

Indirect Questions with Direct Question Word Order in English

Stanislav Pekař

Bachelor's Thesis
2025



Tomas Bata University in Zlín
Faculty of Humanities

Univerzita Tomáše Bati ve Zlíně

Fakulta humanitních studií

Ústav moderních jazyků a literatur

Akademický rok: 2024/2025

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení:	Stanislav Pekař
Osobní číslo:	H22578
Studijní program:	B0231P090005 Anglický jazyk pro manažerskou praxi
Forma studia:	Prezenční
Téma práce:	Nepřímé otázky se slovosledem přímé otázky v angličtině

Zásady pro vypracování

Studium odborné literatury
Teoretické zpracování tématu
Shromáždění textů k analýze
Analýza vybraného jazykového materiálu
Vyvození a formulace závěrů práce

Forma zpracování bakalářské práce: **tištěná/elektronická**
Jazyk zpracování: **Angličtina**

Seznam doporučené literatury:

Baker, C. L. 1995. *English Syntax*. 2nd ed. Cambridge, MA: MIT Press.
Brinton, Laurel J., and Donna M. Brinton. 2010. *The Linguistic Structure of Modern English*. Amsterdam: John Benjamins.
Huddleston, Rodney, and Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
Koenenman, Olaf, and Hedde Zeijlstra. 2017. *Introducing Syntax*. Cambridge: Cambridge University Press.
Radford, Andrew. 2009. *Analysing English Sentences: A Minimalist Approach*. Cambridge: Cambridge University Press.

Vedoucí bakalářské práce: **Mgr. Dagmar Masár Machová, Ph.D.**
Ústav moderních jazyků a literatur

Datum zadání bakalářské práce: **2. ledna 2025**
Termín odevzdání bakalářské práce: **6. května 2025**

L.S.

Mgr. Libor Marek, Ph.D.
děkan

doc. Mgr. Roman Trušník, Ph.D.
ředitel ústavu

Ve Zlíně dne 7. února 2025

PROHLÁŠENÍ AUTORA BAKALÁŘSKÉ PRÁCE

Beru na vědomí, že

- odevzdáním bakalářské práce souhlasím se zveřejněním své práce podle zákona č. 111/1998 Sb., v platném znění bez ohledu na výsledek obhajoby;
- bakalářská práce bude uložena v elektronické podobě v univerzitním informačním systému a bude dostupná k nahlédnutí;
- jedno vyhotovení bakalářské práce v listinné podobě bude ponecháno Univerzitě Tomáše Bati ve Zlíně k uložení;
- na moji bakalářskou práci se plně vztahuje zákon č. 121/2000 Sb. o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon) ve znění pozdějších právních předpisů, zejm. § 35 odst. 3;
- podle § 60 odst. 1 autorského zákona má Univerzita Tomáše Bati ve Zlíně právo na uzavření licenční smlouvy o užití školního díla v rozsahu § 12 odst. 4 autorského zákona;
- podle § 60 odst. 2 a 3 mohu užít své dílo – bakalářskou práci – nebo poskytnout licenci k jejímu využití jen s předchozím písemným souhlasem Univerzity Tomáše Bati ve Zlíně, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly Univerzitou Tomáše Bati ve Zlíně na vytvoření díla vynaloženy (až do jejich skutečné výše);
- pokud bylo k vypracování bakalářské práce využito softwaru poskytnutého Univerzitou Tomáše Bati ve Zlíně nebo jinými subjekty pouze ke studijním a výzkumným účelům (tj. k nekomerčnímu využití), nelze výsledky bakalářské práce využít ke komerčním účelům;
- pokud je výstupem bakalářské práce jakýkoliv softwarový produkt, považují se za součást práce rovněž i zdrojové kódy, popř. soubory, ze kterých se projekt skládá; neodevzdání této součásti může být důvodem k neobhájení práce.

Prohlašuji, že

- jsem na bakalářské práci pracoval(a) samostatně a použitou literaturu jsem řádně citoval(a); v případě publikace výsledků budu uveden(a) jako spoluautor;
- odevzdaná verze bakalářské práce a verze elektronická nahraná do IS/STAG jsou obsahově totožné.

Ve Zlíně, dne

.....

podpis autora

ABSTRAKT

Tato bakalářská práce se zabývá fenoménem přímého slovosledu v nepřímých otázkách v angličtině. Cílem výzkumu je analyzovat přetrvávání inverze podmětu a pomocného slovesa v nepřímých otázkách, kde je tato inverze podle standardních syntaktických pravidel běžně blokována. Práce zkoumá, zda a za jakých podmínek si podřadné nepřímé otázky zachovávají slovosled přímých otázek. Teoretická část popisuje deklarativní slovosled v podřadných interrogativech, mechanismus T-to-C a syntaktické principy odpovědné za blokování inverze. Praktická část vychází z dat Korpusu současné americké angličtiny (COCA) a zaměřuje se na pět vybraných sloves: *wonder*, *ask*, *find out*, *remember* a *know*. Tyto výskyty jsou analyzovány ve spojení s vybranými *wh*-zájmeny a pomocnými slovesy *do*, *have*, a *be*. Dále je věnována pozornost gramatické funkci slovesa *be* a rozložení jeho plného a zkráceného tvaru. Analýza ukázala, že na inverzní slovosled působí více faktorů, včetně typu slovesa, kontrakcí a komunikačního stylu. Variace se nejčastěji objevuje v neformálních kontextech a odráží reálné užití v mluvené angličtině.

Klíčová slova: vnořené otázky, nepřímé otázky, přímé otázky, inverze podmětu a pomocného slovesa, inverzní slovosled, *wh*-výrazy, kontrakce, syntax, korpus COCA

ABSTRACT

This bachelor thesis explores the phenomenon of non-canonical word order in embedded *wh*-questions in English. The aim is to analyse the persistence of subject-auxiliary inversion in indirect questions, where inversion is generally blocked according to standard syntactic rules. The thesis investigates whether and under what circumstances embedded clauses preserve the word order of direct questions. The theoretical part outlines the canonical structure of embedded interrogatives, the mechanism of T-to-C movement, and the syntactic principles responsible for blocking inversion. The practical part is based on data from the Corpus of Contemporary American English (COCA) and focuses on five chosen verbs: *wonder*, *ask*, *find out*, *remember*, and *know*. These are examined in combinations with selected *wh*-words and the auxiliaries *do*, *have*, and *be*. Further attention is paid to the grammatical function of *be* and to the distribution of its full and contracted forms. The analysis revealed that non-canonical word order is influenced by multiple factors, including

the main verb, contraction, and register. Variation appears particularly in informal contexts and reflects actual usage in real-life English.

Keywords: embedded questions, indirect questions, direct question, subject-auxiliary inversion, non-canonical word order, wh-words, contraction, syntax, COCA corpus

ACKNOWLEDGMENTS

First of all, I would like to express my sincere gratitude to my supervisor, Mgr. Dagmar Masár Machová, Ph.D., for her valuable advice, support, and patience throughout the process of writing this bachelor thesis. Her guidance and expertise were truly helpful at every stage of the work.

I would also like to thank my parents and friends for their support and encouragement during my studies.

