrayed in negated clauses where I° <u>must</u> be filled, but main verbs <u>cannot</u> move out of V°, so *do*-insertion is needed instead. As will be seen in section 5, this is an important difference between PDE and EModE verb movement. This difference in syntactic rules could be the key to understanding *how come [emb]*.

4. The Structure of Interrogative Main Clauses in Present Day English

4.1 The Typical Structure

PDE is not a V2 language, but a few PDE clause types do have V2 structures. One of them is the interrogative main clause with an initial *wh*-element (Vikner 2020d, 368). One such clause is the *how come [emb]* structure, but as this clause, as mentioned above, does not adhere to the standard syntactic rules of PDE, considerations of the structure will be saved for later sections.

The V2 structure in more 'typical' PDE interrogative main clauses can best be seen when the initial *wh*-element is not the subject as the finite verb will then precede the subject as well as any medial adverbials there might be.

	CP-spec	C°	IP-spec	I°	Adverbial	VP-spec	V°	
15)			We$_1$		never	t$_1$	dance	at the theatre
16)	Where$_2$	do$_3$	we$_1$	t$_3$	never	t$_1$	dance	t$_2$?
17)	*Where$_2$		we$_1$		never	t$_1$	dance	t$_2$?

While the declarative clause is grammatical without the verb as the second constituent, the interrogative clause is ungrammatical if the word order is the same as in the declarative clause. The finite verb <u>must</u> be the second constituent and has moved to C°, so the structure is V2. Furthermore, the finite verb in C° <u>cannot</u> be the main verb:

	CP-spec	C°	IP-spec	I°	Adverbial	VP-spec	V°	
18)	*Where$_2$	dance$_3$	we$_1$	t$_3$	never	t$_1$	t$_3$	t$_2$?

If the clause does not contain an auxiliary, as e.g. (15), the interrogative clause has *do*-insertion: PDE interrogative main clauses are only grammatical with an overt non-main verb in the second position.
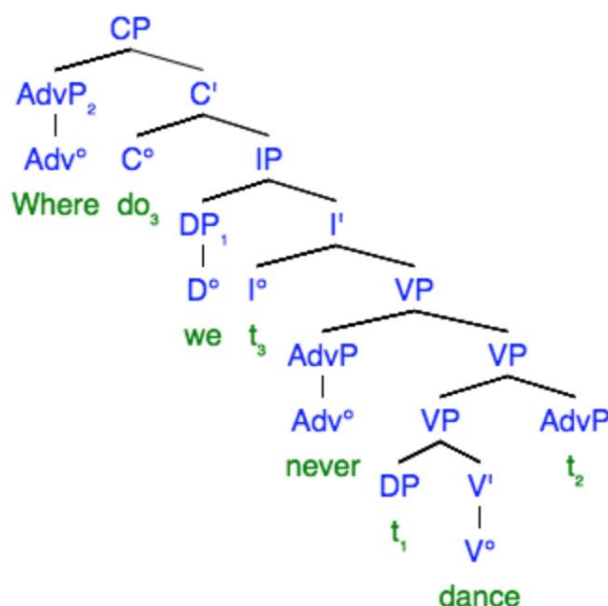

Figure 5   Analysis of the syntactic structure of sentence (16)[3].

In (16), the maximal projection is a CP. Instead of the main verb *dance* moving from its base position at V°, the auxiliary verb *do* is inserted and moves from its base position at I° to the second position at C°, meaning that PDE interrogative clauses V2 structure like all Danish main clauses, e.g. (3) and (4). However, in PDE, only auxiliaries and *be* can move to C°, which means that C° in PDE cannot be filled by a lexical main verb.

4.2 An Outliner from the Standard: How Come [*emb*]

Even so, PDE interrogative main clauses are not so neat as to always adhere to the standard syntax. Consider the phrase *how come [emb]*. It is an interrogative main clause with an initial *wh*-element, *how*, and yet *come*, a main verb, is the second constituent. According to the analysis of PDE interrogative main clauses above that should not be possible. Furthermore, it is unclear whether the matrix clause in *how come [emb]* has a subject. The embedded clause might be the subject in the matrix clause, but whether it is so, is not easily concluded. *come* in standard PDE use is an intransitive verb and does not take a complement. However, since *come* in *how come* moves differently and behaves

---

[3] CP-spec does not need to be filled by AdvP. It can be filled by other phrases as well.

differently semantically (cf. section 9) than 'ordinary' *come* does, it is possible that the question regarding complement is also different.

Embedded clauses rarely function as subjects and even when they do, it is problematic (Haegeman and Guéron (1999, 114-119). Changing the clause to a declarative clause in an attempt to investigate the subject situation results in ungrammaticality:

19)   *[I'm not like that] come.
20)   *It come [(that) I'm not like that].

In this manner, it is not possible to determine whether the embedded clause is the subject of the matrix clause. It is possible that the subject-dilemma can be examined more closely and with a broader variety of tests, but I will simply refer to the embedded clause as the sister of the verb rather than pass judgement on whether it is the subject of the matrix clause or the complement of *come*. Despite of this, the phrase *how come [emb]* is a somewhat frequent occurrence in American English often used in American movies and television series and with close to 10.000 attested uses in Corpus of Contemporary American English (COCA n.d., s.v. *how come*):

21)   How come you haven't cut a deal with him already … ?
22)   How come your mother, with 10 children, was able to do what she did?
23)   How come I'm not like that?

Compared to the roughly 10.000 occurrences of *how come*, COCA only has 41 results for *how comes*, suggesting that the PDE 3rd person, singular inflection morpheme, -*s*, renders the clause unacceptable to most speakers (COCA n.d., s.v. *how comes*). This further alienates the phrase from other PDE clauses.

