SYNCRETISM

A CASE STUDY OF THE PARTICLE ŽE IN CZECH

Jiri Kaspar

University College London

2017

Abstract

This thesis presents a case study of *že* in Czech. The particle can appear in a number of seemingly unrelated constructions: it can be used in declaratives, interrogatives and exclamatives of various types. Below, three of its uses are considered in detail: subordinating conjunction, question tag, discourse particle.

When used as a subordination conjunction or a question tag, *že* might lexicalise one of the following functional heads: C or I. In both cases, *že* functions as a semantically vacuous linker, the purpose of which is to mark a relation between a head of one extended projection and some dependent extended projection. As a subordinating conjunction, it links a matrix clause to an embedded clause. As a question tag, it links a matrix clause to a minimal answer.

When used as a discourse particle, *že* might appear in C or SpecCP. In both cases, *že* functions as a a semantically contenful linker, the purpose of which is to mark a relation between its containing utterance and some other utterance that is part of the previous discourse.

In the process of establishing the above, I outline a number of properties of the constructions in which *že* appears, and I also make a number of general theoretical claims. First, I argue that contrastive topics in Czech appear in a dedicated syntactic position. Second, I propose a new definition of the category contrastive topic. Third, I demonstrate that embedded clauses involving contrastive left dislocation encode an assertoric speech act. Fourth, I argue that question tags comprising *že* are embedded minimal answers to questions expressed by the clauses to which they are attached. Fifth, I claim that a mechanism akin to feature percolation must exist as a theoretical primitive. Sixth, I show experimentally that discourse is more granular than usually assumed in the realm of information structure.

Dedication

I would like to dedicate this work to Leon A. Dogan.

Declaration

I confirm that the work presented in this thesis is my own, and that where information has been obtained from other sources, this has been indicated in the text.

Jiri Kaspar, May 2017

Acknowledgements

First, I would like to thank Klaus Abels for guiding me before, during and after my PhD. His support went above and beyond the call of duty. Apart from teaching me a lot of syntax, Klaus managed to radically transform the way I think. Second, I would like to thank Malte Zimmermann for mentoring me during my stay in Germany and for helping me after my return to England. In addition to teaching me some semantics, Malte made me realise how important it is to strike the right balance between professional and personal life. Third, I would like to thank Radek Šimík for discussing my research on numerous occasions. Radek taught me the importance of considering data from all possible angles. Fourth, I would like to thank my wife for supporting me during and after my PhD. It would have been a great deal more difficult to achieve this without her help. Fifth, I would like to thank David Adger and Vieri Samek-Lodovici for examining my thesis. Their feedback will undoubtedly help me to improve my work further. Sixth, I would like to thank my computer for allowing me to do my research. We have been through all this together! Seventh, I would like to thank everyone else who helped me along the way.

Contents

Chapter 1

Introduction

1.1 Syncretism: Form and Function

In the realm of linguistics, syncretism is described as a phenomenon whereby one form serves more than one function. The term is almost exclusively encountered in the literature on the syntax-morphology interface, and is therefore understood as meaning 'one morphological form serving more than one morphosyntactic function'. This thesis presents a case study of some of the linguistic properties of the Czech particle *že*, pronounced [ʒɛ][1], which is invariable in form and variable in function[2]. The research topic was motivated by the relatively low number of studies on syncretism at lexical level, and by the possibility of splitting two competing hypotheses about homonymy[3]: distributed morphology and nano syntax.

1.2 Etymology

According to the Czech Language Institute[4], the particle *že* originates from the Proto-Indo-European pronominal root morpheme *ghe*. This claim is supported by Hagstrom & McCoy (2003: p.201), who report exactly the same source of origin for Russian *že*. Furthermore, a brief look at the Proto-Indo-European Etymological Dictionary (2009), published by the Dnghu Association, reveals strong correlations between the letters <g> and <e> of Proto-Indo-European lexical items and the respective letters <ž> and <e> of etymologically-related lexical items currently

---

[1]In this thesis, the symbols used in phonetic transcriptions are those recommended by the International Phonetic Association.

[2]The particle can be either stressed or unstressed in connected speech. It can be realised as a free-standing morpheme, or else it can cliticise onto an immediately adjacent lexical item. The present text concentrates solely on its use as a (stressed or unstressed) free-standing morpheme.

[3]The term homonymy here is understood to mean 'a set of lexical items that are simultaneously homographs and homophones'.

[4]A significant amount of the information on the particle's etymology was obtained via electronic mail.

present and used in Slavic languages. In relation to this, it is important to note that the particle *že* is, in one form or another, widespread amongst all Slavic languages.