CONTENTS

INTRODUCTION	13
I THEORETICAL PART	14
1 ENGLISH WORD ORDER AND ITS SYNTACTIC MODIFICATIONS	15
1.1 ENGLISH WORD ORDER	15
1.2 SUBJECT-AUXILIARY INVERSION	16
1.3 FRONTING IN ENGLISH	17
2 INTERROGATIVES	19
2.1 OPEN AND CLOSED INTERROGATIVES	19
2.2 ECHO QUESTIONS	20
2.3 WH-CONSTITUENTS: SUBSTITUTION	21
2.4 WH-MOVEMENT	22
3 DIRECT AND INDIRECT QUESTIONS	23
3.1 DIRECT QUESTIONS	23
3.2 INDIRECT QUESTIONS	24
3.3 T-TO-C MOVEMENT	26
4 INDIRECT QUESTIONS WITH DIRECT QUESTION WORD ORDER	30
4.1 IQS WITH SUBJECT WH-CONSTITUENTS	30
4.2 SUBJECT-AUXILIARY INVERSION IN INDIRECT QUESTIONS	31
II PRACTICAL PART	33
5 METHODOLOGY	34
6 SYNTACTIC ANALYSES OF SELECTED VERBS	36
6.1 FREQUENCY OF VERBS + WH-WORDS + AUXILIARIES	36
6.1.1 Wonder	36
6.1.2 Ask	37
6.1.3 Find Out	37
6.1.4 Remember	38
6.1.5 Know	38
6.2 FUNCTIONS OF THE AUXILIARY <i>BE</i>	38
6.2.1 Wonder	39
6.2.2 Ask	39
6.2.3 Find Out	40
6.2.4 Remember	41
6.2.5 Know	41
6.3 DISTRIBUTION OF <i>Is</i> AND 'S (CONTRACTED FORM)	42
6.3.1 Wonder	42
6.3.2 Ask	43
6.3.3 Find Out	43

6.3.4	Remember	44
6.3.5	Know	44
CONCLUSION.....		46
BIBLIOGRAPHY		48
LIST OF FIGURES		49
LIST OF TABLES		50
LIST OF ABBREVIATIONS		51

INTRODUCTION

The aim of the thesis is to analyse phenomenon of indirect questions in English that retain the word order of direct questions. This phenomenon is later referred to as embedded wh-questions with non-canonical word order. While embedded questions typically follow the structure of declarative sentences, examples where subject-auxiliary occurs are also found, especially in informal or spoken English. These non-canonical structures, though less frequent, reveal interesting patterns related to main verbs that this thesis explores.

Embedded questions have attracted attention because they reveal interactions between the sentence structure and meaning, especially in contexts where speakers unconsciously deviate from expected grammar rules. Studying these structures can thus provide insights into how syntax operated in real, dynamic use.

The theoretical part of the thesis outlines the basics of the standard English word order followed by the explanation of subject-auxiliary inversion and fronting. The second part deals with English interrogatives, the substitution by wh-constituents, and the wh-movement. The third chapter elaborates on the second chapter by explaining how direct and indirect questions are formed, and why is the subject-auxiliary inversion blocked in embedded questions. The fourth chapter focuses specifically on indirect questions with direct question word order, providing an insight into the daily usage of such occurrences.

The practical part of the thesis is based on data collected from the Corpus of Contemporary American English (COCA). Five main verbs (*wonder*, *ask*, *find out*, *remember*, and *know*) are selected to provide a range of verbs that typically embed with wh-clauses. The data collection focuses on combination of these verbs with selected wh-words (*where*, *when*, *why*, and *how*) followed by auxiliaries (*do*, *have*, and *be*). Further analysis is carried out to classify the grammatical functions of *be* (copula, progressive aspect, and passive voice) and to observe possible effects of the contracted form of *is* on the overall word order.

The goal of the thesis is to identify the circumstances under which non-canonical embedded wh-questions occur, and to determine which syntactic factors, if any, contribute to the preservation of direct question word order in embedded contexts.

I THEORETICAL PART

1 ENGLISH WORD ORDER AND ITS SYNTACTIC MODIFICATIONS

This section explores fundamental aspects of English word order and syntactic modifications that influence a clause structure.

The first subchapter establishes that English follows a subject-predicate structure is the foundation of canonical word order. This ensures clarity in communication and systematically organize syntactic constituents in sentences. Additionally, this subchapter distinguishes between surface structure and deep structure and introduces canonical and non-canonical clauses, emphasizing that syntactic modifications such as inversion and fronting disrupts the default word order.

The second subchapter focuses on one major syntactic modification – subject-auxiliary inversion. It results in placing an auxiliary verb before the subject. The subchapter also addresses do-support inversion, that is necessary when no auxiliary is present in the sentence.

The last subchapter examines fronting as another syntactic process. Fronting involves moving a constituent, such as a noun phrase to a different position. Together, these chapters provide understanding of English sentence structure and its possible modifications and sets the base for a deeper discussion about indirect questions.

1.1 English Word Order

English syntax deals with specific combinations of words that carry a certain meaning. Words can be perceived as parts that need to fit together while following grammatical rules. These rules are bound to syntactic categories that words represent and each category manifests different syntactic behaviour. Every word combined with another word or set of words forms constituents and demonstrates a specific hierarchy within the constituent that is bound to the essential element in it, called the head. In more complex sentences, the head that behaves as a whole and can be pre-modified or post-modified by other syntactic categories or whole phrases (Koenenman and Zeijlstra 2017, 9-12).

These constituents are furthermore described by Brinton and Brinton (2010, 185-192). They explain it as generative grammar that describes how speakers produce an unlimited number of possible sequences while following a finite number of syntactic rules in a language. Constituents function as units and follow a hierarchical order called deep structure (D-structure). In contrast, the surface structure (S-structure) refers to the actual word order

that speakers produce. While S-structure shows the visible arrangement of words, D-structure captures the underlying syntactic relationships that remain constant even when the surface order changes, such as in syntactic processes like passivization or question formation (Brinton and Brinton 2010, 186-192).

Huddleston and Pullum (2002, 44-48) explain this constituent-forming theory with constructions consisting of subject and predicate that are called clauses. They might represent either a sentence or be a part of a larger, more complex sentence. When these two constituents are in this order, where a predicate follows the subject, it is called a declarative. It might be also called a canonical clause, and whenever any syntactic process, e.g. inversion, fronting or passivation, changes the default word order, it is called non-canonical clause (Huddleston and Pullum 2002, 44-48).

The general pattern of English sentences formation is described by Brinton and Brinton (2010, 188-192) through a general pattern where two main components are combined to form a sentence: the determiner phrase (DP) which functions as the subject, and a verb phrase (VP) that represents the predicate. This subject-predicate order represents typical English declarative clauses. It demonstrates how English depends on word order to organize and recognize grammatical categories, ensuring the meaning of the sentence is clear through the consistent placement of its core elements (Brinton and Brinton 2010, 188-192).

1.2 Subject-Auxiliary Inversion

The previous subchapter suggests that the default English word order is formed by a subject and a predicate that follows the subject. One of the syntactic processes that changes the word order is a subject-auxiliary inversion. To understand, how the subject-auxiliary inversion works, first it must be determined what is an auxiliary verb. “The auxiliary verbs of English are the following: can, may, will, shall, must, ought, need, dare, be, have, do” (Huddleston and Pullum 2002, 92).

According to Huddleston and Pullum (2002, 92-93), an auxiliary verb occupies a special position in a sentence. It allows to create negation (1b), inversion (1c), code, and emphasis of the canonical sentence (1a) (Huddleston and Pullum 2002, 92-93). The most crucial modification is in this case inversion as it is related to main topic of this thesis.

(1) a. *She has finished her homework.*

b. *She has not finished her homework.* [NEGATION]

c. Has she finished her homework? [INVERSION]

If the sentence does not include any auxiliary in the canonical structure and a non-canonical structure is to be produced, it is necessary to use the auxiliary *do*. This support process of using verb *do* is used when the sentence contains only a lexical verb (2a) (Huddleston and Pullum 2002, 93).

(2) *a. Tom likes chocolate.*

b. Does Tom like chocolate? [DO-SUPPORT INVERSION]

The example (2b) demonstrates do-support inversion of a canonical sentence which is absent of any other auxiliary that can be inverted, therefore *do* is used. In this example (3b) the change in word order is obvious. The subject *Tom* moves to a pre-verbal position. This process is called “a primary inversion” (Huddleston and Pullum 2002, 94-95). Inversions are observable in the case of closed (3a-b) and open interrogatives (4a-b) that are explained in chapter 2 (Huddleston and Pullum 2002, 95).

(3) *a. They saw her.* *b. Did they see her?*

(4) *a. He broke the glass.* *b. Who broke the glass?*

1.3 Fronting in English

Another syntactic modification that influences word order is fronting. As mentioned earlier, English follows subject-predicate word order, but fronting disrupts this structure.