5. V-to-I Movement in Early Modern English

Former variants of English, e.g. Early Modern English (EModE), allowed main verb movement out of V°: 'In the Early Modern English period and before, then, lexical verbs – just like auxiliaries – occupied the I position rather than the V position' (Hejná and Walkden, 2022, 142). It entails EModE main verbs having the ability to precede clause medial adverbials and form negated clauses without the aid of *do*-insertion. Looking at examples from EModE texts, in this case theatre plays, negated clauses formed without *do*-insertion are seen:

24)   Sir, I have not you by th' hand (1.3.64)
25)   I think not so, my lord (1.4.29)
26)   I know not, madam (1.5.98)

(Shakespeare (1602) 2019)

27)   … the thresher [that] dares not taste the smallest grain (1.1.53-55)
28)   But I come not here to discourse of that matter (1.2.25)
29)   This knight [h]ad not his name for nothing (1.5.102-103)

(Jonson (1606) 2001)

While it is important to remember that the examples above are from art rather than from everyday language, a certain connection and resemblance between the two types of language are expected. Especially the Jonson examples are persuasive. Shakespeare often wrote in rhyme and therefore might have used ungrammatical constructions to make his rhyme scheme work. However, since both texts display a variety of main verbs moving from V° to I°, it is an indication that main verb movement was a productive syntactic feature of the time. Even so, it should be noted that the syntactic structure with an auxiliary in I°, as we know it from PDE, was also used in EModE:

30)   I might not be admitted (1.1.23)
31)   [I]t will not curl by nature (1.3.95-96)
32)   If I do not usurp myself, I am (1.5.181)

(Shakespeare (1602) 2019)

33)   I do not mean it (1.4.26)
34)   He has not made his will (1.4.58)
35)   I will not trouble him now (1.5.75)

(Jonson (1606) 2001)

Here it is worth noting that most auxiliaries (e.g., *might, will, have*) have semantic meaning while *do* in the case of *do*-support is semantically empty (Hejná and Walkden 2022, 143). Therefore, (32) and (33) are arguably more notable when considering the syntactic structures of EModE as the auxiliary in I° does not carry semantic meaning. However, the other sentences aid to show that the structure with an auxiliary in I° is not limited to one specific type of auxiliary.

There being examples of auxiliaries in I°, however, does not alter or negate the fact that EModE main verbs are able to move to I°, a movement attested both in the examples above and in the secondary source material by Hejná and Walkden (2022). Thus, EModE negated clauses without an auxiliary verb exist. (26) is an example of this as the main verb *know* has moved from V° to I°. Simultaneously, EModE also has negated clauses with the same structure as PDE negated clauses, meaning that the main verb remains at V°, while the auxiliary verb is born and remains in the I°. (33) is an example of a EModE structure identical to the PDE structures:
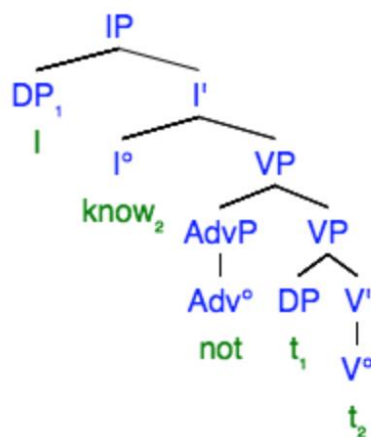

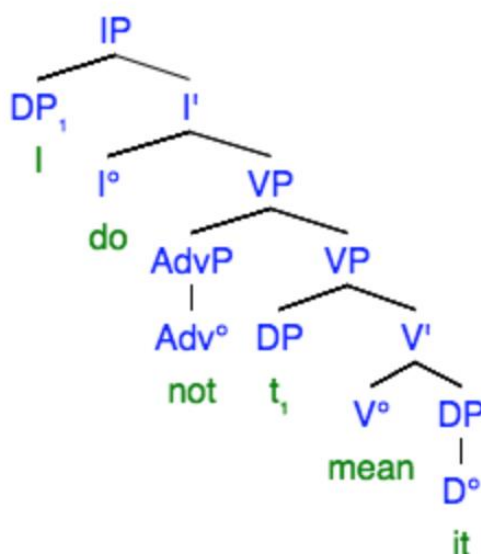Figure 6   Analysis of the syntactic structure of sentence (26).

Figure 7   Analysis of the syntactic structure of sentence (33).

6. The Typical Structure of Interrogative Main Clauses in Early Modern English

In PDE, only auxiliary verbs and main verb *be* can move to I° and thereby continue their movement to fill the second position at C° in interrogative main clauses as shown in section 4.1. In EModE, on the other hand, *all* verbs can be observed in I° so, as expected, all EModE verbs, including main verbs, can continue their movement to C° (Hejná and Walkden, 2022, 142-143). This is again observable in EmodE sources:

36)   why mourn'st thou? (1.5.62)
37)   How say you to that, Malvolio? (1.5.78)
38)   Where lies your text? (1.5.201)

(Shakespeare (1602) 2019)

39)   Why droops my Celia? (3.7.184)
40)   How fare you, sir? (1.3.16)
41)   How does your patron? (1.4.6)

(Jonson (1606) 2001)