What is significant in terms of the development of its functional properties is that *že* was originally used in Proto-Slavic[5] as an emphatic suffix to the personal pronoun *jь*[6] ('it'), which had come about as a result of merging two Proto-Indo-European pronouns: demonstrative *is* ('this') and from-it-derived relative *qos* ('who'/'which') or *jos* ('that'/'the one that')[7]. Given that the Czech Language Institute did not reveal any information on the possible pronunciation of *jьže*, it can only be guessed that it was once pronounced as [jɪʒɛ][8]. The above findings about the particle's etymology can be summarised in the form of a graph[9].

(1)   An overview of the etymology of the particle *že.*
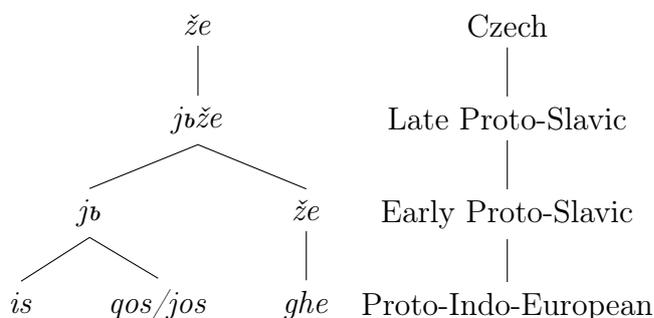
1.3   Functions

In Czech, *že* can be found in sentences with declarative, interrogative and exclamative force. The particle always appears at the left periphery of the clause. While it can often be preceded by a phrase-level category, it may never be preceded by a word-level category. The various functions that the particle has can be divided into the following categories: subordinating conjunction, question tag, discourse particle. This division is only tentative, and the terms selected to represent the

---

[5]Lexical items from this period are, similarly to those from the Proto-Indo-European period, written in Latin script, for they are products of linguistic reconstruction which are unattested in writing.

[6]This pronoun has the following grammatical categories: third (person), singular (number), neuter (gender) and accusative (case).

[7]The Czech Language Institute use *ios* as a term representing the relative pronoun which merged with the demonstrative pronoun *is*; however, there is no entry in the Proto-Indo-European Etymological Dictionary (2009) that corresponds to this lexical item. The two relative pronouns, *qos* and *jos* are the closest matching entries.

[8]This is an informed guess, as it is based on the information provided in Kortlandt (1983).

[9]The developmental stages are based on those outlined in Kortlandt (1983). However, it is important to bear in mind that there were other intervening stages on the way from Proto-Indo-European to Czech, or, for that matter, other Slavic languages.

three categories are not fully adequate. However, the thesis had to be organised into chapters, and the three-way division mentioned above offered itself as a viable solution. In what follows, each function of the particle is briefly outlined. The reader should bear in mind that this is a mere overview, and that each of the particle's functions is scrutinised in the following chapters[10].

1.3.1 Subordinating Conjunction

Perhaps the most frequently noted function of *že* is that of a subordinating conjunction. This use of the particle is abundant in both spoken and written Czech. As the translations below suggest, this type of *že* has a function similar to *that* in English[11].

(2) Petr   řekl,   že Marie   miluje   koblihy.
     Peter.NOM say.PST že Mary.NOM love.PRS doughnuts.ACC
     Peter said that Mary loves doughnuts.

The subordinate clause in the example above can undergo fronting. This is a means of deriving sentential subjects in Czech.

(3) Že Marie   miluje   koblihy,   Petr   řekl.
     že Mary.NOM love.PRS doughnuts.ACC Peter.NOM say.PST
     That Mary loves doughnuts, Peter said.

In addition, the subordinate clause containing *že* can also be conjoined with another clause containing *že*.

(4) Petr   řekl,   že Marie   miluje   koblihy,   a   že
     Peter.NOM say.PST že Mary.NOM love.PRS doughnuts.ACC and že
     Jakub   miluje   banány.
     Jacob.NOM love.PRS bananas.ACC
     Peter said that Mary loves doughnuts, and that Jacob loves bananas.

When *že* functions as a subordinating conjunction, it cannot be omitted without altering the interpretation of the sentence. When the particle is present, the subordinate clause is interpreted as indirect speech. However, when it is absent, the subordinate clause is interpreted as direct speech.

---

[10]Exclamatives are not considered in this thesis.

[11]The two lexical items do not exhibit identical behaviour. For instance, *that*, but not *že*, can be dropped under certain conditions without altering the interpretation of the sentence. In order not to confuse the reader, *že* is always glossed as 'že'.

1.3.2 Question Tag

In addition to the above, *že* can also function as a question tag. This use of the particle is found only in spoken Czech. There are two kinds of context in which this type of *že* can occur. First, each of the questions below can be used as a request for verification of the proposition expressed by the material preceding *že*. As the translations below suggest, this type of *že* has a function similar to *right* in English[12]. Such questions cannot be used to ask for a verification of something that has just been uttered.