Huddleston and Pullum (2002, 1366-1372) describe it as a movement of a constituent from its canonical position to the initial position in the sentence. The object *this problem* in (5a) is a complement of the verb *to solve*, meaning that it is required by the verb, otherwise the sentence would be ungrammatical (5c). In (5b) the object complement of *to solve* is fronted (Huddleston and Pullum, 2002, 1366-1372).

(5) *a. I have no idea how to solve this problem.*

b. This problem, I have no idea how to solve.

*c. *I have no idea how to solve.*

According to Birner and Ward (1998, 4), fronting is not restricted to the object only (6a). Several other phrases can be fronted, for instance a prepositional phrases (PP) (6b), verb phrases (VP) (6c), and adverbial phrases (AdvP) (6d) (Birner and Ward 1998, 4).

- (6) a. *I don't like this book.* > *This book, I don't like.*
b. *He waited at the bus stop.* > *At the bus stop, he waited.*
c. *He should have told the truth earlier.* > *Told the truth earlier, he should have.*
d. *We will meet again someday.* > *Someday, we will meet again.*

2 INTERROGATIVES

This chapter focuses on syntactic formation of interrogatives in English, focusing on their internal structure and the position of *wh*-words.

The first subchapter distinguishes between two main types of interrogatives – open and closed interrogatives. It provides examples of *yes/no* questions and *wh*-questions particularly through presence or absence of subject-auxiliary inversion.

The second subchapter explores a special type of interrogatives – echo questions. These questions contain a *wh*-word that is positioned *in situ* (i.e. at the end of a clause), therefore they resemble a declarative (canonical) sentence.

The third subchapter focuses on the substitution of phrases (XPs) by *wh*-constituents, explaining how different types of phrases can be replaced by *wh*-constituent to form a question.

The final subchapter addresses a syntactic process by which a *wh*-constituent is moved to the front of a clause – fronting. It also discusses two strategies of fronting prepositional phrases: *pied-piping* and *preposition stranding*.

2.1 Open and Closed Interrogatives

Primarily, it is necessary to explain the difference between questions and interrogatives. Sigmund (2017, 17-18) explains that it is important to distinguish between the syntactic form of a sentence and its communicative function. While interrogatives use a specific syntactic structure (such as subject-auxiliary inversion), the term question is applied more broadly to any sentence that intends to elicit information. Sigmund (2017, 17-18) calls closed and open interrogatives, explained in 2.1, *polar (yes/no) interrogatives* (1) and *open (wh-) interrogatives* (2) (Sigmund (2017, 17-18)).

(1) Are you ready? [Yes/No]

(2) What is wrong? [Wh-]

In (1) the only possible answer is *yes* or *no* while the interrogative exhibits subject-auxiliary inversion that is always necessary to create a closed interrogative. The example (2) is an open interrogative that aims to obtain information. The *wh*-word *what* in this example replaces the subject.

Sobin (1990, 142-145) calls these structures syntactic echo question. These questions preserve the original sentence's word order as Radford (2009) explain. The *wh*-word remains in situ, without any fronting and inversion. This is illustrated in (7a-b) (Sobin 1990, 142-145).

- | | |
|--|---------------------------|
| (7) a. <i>She kissed your brother.</i> | [DECLARATIVE] |
| b. <i>She kissed who?</i> | [SYNTACTIC ECHO QUESTION] |

2.3 Wh-Constituents: Substitution

This section of the thesis focuses on substitution by *wh*-words in open interrogatives that are introduced by “one of the interrogative words *who*, *whom*, *whose*, *which*, *what*, *when*, *where*, and *how*, etc” (Huddleson and Pullum 2002, 856) that are essential for creating open interrogatives as the answer to these questions is not just *yes* or *no*.

These words replace or substitute whole phrases. For example, in (8a) *where* substitutes the whole prepositional phrase (PP), e.g. *pp*[to the park]. The example (8b) shows substitution of adverbial phrase (AdvP) by *when*, e.g. *AdvP*[in the morning], and in the last example (8c), *what* replaces the determiner phrase (DP), e.g. *DP*[her homework].

- | | | |
|--------------------------------------|--------|--------------------------------------|
| (8) a. [<i>Where</i>] did Jane go? | [PP] | <i>She went [to the park].</i> |
| b. [<i>When</i>] did Jane arrive? | [AdvP] | <i>She arrived [in the morning].</i> |
| c. [<i>What</i>] did Jane do? | [DP] | <i>She did [her homework].</i> |

Understanding what phrase is substituted is essential for determining a sentence structure and syntactic operations that must follow. This particularly applies to subject-auxiliary inversion that is triggered by substitution of non-subject constituents as shown in (8a-c). Although, when a subject is substituted by a *wh*-constituent, it does not trigger the inversion (Huddleston and Pullum 2002, 873). Examples (9a-b) demonstrate subject substitution by *who* without any change in word order.

- | | |
|--|------------------------|
| (9) a. <i>The man opened the window.</i> | |
| b. <i>Who opened the window?</i> | [subject substitution] |

Even when the subject of the source sentence (10a) is syntactically complex, the word order of the formed declarative, where the subject is substituted by *wh*-constituent (10b), preserves the canonical subject-predicate order. In (10a) the noun is both pre-modified and post-modified, still the replacement by *who* does not change the word order.

(10) a. *The tall man in a blue coat opened the window.*

b. *Who opened the window?*

2.4 Wh-Movement

Apart from the substitution of a phrase, another syntactic operation takes place – fronting. This operation is explored in general terms in section 1.3, but this subchapter focuses specifically on fronting of wh-constituents. While substitution identifies what is being questioned, fronting shows how wh-elements are positioned in the sentence.

As the previous chapters suggested, wh-constituents are typically occupying the initial position in interrogatives. The only type where they remain in situ are echo questions. This fronting of the wh-constituents Radford (2009, 183-186) calls wh-movement.

Fronting is not bound to determiner phrases (DP) only. Several other phrases can be fronted. For example, in (11a-c) Radford (2009, 202-205) suggests that fronting of a prepositional phrase (PP) can be done in two possible ways. The first kind of fronting of the PP can be seen “in formal English”, when “the superordinate preposition *to* is pied-piped” (i.e. dragged) “along with the interrogative pronoun *whom*, so that the whole PP *to whom* moves” to the initial position of the clause (11b). The second possibility is that *to whom* is fronted just partially. Radford argues, that “prepositions can be stranded (i.e. left behind) under wh-movement in informal styles of English”, e.g. in colloquial English. This is demonstrated in (11c) (Radford 2009, 202-205). (12) is example of fronting the adverbial phrase (AdvP).

- | | |
|--|-----------------|
| (11) a. <i>You were talking to whom?</i> | [PP in situ] |
| b. <i>To whom were you talking?</i> | [PP pied-piped] |
| c. <i>Who were you talking to?</i> | [PP stranded] |
| (12) <i>Where is he going tomorrow?</i> | |

3 DIRECT AND INDIRECT QUESTIONS

After the basic concepts of interrogatives formation have been introduced in chapter 2, it is necessary to elaborate specifically on how direct and indirect questions are formed.

The first subchapter describes formation of direct question in English and how their word order differs from declarative sentences and mentions usage by do-support constructions in clauses where no auxiliary is present. It distinguishes open and closed direct questions and demonstrates their syntactic differences.

The second subchapter deals with syntactic structure of indirect (embedded) questions. It begins with explaining that indirect questions serve as a complement to a verb in the clause and that they are marked by a wh-word or a complementizer *if* or *whether*. These structures contain a fronted wh-word that leaves a gap position, forming a wh-movement structure. However, this wh-movement does not trigger subject-auxiliary inversion as it is in open direct questions.

The last subchapter explains a syntactic mechanism known as T-to-C movement. It plays a crucial role in forming direct questions. In direct questions the auxiliary verb moves from T (head tense position) to C (head complementiser position), while the wh-word moves to the specifier position of CP. However, this movement does not apply in indirect questions. As a result, subject-auxiliary inversion in IQs is blocked and ungrammatical in standard English.

3.1 Direct Questions

This subchapter introduces the first type of interrogative constructions – direct questions. According to Baker (1995, 455-460), direct questions (DQ) differ from declarative sentences by usage of subject-auxiliary inversion (1a-b), or by usage of do-support construction (2a-b) when no auxiliary is present. In the case of closed interrogatives, examples (1-2) demonstrate how closed DQ are formed from declarative sentences (Baker 1995, 455-460).