With main verb-movement to C°, the EModE syntactic structure and verb-movement resembles Danish structures:
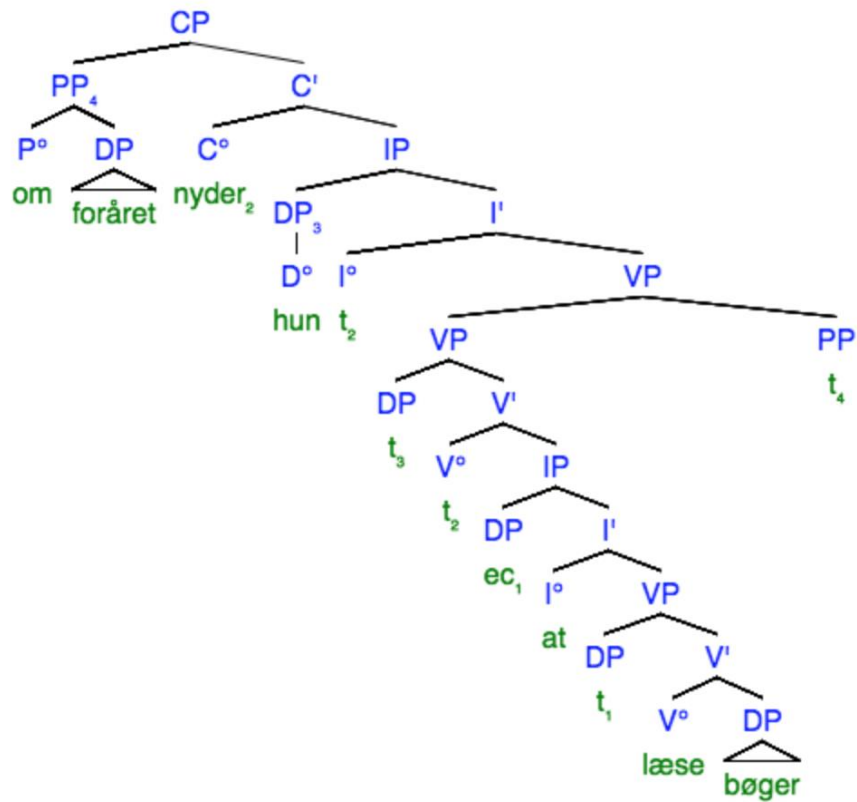
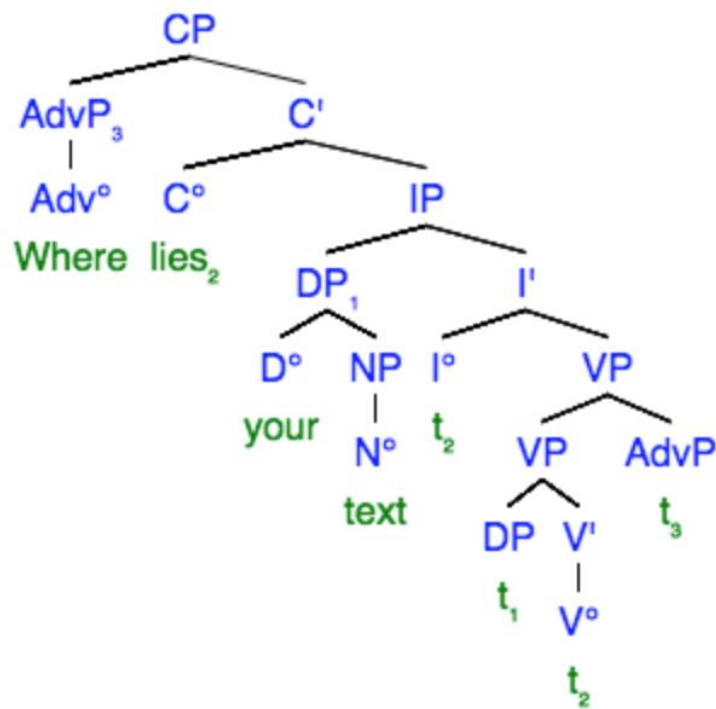Figure 8   Analysis of the syntactic structure of sentence (4).

Figure 9   Analysis of the syntactic structure of sentence (38).

The maximal projection of both Danish (4) and EModE (38) is a CP. In both structures, the finite main verb has moved to the second position at C°. In (4), it is the PP that has moved from its base position to the CP-spec, while it in (38) is the Adverbial Phrase (AdvP) that has moved from its base position to the CP-spec. In both structures, the subject DP remains at the IP-Spec position, having moved there from the VP-spec. The key to understanding *how come [emb]* might very well lie in this structure as we will see later.

As with EModE V-to-I movement, there are also examples of auxiliaries moving to C° in EModE:

42) Why do you speak to me? (5.1.182)
43) How does he love me? (1.5.246)
44) When did I see thee so put down? (1.3.78-79)

(Shakespeare (1602) 2019)

45) Why dost thou laugh so, man? (1.2.97)
46) How might I see her? (1.5.117)
47) How has my judgement wandered? (4.3.11)

(Jonson (1602) 2001)

Again, *do* is semantically empty (Hejná and Walkden 2022, 143) and the examples with *do*-support are therefore especially noteworthy. Most importantly, however, the structures above are the same as in PDE interrogative main clauses, meaning that the maximal projections for (16) and (46) are CPs. The sentences have the infinitive main verb staying at its base position in V°, while the auxiliary verb moves from I° to C°. In both sentences, the AdvP moves from its base position to the CP-spec and the subject DP remains at the IP-Spec position, having moved there from the VP-spec:
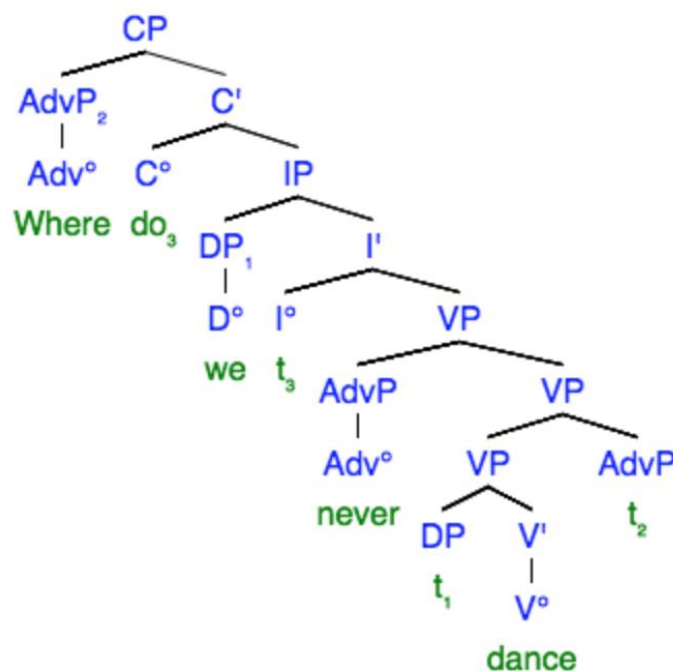