(5) Petr   pojede do Brightonu,   že?
     Peter.NOM go.FUT to Brighton.GEN že
     Peter will go to Brighton, right?

(6) Petr   nepojede   do Brightonu,   že?
     Peter.NOM not-go.FUT to Brighton.GEN že
     Peter will not go to Brighton, right?

Second, each of the questions below can be used as an order to accept the proposition expressed by the material preceding *že*.

(7) Ty   mi   k narozeninám koupíš   počítač,   že?
     you.NOM me.DAT.CL for birthday.DAT buy.FUT computer.ACC že
     You will buy me a computer for birthday, right?

(8) Ty   mi   k narozeninám nekoupíš   ponožky,   že?
     you.NOM me.DAT.CL for birthday.DAT not-buy.FUT socks.ACC že
     You will not buy me socks for birthday, right?

Apart from following it, it is also possible for *že* to precede the clause. The clause-initial position of *že* does not change the licensing conditions on the use of verification questions. Such questions cannot be used to ask for a verification of something that has just been uttered[13].

(9) Že Petr   pojede   do Brightonu?
     že Peter.NOM go.FUT to Brighton.GEN
     Peter will go to Brighton, right?

---

[12]The two lexical items are not identical. For instance, *že*, but not *right*, can be combined with positive and negative polarity particles: *ano* ('yes') and *no* ('no'). In order not to confuse the reader, *že* is always glossed as 'že'.

[13]Note that when *že* functions as a discourse particle, it can also precede the entire proposition. However, the interpretation changes. The verification in such cases concerns what has been uttered in the immediately preceding discourse, which is not the case here. This particular function of *že* as a question tag is compatible with the question being interpreted as a request for a verification of the proposition that has not been introduced in the immediately preceding discourse.

(10) Že Petr   nepojede   do Brightonu?
      že Peter.NOM not-go.FUT to Brighton.GEN
      Peter will not go to Brighton, right?

Similarly, the constructions which function as commands remain felicitous under the same conditions even if *že* precedes the rest of the sentence.

(11) Že mi   k narozeninám koupíš   počítač?
      že me.DAT.CL for birthday.DAT buy.FUT computer.ACC
      You will buy me a computer for birthday, right?

(12) Že mi   k narozeninám nekoupíš   ponožky?
      že me.DAT.CL for birthday.DAT not-buy.FUT socks.ACC
      You will not buy me socks for birthday, right?

When *že* functions as a question tag, it cannot be omitted without altering the meaning of the sentence. When the particle is present, the sentence is interpreted as a tag question. However, when it is absent, the sentence is interpreted as a polar question.

1.3.3 Discourse Particle

Apart from the above, *že* can also function as a discourse particle. This use of *že* is found almost exclusively in spoken Czech. The following example shows *že* in a *wh*-question, which is perhaps the most commonly encountered construction in which this discourse particle appears.

(13) Kam že Petr   pojede?
      where že Peter.NOM go.FUT
      Where will Peter go?

The question above can be used in four types of context. In each case, the speaker relates the question to a proposition that has already been contributed to the discourse by the hearer. First, (13) can function as a request for repetition of some part of the proposition that the speaker did not hear. Such sentences function as echo questions (i.e., 'Peter will go where?'). Second, (13) can function as a request for a reminder of some part of the proposition that the speaker forgot. Such sentences function as reminder questions (i.e., 'Where was it again that Peter will go?') Third, (13) can function as a request for verification of some part of the proposition that the speaker is unsure of. Such sentences function as verification questions (i.e., 'Peter will go where?'). Fourth, (13) can function as an expression of surprise over some part of the proposition that the speaker did not expect. Such sentences function as surprise questions (i.e., 'Peter will go where?!').

It is also possible to find all these questions under subordination, provided that they are selected by a compatible predicate.

(14)   Jakub   chtěl   vědět,   kam   že Petr   pojede.
        Jacob.NOM want.PST know.INF where že Peter.NOM go.FUT
        Jacob wanted to know where Peter would go. (∼echo)

(15)   Jakub   chtěl   připomenout, kam   že Petr   pojede.
        Jacob.NOM want.PST remind.INF   where že Peter.NOM go.FUT
        Jacob wanted to be reminded where Peter would go. (∼reminder)

(16)   Jakub   si   nebyl   jistý, kam   že Petr   pojede.
        Jacob.NOM REFL.CL not-be.PST sure where že Peter.NOM go.FUT
        Jacob was not sure where Peter would go. (∼verification)

(17)   Jakub   byl   překvapený, kdo že mu   to klepe   na
        Jacob.NOM be.PST surprised   who že him.DAT.CL it knock.PRS on
        okno.
        window.ACC
        Jacob was surprised who was knocking on his window. (∼surprise)

In addition, the use of the particle in *wh*-questions partially overlaps with its use in non-*wh*-questions. Depending on the context, the question below may be interpreted either as a request for verification or as an expression of surprise.