- | | |
|--|---------------|
| (1) a. <i>She has finished her work.</i> | [DECLARATIVE] |
| b. <i>Has she finished her work?</i> | [CLOSED DQ] |
| (2) a. <i>He likes apples.</i> | [DECLARATIVE] |
| b. <i>Does he like apples?</i> | [CLOSED DQ] |

In contrast to closed DQ, open direct questions require more than a yes/no answer. They are introduced by wh-constituent that substitutes the whole phrase and that is fronted. They also apply subject-auxiliary inversion or do-support (3a-b), with one exception – the substitution of subject (4a-b) (Baker 1995, 455-460).

- | | |
|---------------------------------------|------------------------|
| (3) a. <i>She went to the cinema.</i> | [DECLARATIVE] |
| b. <i>Where did she go?</i> | [OPEN DQ] |
| (4) a. <i>Someone ate the cake.</i> | [DECLARATIVE] |
| b. <i>Who ate the cake?</i> | [OPEN DQ, Ø INVERSION] |

Direct questions function as independent clauses, consisting of subject and predicate, as shown in all examples (3-4) above. They apply subject-auxiliary inversion, except when the wh-constituent substitutes the subject. Sentences can consist of two or more coordinate clauses, all of them serving as direct questions (5) (Baker 1995, 455-459).

- (5) *Did you call you, or did she call you?*

3.2 Indirect Questions

Having outlined the formation of direct question, this section focuses on syntactic structure of indirect questions. They resemble direct questions but are syntactically quite different.

The first major difference is that indirect questions are introduced by a clause that contains a verb that requires a complement (Huddleston and Pullum 2002, 949), i.e. these verbs need to be combined with another constituent to be grammatical (Koenenman and Zeijlstra 2017, 55-57). Examples (6a-b) demonstrate the verb *said* that requires a complement, in this case a DP.

- | | |
|-------------------------------|------|
| (6) a. <i>He said [that].</i> | [DP] |
| b. <i>*He said.</i> | |

An indirect question is always indicated by presence of verbs in a clause, that must be completed by another phrase, and according to Baker (1995, 103), all indirect questions “begin with phrases that contain what we call “wh-words”” or complementizers *if* or *whether* (Baker 1995, 103).

Examples of such verbs are: *ask, wonder, know, remember, tell, predict*, etc. (Huddleston and Pullum 2002, 976)

For instance, in (7a) the verb *wonder* indicates, that another phrase must follow, otherwise the sentence would be ungrammatical (7b). Huddleston and Pullum (2002, 949) explain, that *[what she said]* clause is subordinate to *[I wonder]* clause as it functions as a complement to *wonder* (Huddleston and Pullum 2002, 949).

(7) a. *I wonder [what she said].* [CI]

b. **I wonder.* [Ø]

In (7a), the clause *[what she said]* resembles an echo question with fronted wh-word. Huddleston and Pullum (2002, 981) explain this resemblance as the clause reporting “an illocutionary act of asking a question” (Huddleston and Pullum 2002, 981). The clause begins with wh-word *what* that substitutes the DP and is fronted. Therefore, this fronted wh-words leaves an empty DP position, represented by a gap ($_\$), withing $_{VP}[she\ said\ [_\]]$ (8) and it demonstrates that the clause is indeed a question, as the complement of *say* is being asked (Baker 1995, 103-104). Furthermore, Huddleston and Pullum (2002, 914) suggest that *[she said]* “clause is embedded withing the interrogative clause” and that the “relation between what and is unbounded”, meaning “there is no grammatical limit on how many clause boundaries may separate them” (Huddleston and Pullum 2002, 914). This utterance in (8) shows wh-movement, yet in case of indirect questions it does not cause subject-auxiliary inversion.

(8) *I wonder [what [she said [_\]]].* [DP [VP [ØDP]]]

Baker (1995, 109) explores the complementation of verbs even further. Aside from interrogative DPs (8), there are several other phrases that be questioned. When questioning an object of preposition, two variants of word order are possible. Either the $_{PP}[to\ whom]$ is fronted and leaves an empty position at the end (9a), or the preposition maintain in its original position and only the wh-word is fronted (9b) (Baker 1995, 109).

(9) a. *She asked [to whom he was speaking _\].* [PP pied-piped]

b. *She asked [who he was speaking to _\].* [PP stranded]

Examples (10a-d) show other types of phrases that are questioned in indirect questions:

(10) a. *I'm not sure when they start .* [AdvP]

b. *He wondered how angry she was.* [AdjP]

c. *They explained why the plan had failed.* [Clause]

d. I know what she did.

[VP]

As it is with DQ, indirect questions can also be open and closed. The examples (6-10) explain open IQ. The case, where IQ can only have yes/no answer or a choice between alternatives is when they are marked with *whether* or *if* (Huddleston and Pullum 2002, 973).

Baker (1995, 112) notes that these questions “do not contain missing phrases” and that “some of them correspond to direct questions that call for yes-or-no answers” (11a) while another type calls “for a choice between alternatives” (11b) (Baker 1995, 112).

(11) *a. I don't know whether/if she is coming.*

*b. She asked whether/*if I wanted tea or coffee.*

Huddleston and Pullum (2002, 973) refer to the choice between more alternatives in (11b) as “the exhaustive conditional construction”, where only *whether* can be used (Huddleston and Pullum 2002, 973).

Another key difference that distinguishes IQ from DQ is that they lack subject-auxiliary inversion. This is related to the fact, that IQs are embedded in a larger sentence, preceded by a clause where the verb requires complementation. Syntactically they resemble declaratives, but their aim is to elicit information (Huddleston and Pullum 2002, 972). In case of (12b), the verb *wonder* requires a complementation of an embedded question [*what he said*].

(12) *a. What did he say?*

b. I wonder what he said.

During transformation from DQ (12a) into IQ (12b), Huddleston and Pullum (2002, 95) argue that subject-auxiliary inversion is blocked because it is “restricted to main clauses” (Huddleston and Pullum 2002, 95).

3.3 T-to-C Movement

Another syntactic operation involved in the formation of DQ with subject auxiliary inversion is T-to-C movement. However, in the case of IQ, subject-auxiliary inversion does not occur since T-to-C movement is blocked. This subchapter aims to explain the mechanism of T-to-C movement and its implications in IQ.

Carnie (2021, 217-218) explains embedded clauses as a part of the main clause that are marked as tense phrase (TP). This embedded TP is part of even larger phrase –

complementizer phrase (CP) that is introduced by a complementizer (Carnie 2021, 217-218). CPs can be perceived as a complement to the verb of the main TP.

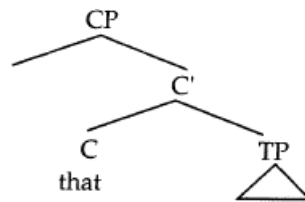


Figure 1 – Structure of CP

(source: Carnie 2021, 221)

CPs consist of C (the head), “a TP complement and an empty specifier position” (Carnie 2021, 221). This empty specifier position is crucial when considering IQs.

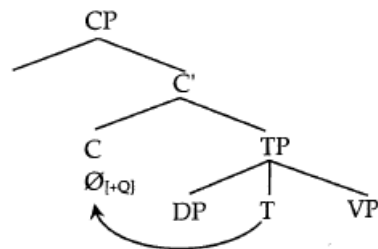


Figure 2 – Movement of Auxiliary

(source: Carnie 2021, 222)

Further (Figure 2), Carnie (2021, 222) demonstrates that a TP containing an auxiliary verb which is during subject-auxiliary inversion moved from the head position T within the TP to the head complementizer position (C) (Carnie 2021, 222).

However, this according to Radford (2009, 183-185) the C position can be occupied by an auxiliary only in the case of DQ.