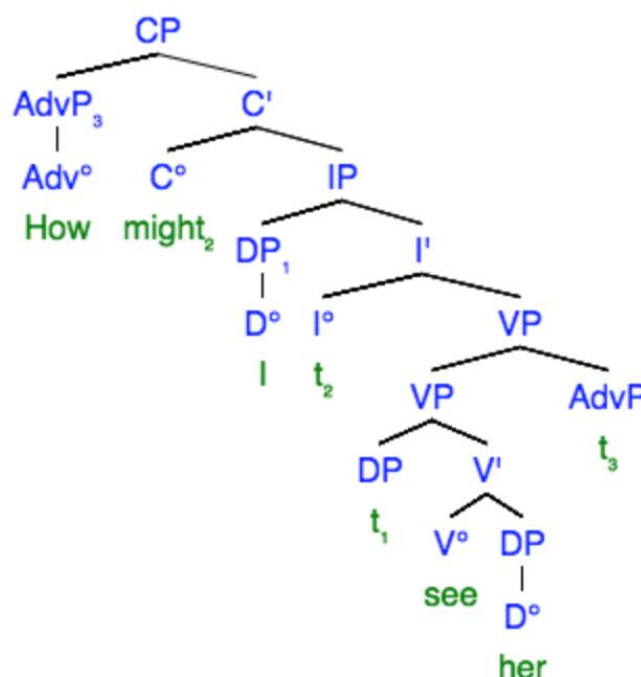Figure 10   Analysis of the syntactic structure of sentence (16).

Figure 11   Analysis of the syntactic structure of sentence (46).

There being multiple ways to structure an interrogative clause in EModE does not alter the fact that there <u>are</u> examples of main verbs in C°. There being both main verb V-to-I movement and main verb V2 in EModE supports the notion that it was <u>not</u> against the grammatical rules at the time to move main verbs. Again, auxiliaries occurring in C° do not negate examples of main verbs also doing so. While the data sets in this paper is too insignificant to reasonably comment on whether one structure was favoured above the other in EModE, it does suffice a wide enough sample of main verbs being in C°, aligning with the findings of e.g. Hejná and Walkden (2022), to support the claim that main verbs in earlier variants of English had the ability to be the second constituent in e.g. interrogative clauses.

7. The Structure of *How Come [Emb]*

7.1 The Difficulty Regarding Adverbial Alterations

As previously mentioned, *how come [emb]* differs from other PDE clauses. One difference is *come* being the second constituent, filling C° in the interrogative main clause. Other PDE interrogative clauses have C° filled by an auxiliary that has moved from I° to C° (or from V° to I° to C°), while the main verb stays in V°. Between the auxiliary and the main verb, the subject resides in IP-spec unless it is the *wh*-element and has been topicalized to CP-spec. However, even when the subject is in CP-spec, C° and V° can be differentiated with the aid of a medial adverbial, a left-hand sister to highest VP, as also V° and I° can (Vikner 2020d, 374). Is the subject not in CP-spec, the medial adverbial will follow the subject as well as the *wh*-element and precede the main verb:

48) Why does he often talk to me?
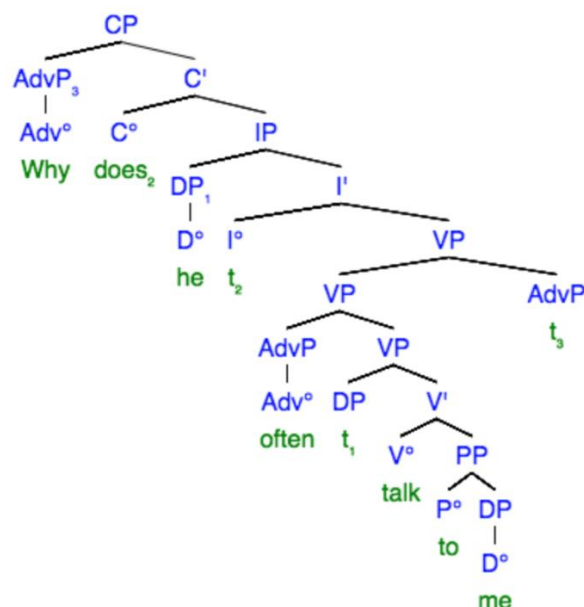49) *Why does often he talk to me?
50) *Why often does he talk to me?

Figure 12   Analysis of the syntactic structure of sentence (48).

The structure of (48) has the AdvP containing the *wh*-ememt *why* topicalized, so it is in the CP-spec position. Following the subject *he* and the trace left by the auxiliary verb *does*, is the AdvP *often*. This AdvP precedes the infinitive main verb *talk*. However, all adverbial additions, such as *often*, to the phrase *how come [emb]* result in ungrammaticality:

51) How come [I'm not like that]?
52) *How come often [I'm not like that]?
53) *How often come [I'm not like that]?
54) *How come [I'm not like that] often?

If the phrase behaved like an ordinary PDE clause, (52) would be grammatical if *come* was in C° and the embedded clause was <u>not</u> the subject, (53) would be grammatical if *come* had remained in V° in which case it preceding the embedded case would make it impossible for the embedded clause to be the subject of the matrix clause, and (54) would be grammatical if *come* had moved to C° and the embedded clause was the subject. However, since all adverbial alterations prove ungrammatical, this test cannot be used to determine *come*'s syntactic position. The failure of the test indicates that the structure of *how come [emb]* is unproductive.