(18)   Do Brightonu že Petr   pojede?
        to Brighton že Peter.NOM go.FUT
        Did you say that Peter will go to Brighton?

Importantly, the expression that is questioned (i.e., *do Brightonu*) does not have to undergo fronting.

(19)   Že Petr   pojede do   Brightonu?
        že Peter.NOM go.FUT where?
        Did you say that Peter will go to Brighton?

When *že* functions as a discourse particle, it can be omitted without altering the meaning of the sentence. Regardless of whether the particle is present or absent, the sentence can be interpreted as either an echo question, a reminder question, a verification question or a surprise question. It is worth noting that sentences with and without the particle are not equivalent in terms of their interpretive range. When the particle is present, the sentence cannot be interpreted as an information question. However, when it is absent, the sentence can be interpreted as a request for new information.

1.4   Information Structure and Constituent Order

This section considers the interaction of information structure and constituent order. Since the main aim is not to test the quality of the existing hypotheses about information structure and its notions, only those analyses that are felt to be most appropriate for the present purposes are outlined.

1.4.1   Information-Structural Categories

The following categories are commonly used in the literature on information structure: GIVEN (G), NEW (N), TOPIC (T), FOCUS (F), CONTRASTIVE TOPIC (CT), CONTRASTIVE FOCUS (CF). Often, linguists working on information structure differ as to which of these categories they consider to be the set of basic theoretical components. Although related, the task of splitting the various competing lines of thought is somewhat orthogonal to the present discussion. In what follows, the information-structural category of each relevant element is marked by a subscript[14].

1.4.2   Marked and Unmarked Constituent Order

The standard claim found in the literature is that Czech is an SVO language. The test that is commonly used to support this claim consists of a question-answer pair in which the question forces the answer to be widely focused. Considering the examples below, the entire sentence in (21) is focused, because it corresponds to the *wh*-element in the question in (20).

(20)   Co   je   nového?
       what AUX.CL new
       What's new?

(21)   [Petr   políbil Marii]$_F$.
       Peter.NOM kissed Mary.ACC
       Peter kissed MARY.

Any other order of the constituents in (21) would not be felicitous in the context at hand. It is, however, also possible to force a subpart of the sentence in (21) to be narrowly focused.

---

[14]Sometimes, elements that qualify as G are not marked as such. This is due to the fact that these elements might be interpreted as T or CT. Whenever G-marking is of importance, it is discussed. It is also worth noting that not everyone assumes the existence of the category T. Büring (to appear), for instance, dedicates an entire section to pointing out various problems with pinpointing a precise definition of T. However, the argument defended below does not hinge on the existence of this category.

(22)   Kdo políbil Marii?
        who kissed Mary.ACC
        Who kissed Mary?

(23)   [Marii   políbil]$_G$ [Petr]$_F$.
        Mary.ACC kissed   Peter.NOM
        PETER kissed Mary.

It follows that any deviation from the unmarked constituent order is reflected in the information-structural marking. Different constituent orders are compatible with different information-structural marking in different contexts. For this very reason, the context of use of each example below is carefully controlled.

1.4.3   Scrambling

Matthesius (1947) argues that emphasis is expressed by means of prosody (i.e., emphatic stress) and/or syntax (i.e., constituent order). One rather fitting pair of examples that the author gives is reproduced below (p.220).

(24)   To je dům,   kde [Neruda]$_G$ [ŽIL]$_F$.
        that is house.NOM where Neruda.NOM lived
        That is the house where Neruda LIVED.

(25)   To je dům,   kde [žil]$_G$ [NERUDA]$_F$.
        that is house.NOM where lived Neruda.NOM
        That is the house where NERUDA lived.

At the end of each sentence, there is a different word: *žil* in (24) and *Neruda* in (25). In each case, the sentence-final element receives emphatic stress. For Matthesius (1947), the necessity to reorder the two constituents (i.e., *Neruda* and *žil*) is driven by interpretation. The author does not use the terms 'given' and 'new', but rather states what context each sentence is compatible with. The sentence in (24) is compatible with the context in (26), and the sentence in (25) is compatible with the context in (27).

(26)   There is a tourist and a guide. The tourist wants to get to know the important sites of Neruda's life. (Neruda is a Czech poet.) The guide takes him to Nerudova street, points at the house where Neruda used to live, and utters (24).

(27)   There is a tourist and a guide. The tourist wants to get to know the memorable houses of Malá Strana. (Malá Strana is a district in Prague.) The guide takes him to Nerudova street, points at the house where Neruda used to live, and utters (25).