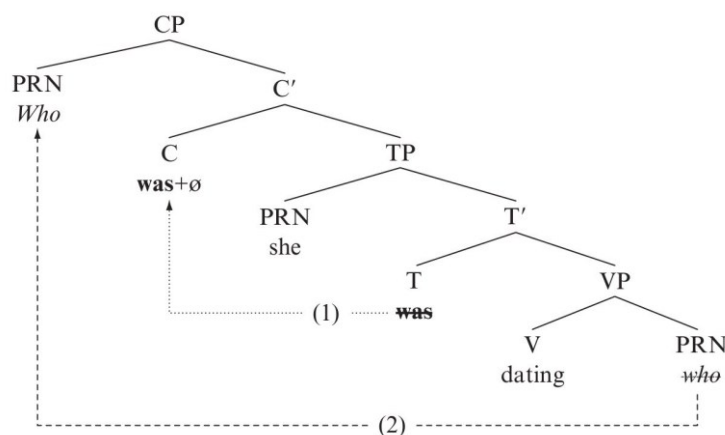


Figure 3 – Wh-Movement

(source: Radford 2009, 184)

Figure 3 shows the combination of the movements. The first movement involves *who* being fronted from its in situ to the initial position of the interrogative. The echo question *She was dating who?* becomes a direct open question with *who* fronted.

The second movement, where *was* changes its position from the head “tense (T) position of tense phrase (TP) to the head complementizer (C) position of complementizer phrase (CP)” is the T-to-C movement (Radford 2009, 184). This is caused by wh-movement, and it triggers subject-auxiliary inversion (i.e. T-to-C movement), giving us *Was she dating...?* (Radford 2009, 183-184).

The wh-word *who* is initially located in the verb phrase (VP), functioning as an object of *dating*. Then it is moved to the specifier position of CP. This results in a sentence *Who was she dating?*. Wh-movement is specific for creating interrogatives while the interrogative wh-word occupy the specifier position of the complementizer phrase, which is structurally required (Radford 2009, 183-185).

While Radford (2009, 184) illustrates how T-to-C movement functions in DQs, Rizzi (2000, 230-231) explains why this movement does not occur in IQs in standard English (13a). In DQs, the wh-words move to the specifier position of CP and auxiliaries move to C. Therefore, the Wh-Criterion proposed by Rizzi is fulfilled. However, this is not the case of IQs (13b) (Rizzi 2000, 230-231).

(13) a. *I wonder what he said.*

b. **I wonder what did he say?*

Rizzi (2000, 230-232) explains that subject-auxiliary inversion is blocked in IQs because the structure of embedded questions does not support T-to-C movement. The C position in IQs might be either filled with *if* or *whether*, or not structurally able to host auxiliary verbs (Rizzi 2000, 230-232).

4 INDIRECT QUESTIONS WITH DIRECT QUESTION WORD ORDER

This chapter focuses on the essential topic of this thesis – indirect questions that retain direct question word order.

The first subchapter explains that the word order of IQs typically differs from DQ, except one case – the *wh*-word occupying the subject position.

The second subchapter focuses on applying subject-auxiliary inversion in embedded questions. It focuses specifically on abundance of these sequences deviating from the standard word order in different dialects.

4.1 IQs with Subject Wh-Constituents

The word order of IQ normally differs from their direct counterparts as mentioned in the previous chapter. Direct questions involve subject-auxiliary inversion (or T-to-C movement). In contrast, IQs do not allow this movement, due to the fact, they are embedded clauses (Huddleston and Pullum 2002, 95).

Yet there is one case, when the word order of both, IQ and DQ, is the same – when the *wh*-word substitutes the subject. These examples (1a-c) show a subject substituted by a *wh*-word. There are only three possible *wh*-words that can be used – *what*, *which* + NP, and *who*.

- (1) a. *She asked what happened.*
 b. *They wanted to know which student failed the test.*
 c. *He wonders who sent the email.*

Huddleston and Pullum (2002, 972) explain when subjects are substituted by *wh*-words, the word order of either subordinate (2a) or main clause (2b) remains identical, applying no subject-auxiliary inversion. This also allows the clause to retain canonical word order even in interrogatives (Huddleston and Pullum 2002, 972).

- (2) a. *She asked [what happened].* [IQ]
 b. *[What happened]?* [DQ]

Normally, the word order of IQs and DQs differ in terms of T-to-C movement. However, when the *wh*-word itself represents the subject, no inversion happens in either form. This occurrence does not deviate from syntactic constraints.

4.2 Subject-Auxiliary Inversion in Indirect Questions

The standard English word order does not allow usage of subject-auxiliary inversion in embedded clauses. These clauses follow canonical (declarative) word order with the subject preceding the auxiliary. However, some IQs that contain *wh*-word in non-subject position may obtain DQ word order. From the perspective of syntax, these non-canonical occurrences in IQs are considered ungrammatical, yet they are still produced by the speakers on daily basis. This subchapter aims to explain the occurrence of this phenomenon in English.

Ronald Butters (1974, 230-234) describes occurrence of subject-auxiliary inversion in indirect questions, a feature that is common in African American English (AAE), but also it occurs across a variety of English dialects. Example (3a-b) shows a comparison to a sentence produced by AAE speaker (3b) compared to a standard English speaker (Butters 1974, 230-234).

(3) a. *Ask John how did he like the game.* (Butters 1974, 230)

b. *Ask John how he liked the game.* (Butters 1974, 230)

This phenomenon may be caused by generalization of syntactic rules – applying the same question forming strategy as it is in DQs, although this results in deviating from standard English. This pattern of applying the same syntactic rule for multiple situations is called Hypercorrection (Butters 1974, 233).

Butters (1974, 231) also suggests occurrence of IQs with DQ word order in Anglo-Irish dialects, and providing examples from English literature, yet arguing that “poets are free to break rules of normal syntax” (Butters 1974, 231).

The last important idea that Butters (1974, 232-233) mentions is that usage of this pattern is widely spread among American Southerners and Afro Americans, and it might be caused by “Scotch-Irish settlers in the American South [bringing] the dialect with them”, nowadays resulting in common daily usage (Butters 1974, 232-233).

Fought (2003, 98) discusses Chicago English dialect, that applies subject-auxiliary inversion in IQs as well. She provides authentic speech examples, e.g. *I don't know **what color are we**, but it doesn't matter* (Fought 2003, 98).

Henry (1995, 105-106) explains the difference between standard English (SE) and Belfast English (BE) – subject auxiliary inversion is unique to main clauses in SE but is common in Hiberno-English dialects (Anglo-Irish), including BE. He highlights that

inversion occurs not only in direct quotations (4a) but also in true IQs (4b) (Henry 1995, 105-106).

(4) a. *I wondered, “Is Bill going?”* (Henry 1995, 106)

b. *She asked who had I seen.* (Henry 1995, 106)

Although these structures deviate from standard English patterns, they might be considered grammatical withing a context of dialects.

II PRACTICAL PART

5 METHODOLOGY

The purpose of the analysis is to examine real-world usage of non-canonical indirect questions. The data used in the analysis are drawn from Corpus of Contemporary American English (COCA). Initially, I wanted to compare the data from COCA with data drawn from British National Corpus (BNC), yet the prevalence of indirect questions with direct question word order was marginal.

The COCA corpus was selected for its wide coverage of contemporary American English, including spoken, fiction, newspaper, magazine, and academic texts. Its balance of register makes it particularly suitable for investigating real-world grammatical phenomena, such as non-canonical word order in embedded *wh*-questions.

The analysis investigated indirect questions with direct question word order using a corpus-based study. The investigation focused on the frequency and patterns of specific verbs – *wonder*, *ask*, *find out*, *remember*, and *know* and their combinations with selected *wh*-words, (*where*, *when*, *why*, and *how*) and all forms of auxiliary verbs (*do*, *have*, *be*). These verbs, according to Huddleston and Pullum, often introduce embedded clauses (Huddleston and Pullum 2002,976). They regularly require complementation by a clause introduced by *wh*-word.

The data for the analysis were collected from Corpus of Contemporary American English (COCA). I accessed the Corpus through <https://www.english-corpora.org/corpora.asp>, where I selected COCA.