7.2 Arguments for V2 in How Come [emb]

As shown above, testing *come*'s position is difficult. However, there are things that favour the reading of *come* having moved to C° rather than having stayed at V°. Firstly, as shown earlier in this paper,

PDE interrogative main clauses have V2 structure. Interpreting a single interrogative structure as not having V2 creates conflict in the established, data-supported theory. It is a theory internal argument and had been too weak on its own, if there had not been other arguments to support it. However, secondly, the rule of main verbs not being able to move to CP° is fairly recent in the history of the English language: E.g., EModE writings contain examples of main verbs being the second constituent in interrogative main clauses. That means that there is precedency for main verbs, such as *come*, occupying C°, even if said precedency is not from PDE. Thirdly, the phrase *how come* itself was used in stages of the English language when main verbs could still move to I° and C° (Merriam-Webster n.d., s.v. *how come…*), making it possible that the structure is a linguistic relic whose structure has been determined by the rules of older variants of English than PDE.

8. Other Explanations of the Structure of *How Come [emb]*

Other ideas as to how to explain the structure of *how come* do exist. One of these is the 'Reduction Theory' which suggests that the structure contains a number of erased elements: 'How ~~did it~~ come ~~about~~ [~~that~~ you were arrested]?' (Radford 2018, 218). Another possible explanation is to consider *how come* a single constituent, meaning that it fills <u>only</u> CP-spec like e.g., *why* does. Yet another explanation is that *come* could be a rare auxiliary verb. The two latter suggestions change the structure of *how come* from a matrix clause to a non-matrix clause. Below, the three ideas are investigated and certain problems arising from them are analysed.

8.1 The Reduction Theory

As mentioned above the 'Reduction Theory' suggests that the underlying structure of *how come [emb]* contains a number of erased elements: 'How ~~did it~~ come ~~about~~ [~~that~~ you were arrested]?' (Radford 2018, 218). While the theory does account for some aspects of the special phrase structure, e.g. the clause becoming ungrammatical if *come* is inflected with the morpheme –*s*, its deletion approach is unprecedented and the deletions arbitrary and across constituents. The theory is unsatisfactory (Radford 2018, 217-220).

8.2 How Come as a Single Wh-Element

If *how come* is considered a single constituent, it fills <u>only</u> CP-spec like e.g. *why* does. PDE does have instances of words and constituents, being spelled with a space such as the reciprocal *each other*. Therefore, a *wh*-element being spelled with a space could theoretically exist. However, '*how* can be postmodified by an aggressive non-D-linker like *the hell*', while *the hell* cannot follow *how come*, indicating that *how come* is not a constituent (Radford 2018, 234-235):

> 55)   How the hell come you didn't listen?
> 56)   *How come the hell you didn't listen?

Furthermore, *how come* solely exists clause initially, whereas other *wh*-elements can form echo questions by remaining at their base position and being stressed:

57)   Why didn't you come to the party?
58)   You didn't come to the party why?
59)   How come you didn't come to the party?
60)   *You didn't come to the party how come?

Had *how come*'s behaviour resembled that of *wh*-elements, it would have been an indication that *how come* functions like a single constituent and therefore should be considered one. However, as it is, even if *how come* was to be considered a single constituent, allowances would have to be made to account for it behaving differently than other, similar constituent, namely the *wh*-elements. Though not proving that it is not so, that makes the interpretation of *how come* being a single constituent unlikely. Additionally, the word order following *how come* is problematic if interpreting *how come* as a *wh*-element: The constituent following *how come* is not a finite verb as it should be in an interrogative main clause with an initial *wh*-element:

	CP-spec	C°	IP-spec	I°	AdvP	V°	PP
61)	How come		I	am_1	not	t_1	like that
62)	Why	am_1	I	t_1	not	t_1	like that?
63)	How	am_1	I	t_1	not	t_1	like that?

Considering *how come* as a single constituent results in a *wh*+ interrogative main clause <u>without</u> V2. The interpretation of *how come* being a *wh*-element is highly unlikely.

8.3 Come as an Auxiliary Verb

Some might consider *come* a rare auxiliary verb. Its position in C° would then follow PDE verb-movement rules. However, clauses can only have one finite verb. In clauses with an auxiliary, the auxiliary is the finite verb while the main verb is in its infinitival form. That is not the case in clauses with *how come*. The main verb in e.g. (23), *am*, is in its 1ˢᵗ person, present form rather than in its infinitival form, *be*. If *come* had been an auxiliary, the sentence would have read:

64)   *How come I not be like that?

Notice how copula main verb *be* has to be in its infinitival form <u>and</u> stay in V° rather than move to I°, since *come*'s trace fills I°:
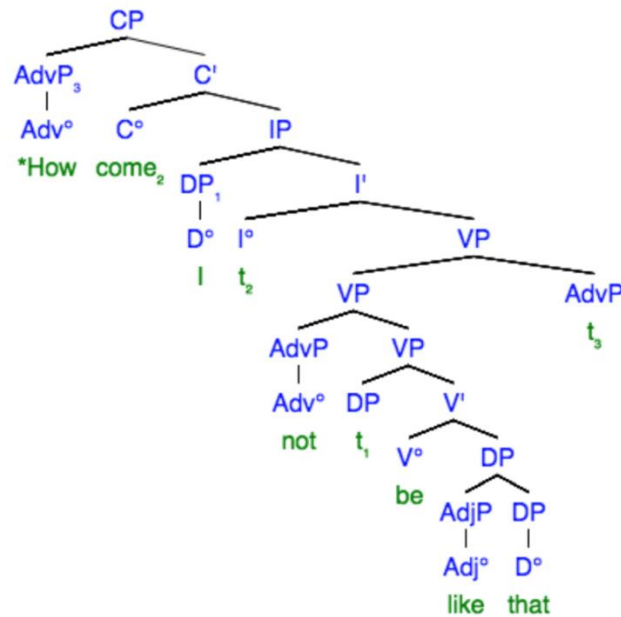
Figure 13   Analysis of the syntactic structure of sentence (64).