The contrast can be translated into the current terminology. In (24), *Neruda* is interpreted as 'given', because the discourse is about Neruda's life, not about memorable houses of Malá Strana. In (25), *Neruda* is interpreted as 'new', because the discourse is about memorable houses of Malá Strana, not about Neruda's life. What is crucial is that, given their respective contexts, the two sentences cannot be interchanged. Even though Mathesius did not formalise his analysis sufficiently, his insightful observation has been, in one form or another, mentioned since by other authors (e.g., Veselovská (1995), Lenertová (2001), Lenertová & Junghanns (2007), Kučerová (2007), Šimík & Wierzba (2015)).

Before I make any attempt to explain alternation patterns similar to those in (24) and (25), I need to specify what is meant by 'given' and 'new'. In the literature on information structure, the two labels are usually taken to be two sides of the same coin[15]. It follows that by defining the notion 'given' one also defines the notion 'new', and vice versa. Since the former seems to be easier to define, there have been many attempts to formalise what it means for an element to be 'given'. Šimík & Wierzba (2015), who provide a state-of-the-art account of data akin to those in (24) and (25), subscribe to the following semi-formalism proposed by Schwarzschild (1999: p.151).

(28)   DEFINITION OF GIVEN
        An utterance U counts as given iff it has a salient antecedent A and

        a.   if U is type e, then A and U corefer;
        b.   otherwise: modulo ∃-type shifting, A entails the existential F-closure of U.

The definition in (28) presents a dissociation of the presupposition of salience from the presupposition of truth. Consequently, definite NPs are not necessarily given, and given NPs are not necessarily definite. This is a desirable result. The clause (a) in the definition in (28) ensures that any two expressions whose referents belong to the domain of individuals and refer to the same individual are interpreted as 'given'. This might seem subtle, but it accounts for cases such as the one below.

(29)   The President of the United States of America ended his term. When asked what he would do next, [Barack Obama]_G expressed his desire to stay in politics.

It follows from (28) that *Barack Obama* counts as given in (29). This is so, because it has a salient antecedent that is of the same semantic type (i.e., type e): *The*

---

[15]There are some exceptions. For instance, Prince (1981) introduces more categories and Firbas (1966) considers givenness to be a kind of continuum.

President of the United States of America. The clause (b) in the definition in (28) ensures that expressions that are not of type e could also be interpreted as 'given'. The example below demonstrates that even entailed NPs can be given.

(30)   A: Did you see Peter's new house?
         B: Yeah. I loved [the kitchen]$_G$.

Following (28), *the kitchen* counts as 'given'. This is so, because $\exists x.house'\ (x)$ entails $\exists x.kitchen'\ (x)$. On this account of givenness, almost any type of linguistic expression can be taken to be 'given'. Assuming a basic dichotomy between 'given' and 'new', all elements that are not 'given' are 'new'.

(31)   DEFINITION OF NEW
         An utterance U counts as new iff it is not given.

Junghanns (2001) claims that constituent order in Czech, and many other Slavic languages, is restricted by the following principle (p.331). The principle is claimed to hold for both clauses with maximal/wide focus and clauses with minimal/narrow focus. Consequently, regardless of whether the focus exponent is included in a maximally focused constituent or a minimally focused constituent, it is required to appear at the right periphery of the clause.

(32)   DEFAULT PRINCIPLE OF FOCUS REALISATION
         The focus exponent appears at the right periphery of the clause.

According to Šimík & Wierzba (2015), the interpretation of the rightmost element within the sentence as 'focused' follows from a combination of prosodic and information-structural restrictions. As far as prosody is concerned, Daneš (1957) notes the following (p.63).

(33)   STRESS ASSIGNMENT RULE
         a.   The intonation centre of a phonological phrase is located on the last
              accented element within that phonological phrase.
         b.   The intonation centre of an intonation phrase is located on the last
              accented element within that intonation phrase.

'The intonation centre' refers to distinct things in the two clauses in (33). In (a), it refers to phrasal stress. In (b), it refers to sentential stress. As such, it is closely related to the notion of prosodic prominence, which, according to Daneš (1957), is a composite of pitch and intensity. Šimík & Wierzba (2015) combine the stress assignment rule above with the following version of the nuclear stress rule proposed

in Truckenbrodt (2013: p.133)[16].

(34) NUCLEAR STRESS RULE
      Strengthen the rightmost phrasal stress in the intonation phrase.

Because a sentence usually corresponds to an intonation phrase, the combination of (33) and (34) ensures that, unless stress is shifted, any element in the sentence-final position will be realised with the sentence stress (i.e., it will contain the focus exponent). In order to account for word order alternations such as those exemplified in (24) and (25) above, Šimík & Wierzba (2015) use the following constraint inspired by a similar restriction proposed in Féry & Samek-Lodovici (2006).

(35) *STRESS-GIVEN
      A given expression does not realise sentence stress.