The survey procedure contained searching for the queries from the list that contain the selected verbs, the *wh*-word, and followed by either all form of the auxiliary *do* (*_vd*), *have* (*_vh*), or *be* (*_vb*). Due to the fact, that *wh*-words *what* and *who* mostly occupy the subject position in the embedded clause, they were excluded from the research as these are not the centre of my focus. Also, these results with *what* and *who* by typing the queries into the list search contained thousands of searched phrases, especially in combination with the verb *know*, yet mostly serving as subjects.

The classification of the results followed the searching procedure. The samples had to be validated to see, whether they were not just grammatical errors or if the queries were the required ones. They I created tables where I noted all 5 verbs and their combinations with the *wh*-words and auxiliaries. From these results only 5 small tables were created, representing the summaries. For instance, the verb *wonder* + the selected *wh*-words resulted

in 60 samples with the auxiliary *do*, 6 samples with *have*, and 91 samples with *be*. These summaries were compared to determine which auxiliary is the most abundant in embedded wh-clauses.

Furthermore, I conducted syntactic analyses. First, I focused thoroughly on the auxiliary *be* to determine whether it is used as a copula verb, in progressive aspect, or in the passive voice. I counted every relevant occurrence from the Corpus and classified it accordingly into these three categories. This also included dividing the auxiliary *be* according to its form (*is*, *'s*, *am*, *are*, *was*, *were*) to get the final numbers that I present. Every query was demonstrated by examples of all three functions of *be* (if it was found in COCA) and the subject-auxiliary inversion underlined.

The second analysis contained analysing *is* and its contracted form *'s*. This was done since when *'s* occurs after the wh-word, there is no other possibility than to create an indirect question with direct question word order. This was done to determine whether the speakers tend to use the full form *is*, or the contracted form *'s* more abundantly since it might be the main reason why these queries occur and if this conditions the overall usage of embedded wh-questions with non-canonical word order.

6 SYNTACTIC ANALYSES OF SELECTED VERBS

This chapter analyses the frequency of the selected verbs (*wonder*, *ask*, *find out*, *remember*, and *know*) and their usage in combination with *wh*-words and auxiliaries.

Secondarily, it deals with determining the syntactic function of *be* in embedded *wh*-questions with non-canonical word order, dividing them into three categories: copula verb, progressive aspect, and passive voice. Every sub-subchapter shows examples of such constructions. A purely quantitative analysis would be insufficient without considering the grammatical role that auxiliaries play within embedded structures, as the function may influence the likelihood of word order variation.

Finally, it determines whether the usage of these ungrammatical sequences is conditioned by the contracted form of *is* ('s).

6.1 Frequency of Verbs + Wh-Words + Auxiliaries

To determine the frequency of non-canonical word order in embedded questions, I searched in COCA for occurrences of verbs *wonder*, *ask*, *find out*, *remember* and *know* because these verbs, suggested by Huddleston and Pullum (2002, 976), often introduce such clauses. The following tables present the results and provide an overview of which combinations of the verbs followed by the chosen *wh*-words (*where*, *when*, *why*, and *how*) and an auxiliary (*do*, *have*, and *be*) and all their forms are the most common in American English.

6.1.1 Wonder

The table 1 presents the total numbers of indirect questions introduced by the verb *wonder* combined with the chosen *wh*-words and the auxiliary verbs *do*, *have*, and *be* in all forms. The most productive auxiliary is *be*, with a total of 91 samples across all its forms, followed by *do*, with a total of 60 samples. The least frequent is *have*, which appears only in 6 examples.

Table 1 – *wonder* + *where/when/why/how* + (AUXILIARY)

<i>wonder</i>	wh-word	(DO)	60
		(HAVE)	6
		(BE)	91

(source: own elaboration)

These results indicate that indirect questions containing the verb *wonder* most often involve the auxiliary *be*, while do-support structures are also common. In contrast, *have* is used only rarely in this type of embedded questions.

6.1.2 Ask

The table 2 presents the verb *ask*. It mostly combines with the auxiliary *be*, which occurs in 127 examples. The second most used auxiliary is *do*, with 91 samples, while *have* appears least frequently, with a total of 9 samples.

Table 2 – *ask* + *where/when/why/how* + (AUXILIARY)

<i>ask</i>	wh-word	(DO)	91
		(HAVE)	9
		(BE)	127

(source: own elaboration)

As it is with *wonder*, the auxiliary *be* is most productive, but *ask* also shows a high number of do-support constructions. The combination of *ask* + wh-word + *have* is relatively rare.

6.1.3 Find Out

The verb *find out* appears much less frequently in embedded questions compared to the other verbs analysed, with a total of 32 occurrences found. The auxiliary *be* is the most used auxiliary, with a total of 18 samples, followed by *do* (14 samples). The auxiliary *have* was not found in combination with *find out* and the wh-words in embedded questions in COCA database.

Table 3 – *find out* + *where/when/why/how* + (AUXILIARY)

<i>find out</i>	wh-word	(DO)	14
		(HAVE)	0
		(BE)	18

(source: own elaboration)

Although the number of all occurrences is low, the results still reflect the general pattern across the other verbs, with *be* being the most productive, and *do* the second most productive.

6.1.4 Remember

The table 4 presents the verb *remember* that occurs relatively rarely in this construction, with a total of only 19 examples found in COCA database. As with the other verbs, the most frequent is *be* (16 samples) and *do* (3 samples). The auxiliary *have* does not occur at all in this pattern in COCA.

Table 4 – *remember* + *where/when/why/how* + (AUXILIARY)

<i>remember</i>	wh-word	(DO)	3
		(HAVE)	0
		(BE)	16

(source: own elaboration)

Despite the low frequency, the results are in line with the general pattern observed for the other chosen verbs. This confirms the dominant role of *be* in embedded questions.

6.1.5 Know

The verb *know* is the most abundant in the COCA database, with a total of 416 samples in combination with wh-words and auxiliaries. The most frequent is *be* (237 samples), followed by *do* (173 samples). The auxiliary *have* appears only in 6 samples.

Table 5 – *know* + *where/when/why/how* + (AUXILIARY)

<i>know</i>	wh-word	(DO)	173
		(HAVE)	6
		(BE)	237

(source: own elaboration)

These results confirm that *know* is highly productive in indirect questions with direct question word order, and the general trend observed among all the verbs remains. *Be* is the most common auxiliary in embedded wh-questions, followed by *do*, while *have* is used only rarely.

6.2 Functions of the Auxiliary *Be*

Since *be* was the most frequent occurring auxiliary in my dataset, its grammatical function was examined in more detail. Each occurrence was analysed and classified as either a copula,

a progressive aspect auxiliary, or as a passive auxiliary. The following tables show how *be* is distributed across these three functions for each of the five chosen verbs.

The aim of the analysis is to determine whether the grammatical function of *be* influences the probability of non-canonical word order in embedded questions. Given that copular constructions, progressive aspect, and passive voice differ structurally, it is relevant to observe whether certain functions of *be* are more frequent associated with subject-auxiliary inversion. This section therefore provides a more detailed background for the subsequent interpretation of the data.

6.2.1 Wonder

The table 6 shows the distribution of the auxiliary *be* in indirect questions introduced by the verb *wonder*. The most frequent function is copula, which occurs 63 times. This is followed by progressive aspect, with 29 samples. The passive voice occurs relatively rarely in the dataset, with just 2 instances. These results suggest that *be* in combination with *wonder* + wh-words commonly functions as a linking (copula) verb, while progressive aspect is also relatively common. Passive structures are very rare in this context.

Table 6 – *wonder* + wh-word + COPULA/PROGRESSIVE/PASSIVE

<i>wonder</i> + wh-word	COPULA	63
<i>wonder</i> + wh-word	PROGRESSIVE	29
<i>wonder</i> + wh-word	PASSIVE	2

(source: own elaboration)

Example (1) contains *be* in form of copula, the second example (2) demonstrates the progressive aspect of *be*, and the last example (3) represents the passive voice with *be*. All three examples contain different wh-word.

(1) *And I wonder where are we now.* (COCA:2019:SPOK)

(2) *I wonder why are you only asking the questions but not providing the answers?*
(COCA:2012:WEB)

(3) *I wonder how am I supposed to understand [...].* (COCA:2014:NEWS)

6.2.2 Ask

In the case of verb *ask*, the auxiliary *be* functions mostly as a copula (81 instances). Progressive aspect occurs in 32 examples, while passive constructions are found in 14 cases.