If we analyse (64) as an 'ordinary' interrogative main clause, using the structure of sentence (48) to guide us, we would have the auxiliary verb *come* moving to the second position at C°. As there is no embedded clause in this case, the subject *I* will move to the IP-spec and the negation *not* is placed between the subject and the infinitive main verb *be* at V°. The result is ungrammatical. Considering *come* to be an auxiliary verb is not a viable theory.

9. *How Come* as an Idiomatic Expression

9.1 The Semantics of How Come [*emb*]
Apart from its PDE rule-breaking syntactic structure, *how come* is also semantically interesting. The basic meaning of the phrase is most comparable with *why*, but it is used for 'conveying an informal tone' (Merriam-Webster n.d., s.v. *how come…*) and unlike some interrogative *why*-clauses *how come* is always unambiguous.

> 65)   How come you think he said that?
> 66)   Why do you think he said that?

(65) can only ever question 'why you think what you think', not 'why he said that'. (66), however, can be interpreted as 'why do you think so' <u>and</u> 'why did he say so'. The ambiguity can be seen in a syntax tree by observing the base position of *why*:
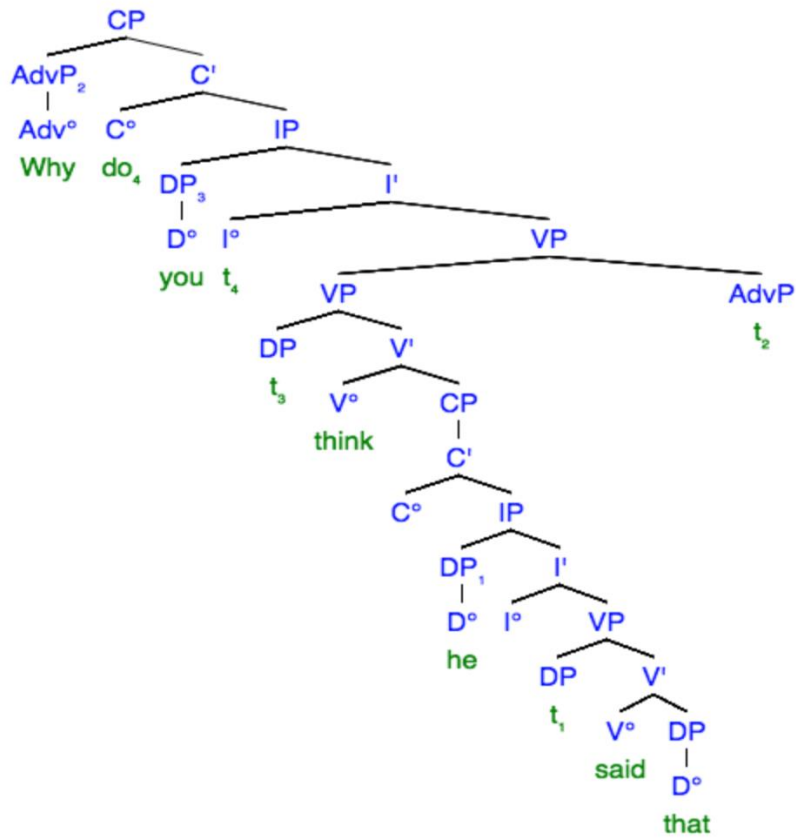
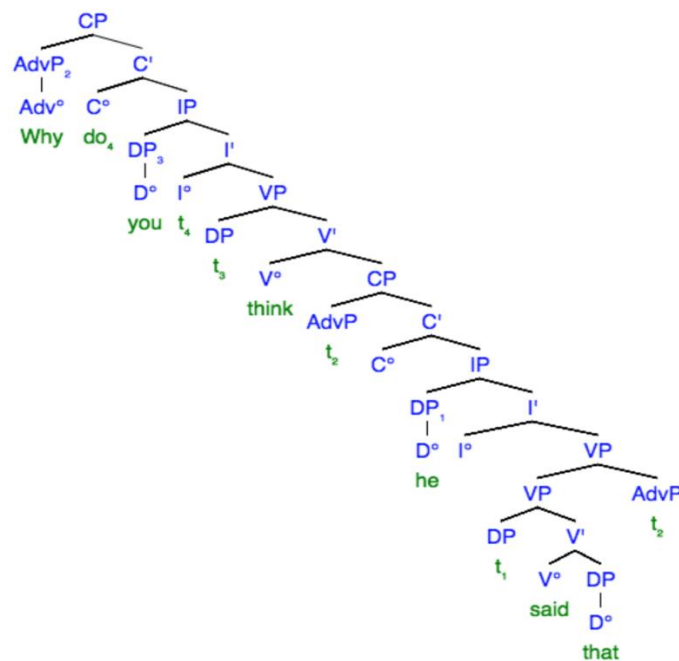Figure 14   Analysis of the syntactic structure of sentence (66).

Figure 15   Analysis of the syntactic structure of sentence (66).

It can be observed in Figure 14, how *why* can be taken to originate in the matrix clause, meaning that *why* is an adjunct to the VP with *think* in the V° and moves directly from its base position to the CP-spec of the main clause. This means that it is 'what you think' that is under scrutiny. In Figure 15, however, *why*'s base position is within the embedded clause, so the base position of *why* is as an adjunct to the VP with *said* in the V°. This means that *why* must first move from its base position to the CP-spec of the embedded clause and then to the CP-spec of the main clause. It is therefore 'what he said' that is being questioned.

How come, however, does not have multiple possible base positions. Unlike *why*, *how come [emb]* is a matrix clause on its own and therefore cannot originate from inside the embedded clause. (65) (repeated here as (67)) contains an embedded clause within an embedded clause within a main clause, while (66) (repeated here as (68)) only consists of an embedded clause within a main clause:

67)   [How come [you think [he said that?]]]
68)   [Why do you think [he said that?]]