If a 'given' element appears in the sentence final position, either (34) or (35) gets violated. Shifting the stress from the final position results in (34) being violated. Realising stress on the given element results in (35) being violated. In such cases, the derivation can be saved by syntax. Perhaps the most commonly applied last resort operation is scrambling. Czech has been argued to involve A-scrambling. Sturgeon (2008) claims that scrambled elements target the middlefield, which she considers to be located between I and v. Since scrambled elements can freely interchange with vP-adjoined material, the author assumes that scrambling is essentially vP-adjunction.

The contrast between (24) and (25) might not be easily accounted for by the analysis proposed in Šimík & Wierzba (2015). The definition of givenness, directly adapted from Schwarzschild (1999) seems to require an overt realisation of the antecedent. One would have to claim that the discourse preceding the utterance of (24) and (25) comprised a mention of *Neruda's life* and *memorable houses of Malá Strana*, respectively. Admittedly, this does not seem to be arbitrary, as it is natural to have a reason to hire a tourist guide. The optimality-theoretic mechanism introduced in Šimík & Wierzba (2015) can therefore formally account for the contrast between (24) and (25). What remains controversial is the syntactic analysis proposed by Sturgeon (2008). The controversiality stems from the fact that the author argues that the verb moves from V to v, but not higher. Assuming that *Neruda* is base-generated in the canonical subject position (i.e., it merges with vP), *žil* would have to be located above vP in (25). The analysis proposed in

---

[16]Junghanns (2001), Truckenbrodt (2013), and Šimík & Wierzba (2015) all acknowledge the resemblance of the generalisations they use to the Nuclear Stress Rule proposed by Chomsky & Halle (1968).

Sturgeon (2008) thus does not seem to be compatible with the attested data.

1.4.4   Rightward Backgrounding

In adherence to (32), or the combination of (33) and (34), the questions in (36) and (38) may be answered by (37) and (39), respectively. The classification of the relevant element as F follows one of the most basic principles of the question-answer congruence: the element of the answer that corresponds to the *wh*-element of the question is interpreted as F (i.e., it is F-marked).

(36)   Kdo ti   uvařil?
       who you.DAT.CL cooked
       Who cooked for you?

(37)   [Uvařil mi]$_G$ [PETR]$_F$.
       cooked me.DAT.CL Peter.NOM
       PETER cooked for me.

(38)   Co   ti   uvařil?
       what you.DAT.CL cooked
       What did he cook for you?

(39)   [Uvařil mi]$_G$ [FAZOLE]$_F$.
       cooked me.DAT.CL fazole.ACC
       He cooked BEANS for me.

However, as Junghanns (2001) notes, the constituent containing the focus exponent may also appear at the left edge of the clause, in an apparent violation of (32). The sentence in (40) can serve as an answer to the question in (36), and the sentence in (41) can serve as an answer to the question in (38).

(40)   [PETR]$_F$ [mi   uvařil]$_G$.
       Peter.NOM me.DAT.CL cooked
       PETER cooked for me.

(41)   [FAZOLE]$_F$ [mi   uvařil]$_G$.
       beans.ACC me.DAT.CL cooked
       He cooked BEANS for me.

Junghanns (2001) argues that the syntactic operation that allows the narrowly focused constituents in (40) and (41) to appear clause-initially consists of three steps. First, a copy of the entire clause is right-adjoined to its source. Second, everything except for the constituent containing the focus exponent is deleted in the source. Third, the constituent containing the focus exponent is deleted in the copy. The label given to this syntactic operation is 'rightward backgrounding'. The name reflects the fact that everything following the constituent containing the focus

exponent must be realised with low, flat intonation: a typical property of focus-background partitions. According to Junghanns (2001), 'rightward backgrounding excludes the possibility of a topic', and it 'is used when the speaker does not want to single out a topic' (p.336). The output of the process is depicted below.

(42)   An abstract representation of rightward backgrounding.
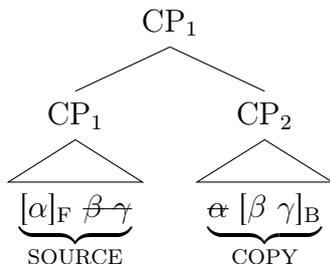
The analysis proposed by Junghanns (2001) ensures that the principle in (32) is not violated. Admittedly, examples such as (40) can be derived by means of stress shift. However, such a derivation would violate (32) and (34). Šimík & Wierzba (2015) mention the possibility of stress shift. The examples that they claim to involve it are marked as (more or less) degraded, but not unacceptable. This follows from their optimal-theoretic analysis, which assumes that (35) is ranked higher than (34). However, (40) and (41) are not even slightly degraded, which makes it hard to analyse them as involving stress shift. This is not to say that stress shift is not an available operation, but rather that it is not the only available operation. Apart from claiming that the operation depicted in (42) exists, Junghanns (2001) does not make any claims regarding its external syntax. In the absence of the evidence to the contrary, the question of whether this type of construction can be coordinated or subordinated, for instance, remains open.