Compared to *wonder*, passive voice occurs more often with the verb *ask*, although the occurrences of copulas and passive voice samples are relatively similar with *wonder* and *ask*. Progressive forms are again common and show consistent usage across the verb *ask*.

Table 7 – *ask* + wh-word + COPULA/PROGRESSIVE/PASSIVE

<i>ask</i> + wh-word	COPULA	81
<i>ask</i> + wh-word	PROGRESSIVE	32
<i>ask</i> + wh-word	PASSIVE	14

(source: own elaboration)

The following examples demonstrate the usage of copula (4), the progressive aspect (5), and the passive voice (6) in the query *ask* + wh-word + *be*.

(4) *I heard him ask where was Alice.* (COCA:2008:MOV)

(5) *I'm sure they didn't even ask how are the athletes doing?* (COCA:2018:SPOK)

(6) *Or I guess I should ask why was I born so long ago?* (COCA:2012:BLOG)

6.2.3 Find Out

In the case of *find out*, the auxiliary *be* most frequently functions as a copula, with 17 out of 18 occurrences. Only one example was classified as progressive, and no passive constructions were found in COCA. This verb shows the least variation in the function of *be*, with a strong preference for copular structures.

Table 8 – *find out* + wh-word + COPULA/PROGRESSIVE/PASSIVE

<i>find out</i> + wh-word	COPULA	17
<i>find out</i> + wh-word	PROGRESSIVE	1
<i>find out</i> + wh-word	PASSIVE	0

(source: own elaboration)

The verb *find out* shows the biggest compatibility with copula *be*, shown in example (7), and only one case of the progressive aspect, shown in example (8).

(7) *I'm supposed to find out when was the last time a play [...].* (COCA:2000:TV)

(8) *[...] time for me to find out why is he making her laugh so much?* (COCA:2011:TV)

6.2.4 Remember

The verb *remember* most frequently co-occurs with the auxiliary *be* functioning as a copula, with 15 out of 16 samples. Only one instance was classified as progressive, and there were no examples of passive constructions. Similarly as *find out*, *remember* shows a strong preference for copular structures and minimal variation in the function of *be*.

Table 9 – *remember* + wh-word + COPULA/PROGRESSIVE/PASSIVE

<i>remember</i> + wh-word	COPULA	15
<i>remember</i> + wh-word	PROGRESSIVE	1
<i>remember</i> + wh-word	PASSIVE	0

(source: own elaboration)

The following examples present the query *remember* + wh-word + *be* in copula function (9), and the progressive aspect (1). Interestingly, the example (9) contains two embedded wh-clauses, both applying subject-auxiliary inversion.

(9) *Do you remember when was your first kiss and how did you feel?*
(COCA:2012:WEB)

(10) *Remember why are you running for.* (COCA:2017:MOV)

6.2.5 Know

The verb *know* shows the highest number of occurrences among all analysed verbs. The most frequent function of *be* with *know* + wh-word is copula, with a total of 158 samples. Progressive aspect appears in 67 cases, and passive voice occurs in 12 instances. Compared to the other verbs, *know* shows the biggest functional diversity in its use of *be*, although the copula remains the dominant type.

Table 10 – *know* + wh-word + COPULA/PROGRESSIVE/PASSIVE

<i>know</i> + wh-word	COPULA	158
<i>know</i> + wh-word	PROG	67
<i>know</i> + wh-word	PASSS	12

(source: own elaboration)

The examples below (11-13) demonstrate the occurrence of copula *be* (11), the progressive aspect (12), and the passive voice (13) in the query *know* + wh-word + *be*.

(11) *And how does your friend know how are the things in the city?* (COCA:2008:MOV)

(12) [...] I'd just like to know when am I ever going to be perfect? (COCA:2012:WEB)

(13) You know why are they called community colleges? (COCA:1996:MOV)

6.3 Distribution of *Is* and *'s* (Contracted Form)

Since contractions such as *'s* (shortened form of *is*) may influence subject-auxiliary inversion in indirect questions, the following chapter compares the frequency of the full form *is* and its contracted form *'s*. The data are drawn from COCA and the comparison focuses on combinations with the selected verbs and the selected wh-words. The aim is to determine whether the contraction might play a role in the occurrences of non-canonical word order. The following tables present the overall distribution of the two forms across the relevant constructions and also the examples from COCA are provided.

6.3.1 Wonder

In indirect questions introduced by the verb *wonder*, the full form *is* appears more frequently than the contracted form *'s*. Out of the total of 56 examples, only 5 contain the contracted form. This suggests that speakers tend to avoid contractions in embedded questions with *wonder*. The contracted form was completely absent in combination with *why* and *how*.

Table 11 – *wonder* + wh-word + *is/'s*

<i>wonder</i> + wh-word	<i>is</i>	51
<i>wonder</i> + wh-word	<i>'s</i>	5

(source: own elaboration)

The example (1) presents a samples from COCA with the full form of *is*, and the second examples (2-3) presents the contracted form *'s* in combination with *where* (2) and *when* (3).

(1) I wonder where is Morgan. (COCA:2014:TV)

(2) [...] I wonder where 's that love you promised me? (COCA:2002:MOV)

(3) You wonder when 's the last train leaving this station? (COCA:2009:MAG)

In constructions introduced by the verb *wonder*, the full form *is* is much more frequent than the contracted form *'s*. Only a small number of contracted *is* was found.

6.3.2 Ask

The verb *ask* (table 13) shows preference for the full form of *is* as well. Out of 50 examples, 44 use *is*, while only 6 use the contracted form. Although contractions appear slightly more often here than with *wonder*, the full form remains dominant in this context. The contracted form appears mostly with *where* (4 samples), with *when* and *how* in one sample, and the combination with *why* shows no samples.

Table 12 – *ask* + wh-word + *is*/'s

<i>ask</i> + wh-word	<i>is</i>	44
<i>ask</i> + wh-word	's	6

(source: own elaboration)

The following examples (4-7) present the dominant full form – *is* (4), and the contracted form 's in the query *ask* + *how* + 's (5), since the contracted form was completely absent in combination with *wonder* + *how*. Furthermore, 's occurs also in combination with *where* (6) and *when* (7).

(4) *May I ask where is the fire?* (COCA:1998:MOV)

(5) *I was going to ask how 's married life, but I guess I know.* (COCA:1994:FIC)

(6) *As players we ask where 's the money going?* (COCA:2012:WEB)

(7) *I didn't ask when 's the last time she mentioned it.* (COCA:2010:TV)

The verb *ask* shows a slightly higher occurrence of contractions compared to *wonder*, but the full form *is* still remains dominant.

6.3.3 Find Out

Find out (table 14) appears less frequently than the other verbs. Therefore, the overall numbers are lower as well. Yet the tendency to use the full form of *is* is more common than the usage of the contracted form. The full form of *is* occurs with all four wh-words, while the contracted form occurs only with *where* (2 samples).

Table 13 – *find out* + wh-word + *is*/'s

<i>find out</i> + wh-word	<i>is</i>	9
<i>find out</i> + wh-word	's	2

(source: own elaboration)

The examples below show the full form of *is* (8) and the contracted form 's with *where* (9-10).

(8) [...] you try to find out where is the sweet spot here? (COCA:2010:SPOK)

(9) To find out where 's the best restaurant [...]. (COCA:2012:BLOG)

(10) And all the investors were calling Bayou to find out where 's the money, where's the money. (COCA:2009:SPOK)

Similarly, *find out* shows a clear preference for the full form *is*, with only a few examples of the contracted form 's.

6.3.4 Remember

Remember as a verb with the lowest prevalence in the database overall co-occurs with *is* in 4 cases. However, it does not occur with 's at all. It might be cause due to the lower overall frequency of this construction in COCA.

Table 14 – remember + wh-word + is/'s

remember + wh-word	is	4
remember + wh-word	's	0

(source: own elaboration)

The examples of query remember + wh-word + *is* (11-12):

(11) I think I remember where is it suddenly. (COCA:2012:TV)

(12) [...] I think it's important to remember why is this area important, [...].
(COCA:2019:SPOK)

In the case of *remember*, only the full form *is* was attested, with no examples of contraction in the COCA database.