Depending on whether one believes the embedded clause to be the subject of the main clause or the complement to *come* the general structure is as follows:
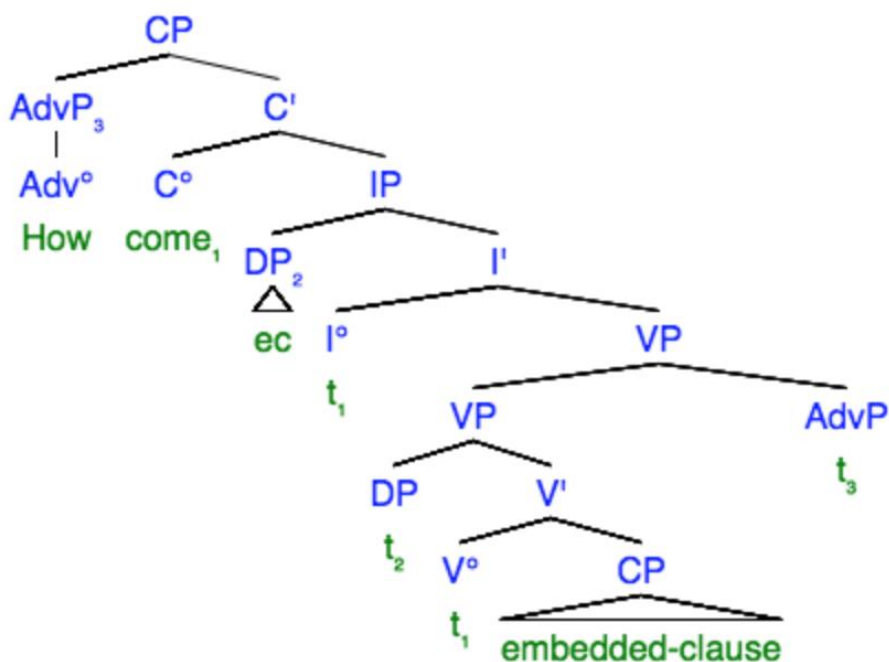

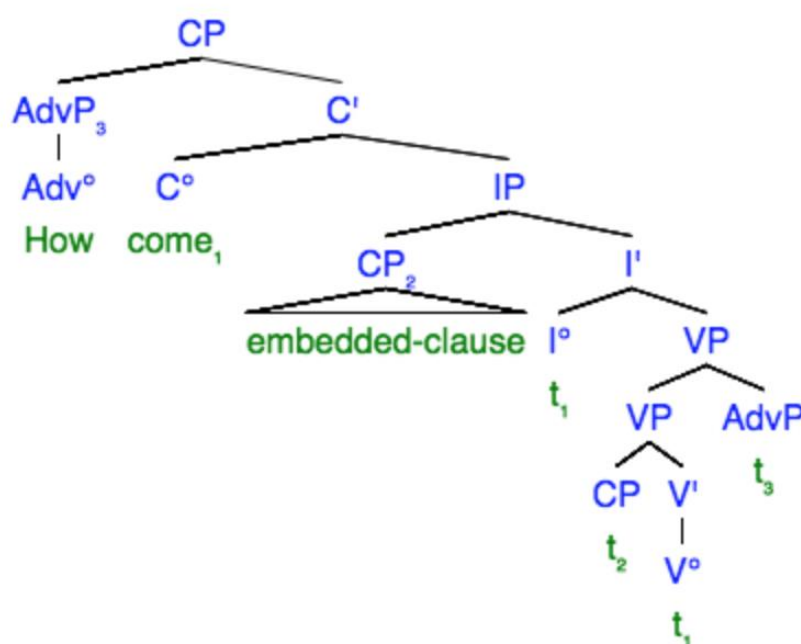Figure 16   The syntactic structure of *how come [emb]* if *[emb]* is the complement of *come*.

Figure 17   The syntactic structure of *how come [emb]* if *[emb]* is the subject of the clause.

If *[emb]* in *how come [emb]* is considered to be the complement of *come*, the embedded clause stays at its base position. The subject, placed in the IP-spec, will then be an empty category. If *[emb]* is considered to be the subject of the clause, the embedded clause moves from its base position to the IP-Spec position.

This difference between *why* and *how come* both in semantic interpretation and syntactic structure and movement suggests that the relation between the two are less straight-forward than what one might expect given the overlap in when they can be used. Considering this, alongside the great difference in syntax and meaning between 'ordinary' *how* and *come* and that of the combination *how come*, the indication is that *how come* might be an idiomatic chunk, as in 'a sentence containing an **idiomatic** element, the interpretation of the sentence is not fully determined by the interpretation of its parts' (Haegeman and Guéron (1999, 494, emphasis not mine).

The contrast in ambiguity, general meaning, and difference from the semantics of its parts indicates that *how come [emb]* might be an idiomatic expression.

9.2 The Differences Between How Come [Emb] and How Came [Emb]

Another argument for *how come* being an idiomatic expression is the phrase *how came [emb]*. *How came* is an archaic expression that only generates 50 results in COCA (n.d., s.v. *how came*), but its presence in beloved classics like *Pride and Prejudice* (Austen (1813) 2007) and *Hamlet* (Shakespeare (1603) 2007) makes it likely that PDE speakers still understand the expression and have a gut feeling about how it works syntactically. *How came [emb]* might <u>look</u> like the past tense version of *how come [emb]*, but that is not the case and, I believe, not how most native speakers would interpret the clause.

Firstly, *how come [emb]* and *how came [emb]* do not contain the same type of embedded clause. *How come* always has a finite embedded clause, while *how came* has a non-finite embedded clause: The embedded clause in *how come [emb]* is a CP without a complementiser, while the embedded clause in *how came [emb]* is an IP. Changing the type of embedded clause, results in ungrammaticality:

69)   How come you know?
70)   *How come you to know?
71)   How came you to know?
72)   *How came you know?