1.4.5   Leftward Backgrounding

Apart from scrambling and rightward backgrounding, given material can be moved leftwards across focused material by means of A- or A'-movement. In adherence to (32), or the combination of (33) and (34), the question in (43) may be answered by (44). The classification of the relevant element as F follows one of the most basic principles of the question-answer congruence: the element of the answer that corresponds to the *wh*-element of the question is interpreted as F (i.e., it is F-marked).

(43)   Kdo uvařil   fazole?
        who cook.PST beans.ACC
        Who cooked the beans?

(44) [[Fazole]$_T$ uvařil]$_G$ [PETR]$_F$.
beans.ACC cook.PST Petr.NOM
The beans, PETER cooked.

According to Junghanns (2001), leftward backgrounding does not involve the mechanism that he claims is required for rightward backgrounding. A perceptive reader would have noticed that *fazole* in (44) is not only marked as G, but also as T. The notation reflects the assumption that, in contradistinction to rightward backgrounding, leftward backgrounding involves topicalisation[17]. The output of the process is depicted below.

(45) An abstract representation of leftward backgrounding.
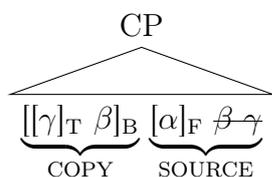
Again, apart from claiming that the operation depicted in (45) exists, Junghanns (2001) does not make any claims regarding its external syntax. In the absence of the evidence to the contrary, the question of whether this type of construction can be coordinated or subordinated, for instance, remains open.

In addition to the above, Junghanns (2001) claims that there exist cases where leftward backgrounding feeds rightward backgrounding to derive constituent orders in which F appears in neither the left nor the right periphery of the clause. Radek Šimík (p.c.) notes that such constructions are attested in Czech, and can be derived by means of stress shift. However, as shifting the stress from the final position induces a violation of (34), such sentences are, at best, slightly degraded. Since

---

[17]It is hard to say anything concrete about the position of the verb in (44). The movement of the verb to I or C has been generally claimed to be motivated by the need to either support clitics or derive questions. However, clitics do not need to be supported by verbs and questions can be derived without verb movement. Similarly to many other languages, head movement in Czech seems to have little motivation syntactically. What it seems to be doing in (44) is essentially what scrambling does. Šimík & Wierzba (2015) consider scrambling in Czech to be a strategy to satisfy (34). According to Sturgeon (2008), scrambling in Czech is vP-adjunction. The fact that the verb can freely exchange positions with vP-adjoined adverbs gives some support to this analysis.

(i) Kdo (včera) uvařil (včera) fazole?
who yesterday cook.PST yesterday beans.ACC
Who cooked the beans yesterday?

(ii) [[Fazole]$_T$ (včera) uvařil (včera)]$_G$ [Petr]$_F$.
beans.ACC yesterday cook.PST yesterday Petr.NOM
The beans, PETER cooked yesterday.

the combination of rightward and leftward backgrounding proposed by Junghanns (2001) is meant to derive perfectly grammatical constructions, it cannot be assumed that this compound operation is available in Czech.

1.4.6   Question-Answer Congruence

In the literature on information structure, the following pairs of categories are often encountered: 'given'-'new'; 'focus'-'background'; 'topic'-'comment'. It might be tempting to try to reduce one category into another. Indeed, there have been attempts to do so. For instance, some authors assumed that 'given'='background', and concluded that 'new'='focus'. Büring (2007) proposes that focus should be identified via a question-answer test. The test consists of answering a particular *wh*-question. The intuition is that the constituent of the answer that corresponds to the *wh*-element of the question is F.

(46)   a.   Who did your father vote for?
         b.   He voted for [a friend of JONES's]$_F$.

In the example above, *a friend of Jones's* in the answer corresponds to *who* in the question. Therefore, it is marked as F. It is widely agreed that F-marking is reflected in the assignment of prominence. In (46-b), the most prominent accent falls on *Jones*, as it is the rightmost stressed element within its containing intonation phrase (i.e., the sentence). In light of this, consider the following.

(47)   a.   Who did Jones's father vote for?
         b.   He voted for [a FRIEND of Jones's]$_F$.