6.3.5 Know

The word *know* displays the highest number of occurrences in this section. The full form *is* appears 109 times, while the contracted form 's occurs in 28 examples. Although *is* remains more frequent overall, *know* shows the highest relative use of the contracted form compared to the other verbs. This may be due to its frequent usage in all section of COCA. The full form of *is* occurs in combinations with all four wh-words, while the contracted form occurs in combination with *where*, *when*, and *how*.

Table 15 – *know* + wh-word + *is*/*'s*

<i>know</i> + wh-word	is	109
<i>know</i> + wh-word	s	28

(source: own elaboration)

The examples present samples from COCA, the query *know* + *where* + *is* that was the most abundant (13), and the query of *know* + wh-words + *'s* (14-16), except *why* that is not present.

(13) *Do we know where is this place?* (COCA:2017:MOV)

(14) *I know where 's the rooster.* (COCA:2011:TV)

(15) *You want to know when 's it gonna rain!* (COCA:1992:MOV)

(16) *We don't know how 's he dealing with the situation, [...].* (COCA:2012:BLOG)

The verb *know* shows the highest relative frequency of the contracted form *'s* among the analysed verbs. Even here, however, the full form *is* remains more frequent overall.

CONCLUSION

The aim of the thesis was to explore non-canonical word order in embedded questions, with a particular focus on the persistence of subject-auxiliary inversion in contexts where it should normally be blocked. The theoretical part provided an overview of standard and non-standard English word order in embedded questions, including the mechanism of T-to-C movement and the factors that typically prevent inversion in embedded interrogative clauses. This section also introduced key syntactic concepts necessary for understanding the word order patterns analysed later in the thesis.

The practical part was based on data from the Corpus of Contemporary American English (COCA). Five verbs (*wonder*, *ask*, *find out*, *remember*, and *know*) were analysed in context of indirect questions, in combination with selected wh-words (*where*, *when*, *why*, and *how*) and auxiliaries (*do*, *have*, and *be*). The data was collected to determine whether any of these combinations tend to preserve direct question word order, if so, under which circumstances.

The results showed that the auxiliary *be* was the most frequent in embedded clauses, followed by *do*, while *have* appeared only marginally. This distribution reflects the dominant role of *be* in forming indirect questions. The low number of *have* examples suggests that perfective aspect constructions are less typical in embedded wh-questions introduced by the selected verbs.

Further analysis of *be* revealed that its most common function was a copula, followed by progressive aspect, and the passive constructions were the least frequent. This suggests that embedded questions most often involve relatively simple linking structures rather than more complex verbal forms.

The comparison between the full form *is* and the contracted form *'s* indicated that the full form was generally preferred across all selected verbs, although *know* showed a noticeably higher number of contractions. This observation supports the idea that contracted form *'s* might correlate with non-canonical word order as it might remind the speakers of a direct question and in informal, fast spoken English, people may not always analyse the structure of a sentence they produced carefully, which can lead to preservation of subject-auxiliary inversion in contexts where standard grammar would block it.

Overall, the results suggest that non-canonical word order in embedded wh-questions is influenced by several factors, including the choice of main verbs, the type of auxiliary used,

the grammatical function of *be*, and the presence or absence of contractions. Although canonical (declarative) word order remains the standard in embedded questions, the data confirms that variation exists that this deviation is naturally used by speakers, especially in everyday communication, especially in spoken or informal register.

Thus, the practical part highlights that indirect questions with direct question word order represent a syntactic area where the boundaries of canonical and non-canonical word order can be blurred. This is particularly observable in cases where the contracted auxiliary 's is present, reflecting the dynamic nature of English language use.

Future research could extend this study by including other matrix verbs or examining varieties of English outside the American context. In addition, further investigation into sociolinguistic factors influencing non-canonical word order could provide a more complete picture of this phenomenon.

BIBLIOGRAPHY

- Baker, C. L. 1995. *English Syntax*. 2nd ed. Cambridge, Mass: MIT Press.
- Birner, Betty J., and Gregory L. Ward. 1998. *Information Status and Noncanonical Word Order in English*. Studies in Language Companion Series 40. Amsterdam and Philadelphia: John Benjamins.
- Brinton, Laurel J., and Donna M. Brinton. 2010. *The Linguistic Structure of Modern English*. Amsterdam: John Benjamins.
- Butters, Ronald R. 1974. "On the Structure of English Indirect Questions." In *Current Progress in Linguistics: Proceedings of the Second International Conference on Historical Linguistics*, edited by William Labov, 153–167. Bloomington: Indiana University Linguistics Club.
- Carnie, Andrew. 2021. *Syntax: A Generative Introduction*. 4th ed. Introducing Linguistics. Chichester, West Sussex: Wiley-Blackwell.
- Davies, Mark. 2008. *The Corpus of Contemporary American English (COCA): 560 million words, 1990–present*. Provo, UT: Brigham Young University.
- Fought, Carmen. 2003. *Chicano English in Context*. New York: Palgrave Macmillan.
- Henry, Alison. 1995. *Belfast English and Standard English: Dialect Variation and Parameter Setting*. New York: Oxford University Press.
- Huddleston, Rodney, and Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Koenenman, Olaf, and Hedde Zeijlstra. 2017. *Introducing Syntax*. Cambridge: Cambridge University Press.
- Radford, Andrew. 2009. *Analysing English Sentences: A Minimalist Approach*. Cambridge: Cambridge University Press.
- Rizzi, Luigi. 2000. *Comparative Syntax and Language Acquisition*. London: Routledge.
- Siemund, Peter. 2017. "Interrogative Clauses in English and the Social Economics of Questions." *Journal of Pragmatics* 119 (October): 15–32.
- Sobin, Nicholas. 1990. "On the Syntax of English Echo Questions." *Lingua* 81 (2–3): 141–167.

LIST OF FIGURES

Figure 1 – Structure of CP	27
Figure 2 – Movement of Auxiliary	27
Figure 3 – Wh-Movement.....	28

LIST OF TABLES

Table 1 – <i>wonder</i> + <i>where/when/why/how</i> + (AUXILIARY)	36
Table 2 – <i>ask</i> + <i>where/when/why/how</i> + (AUXILIARY).....	37
Table 3 – <i>find out</i> + <i>where/when/why/how</i> + (AUXILIARY)	37
Table 4 – <i>remember</i> + <i>where/when/why/how</i> + (AUXILIARY)	38
Table 5 – <i>know</i> + <i>where/when/why/how</i> + (AUXILIARY)	38
Table 6 – <i>wonder</i> + <i>wh-word</i> + COPULA/PROGRESSIVE/PASSIVE	39
Table 7 – <i>ask</i> + <i>wh-word</i> + COPULA/PROGRESSIVE/PASSIVE	40
Table 8 – <i>find out</i> + <i>wh-word</i> + COPULA/PROGRESSIVE/PASSIVE	40
Table 9 – <i>remember</i> + <i>wh-word</i> + COPULA/PROGRESSIVE/PASSIVE	41
Table 10 – <i>know</i> + <i>wh-word</i> + COPULA/PROGRESSIVE/PASSIVE	41
Table 11 – <i>wonder</i> + <i>wh-word</i> + <i>is/'s</i>	42
Table 13 – <i>ask</i> + <i>wh-word</i> + <i>is/'s</i>	43
Table 14 – <i>find out</i> + <i>wh-word</i> + <i>is/'s</i>	43
Table 15 – <i>remember</i> + <i>wh-word</i> + <i>is/'s</i>	44
Table 16 – <i>know</i> + <i>wh-word</i> + <i>is/'s</i>	45

LIST OF ABBREVIATIONS

AAE – African American English

AdjP – adjective Phrase

Cl – clause

COCA – Corpus of Contemporary American English

C – complementizer

CP – complementizer phrase

DQ – direct question

DP – determiner phrase

IQ – indirect question

NP – noun phrase

PP – prepositional phrase

T – tense

TP – tense phrase

VP – verb phrase