Had the difference between *how come [emb]* and *how came [emb]* been one of tense, they should have contained the same type of embedded clause. Secondly, *how come* can be grammatical with or without an embedded clause. *How came* is only grammatical with an embedded clause:

73)   (You know.) How come?
74)   *(You know.) How came?

How come can only be grammatical without an embedded clause if it immediately follows a main clause that could have been the embedded clause in *how come [emb]*. However, it is ultimately possible for *how come* to form a clause on its own, something *how came* can never do.

Thirdly, *how come* and *how came* cannot be used interchangeably when it comes to conveying meaning. *How come* has meant '"why?" since at least the mid-19th century' (Merriam-Webster n.d., s.v. *how come...*), while the semantics of *how came* are better described as a 'how' as a PDE translation of *Hamlet* shows (SparkNotes Editors 2005). The two clauses, while looking similar, do not convey the same question semantically.

In short, the difference between the two phrases cannot be ascribed to one of past and present tense. The two clauses are so semantically and syntactically different that they have no considerable relation to each other save for the relation that exists between any two interrogative clauses. This relation, of course, being that both are interrogative clauses.

9.3 Arguments for How Come being an Idiomatic Expression

How come [emb] stands out from other PDE clauses for several reasons: 1) the archaic V2 main verb structure, 2) its lack of tense and inflection, 3) the differences between it and *how came [emb]*, and 4) the contrasting level of ambiguity between *how come* clauses and certain *why* clauses. The most fitting description of *how come [emb]* is that it is an idiomatic expression with a syntactic structure dating back to at least the EModE period and in which only the constituents within the embedded CP clause may be changed.

10. Conclusion

PDE does not have V2 save for a few special cases, one of which is the interrogative main clause construction. However, even in that case, not all verbs can move to the second position. Only auxiliary verbs and main verb *be* can move to C°. These verbs are likewise the only PDE verbs which are found in I° when finite. C° and I° alike can only be filled by auxiliary verbs and *be*.

However, the phrase *how come [emb]* is an interrogative main clause, which means it must have a V2 structure, but it contains no auxiliary verb. It has only *come*. It opens up for several interpretations of which the most reasonable is that *come*, despite being a main verb, has moved to C°: The history of *how come [emb]* shows that the phrase has been used at least as far back as the EmodE period. Furthermore, examination of interrogative main clause structures and V-to-I movement in EModE shows auxiliary verbs and main verbs alike having the ability to fill both I° and C°. This indicates that main verb *come* moving to C° has precedency as earlier variants of English allowed main verb-movement to the second position and to I°. Thus, the structure in *how come [emb]* likely follows non-PDE verb-movement rules.

The structure not having changed with the tides of the English language indicates that the clause is 'frozen' - syntactically and semantically. At some point, it has become an unchangeable idiomatic expression and thus survived. This idea is further supported by the phrase's lack of ambiguity, the numerous differences between it and *how came [emb]*, and its inability to be changed by either inflection, word order, or medial adverbials. It is not a productive PDE clause, but a syntactic relic from earlier stages of the English language with a set semantic use.

Reference list

Austen, Jane. (1813) 2007. *Pride and Prejudice*. Reprint: Wordsworth Classics. Wordsworth.

COCA (Corpus of Contemporary American English). n.d. 'how came'. Accessed 06.01.2021, https://www.english-corpora.org/coca/.

COCA (Corpus of Contemporary American English). n.d. 'how come'. Accessed 06.01.2021, https://www.english-corpora.org/coca/.

COCA (Corpus of Contemporary American English). n.d. 'how comes'. Accessed 06.01.2021, https://www.english-corpora.org/coca/.

Haegeman, Liliane, and Jacqueline Guéron. 1999. *English Grammar: a generative perspective*. Blackwell Publishing.

Hejná, Michaela, and George Walkden. 2022. *A History of English*. Language Science Press. https://doi.org/10.5281/zenodo.6560337

Jonson, Ben. (1606) 2001. 'Volpone'. In *Ben Jonson's Plays and Masques*, edited by Richard Harp, 3-110. W. W. Norton & Company, Inc.

Merriam-Webster. n.d. 'how come people say how come'. 2020 Merriam-Webster, Incorporated. Accessed 13.12.2020, https://www.merriam-webster.com/words-at-play/how-come-people-say-how-come.

Radford, Andrew. 2018. '4 – How Come'. In *Colloquial English: Structure and Variation*, 216-292. Cambridge: Cambridge University Press.

Shakespeare, William. (1602) 2019. 'Twelfth Night'. In *Twelfth Night*, edited by Keir Elam, 155-354. Reprint, The Arden Shakespeare, Bloomsbury Publishing Plc.

Shakespeare, William. (1603) 2007. 'Hamlet – Prince of Denmark'. In *The Complete Works of William Shakespeare*, edited by Helen Taylor, 670-713. Wordsworth Editions Limited. Wordsworth.

SparkNotes Editors. 2005. 'No Fear Hamlet' SparkNotes.com. Act 5, Page 7. Accessed 06.01.2021. https://www.sparknotes.com/nofear/shakespeare/hamlet/page_286/

Vikner, Sten. 2020a. 'English Linguistics II/Advanced Syntax – 2. Functional Projections II: IPs'. Course handout, Department of English, Aarhus University.

Vikner, Sten. 2020b. 'Focus Area/Comparative Syntax – 11.-12. *wh*-movement and NP-movement'. Course handout, Department of English, Aarhus University.

Vikner, Sten. 2020c. 'Focus Area/Comparative Syntax – 7. & 8. CP, C°, and V2'. Course handout, Department of English, Aarhus University.

Vikner, Sten. 2020d. 'The Placement of Finite Verbs'. In *The Cambridge Handbook of Germanic Linguistics*, edited by Michael T. Putnam and B. Richard Page, 365–88. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108378291.017.