In the example above, *a friend of Jones's* in the answer corresponds to *who* in the question. Therefore, it is marked as F. However, this time, the most prominent accent falls on *friend*, not *Jones*. Büring (2007) uses this example to argue that the relation between focus and accent has to be accounted for independently from the question-answer test. In other words, the question-answer test alone cannot account for the difference distribution of prominance in (46-b) and (47-b). The questions of each question-answer pair above (i.e., (46) and (47)) are different, and they are meant to represent different contexts. One can try to explain the different patterns of prominence by givenness, or the lack thereof. In (46-b), both *friend* and *Jones* are 'new'. Since the most prominent accent in English is realised on the last stressed element, *Jones* is accented. In (47-b), on the other hand, only *friend* is 'new'. Since *Jones* is 'given', the stress shifts, and the most prominent accent is realised on *friend*. It might be tempting to try to account for the distribution of focus purely in terms of givenness. The most prominent accent could be said to be

realised on the rightmost element that is 'new'. However, as Büring (2007) notes, the following example would be wrongly predicted not to contain any accent, as all the expressions in the answer are 'given' by virtue of their being mentioned in the question.

(48)   a.   Who did Jones's father vote for?
   b.   He voted for [JONES]$_F$.

In the example above, the most prominent accent is realised on *Jones*, in spite of the fact that it is 'given'. Therefore, it seems that the best strategy is to determine focus by means of the question-answer test, and then use givenness to determine accentuation.

To formalise the intuition that the F-marked element in the answer corresponds to the *wh*-element in the question, Büring (2007) introduces the following two concepts: the ordinary meaning of the question (i.e., $[\![Q]\!]^o$), and the focus value of the answer (i.e., $[\![A]\!]^f$). The ordinary meaning of the question corresponds to a set of propositions obtained by substituting the *wh*-element by a variable, and the focus value of the answer corresponds to a set of propositions obtained by substituting the F-marked element by a variable. For instance, the ordinary meaning of (48-a) and the focus value of (48-b) is the following.

(49)   {Jones's father voted for x | x is an individual}

The question-answer congruence takes advantage of the possibility to represent questions as sets of propositions that correspond to all possible answers (Hamblin (1958)) or to all true answers (Karttunen (1977)). The idea is that an answer is congruent to a question if the set of propositions expressed by the answer is a subset of the set of propositions expressed by the question. Büring (2007: p.451) defines question-answer congruence as follows.

(50)   QUESTION-ANSWER CONGRUENCE (QAC)
   A is a felicitous answer to Q only if

   a. $[\![Q]\!]^o \subseteq [\![A]\!]^f$, and
   b.   there is no alternative focusing A' of A which has less F-markings and meets (3-a).

(3-a) ensures that the focus value of the answer is a subset of the ordinary meaning of the question. (3-b) ensures that the answer contains as few F-markings as is sufficient for the satisfaction of (3-a). The above is coupled with givenness, which may sometimes influence the placement of the accent within an F-marked element.

Due to the analytical complexity of sentences with wide focus, the following text considers only sentences with narrow focus. The question-answer test is used to determine the position of F and to introduce a partition between focus and background. The core assumption is that question-answer congruence must be satisfied at some level of discourse representation. More about this will be said below.

Chapter 2

Contrastive Topicalisation

2.1 Introduction

This chapter outlines a number of properties of contrastive topicalisation in Czech. This syntactic operation consists of A'-moving a phrase-level category into the specifier of IP, a position in which it can realise a rise accent typical of elements interpreted as CT. In certain cases, the particle *že* may either precede or follow the element interpreted as CT. To explain this and to at the same time maintain that SpecIP is the only position in which elements interpreted as CT may occur, it has to be assumed that *že* may appear in either I or C. This positional flexibility is not manifested in the first conjunct of a coordinate structure under subordination, because the clause containing the subordinating predicate has to be immediately adjacent to the head of the phrase it selects (i.e., *že*).

2.2 Question-Answer Pairs

Jackendoff (1972) uses question-answer pairs to demonstrate how the form of the question can influence the form of the answer. In his discussion of data from English, he distinguishes between an A-accent and a B-accent: the former corresponds to a fall accent and the latter to a fall-rise accent. Let us consider the following pair of examples taken from Jackendoff (1972: p.261)[1,2].

(1) a. Well, what about FRED? What did HE eat?
   b. [Fred]$_{CT}$ ate [the BEANS]$_F$.

---

[1]The example numbering has been altered and the CT- and F-marking has been added.

[2]In the discussion of (1) and (2) below, the first questions in (1-a) and (2-a) are disregarded. In addition, it is assumed that the relevant assignment functions assign the value *Fred* to the pronoun *he* in (1-a) and the value *the beans* to the pronoun *them* in (2-a). Hence, unless indicated otherwise, (1-a) and (2-a) in the following text refer to *What did Fred eat?* and *Who ate the beans?*, respectively.

(2)   a.   Well, what about the BEANS? Who ate THEM?
        b.   [FRED]$_F$ ate [the beans]$_{CT}$.

In (1-b), *Fred* is realised with a B-accent, while *beans* is realised with an A-accent. In (2-b), on the other hand, *Fred* is realised with an A-accent, while *beans* is realised with a B-accent. Büring (2003), who provides a formal account of the data in (1) and (2), refers to any constituent realised with the A-accent as F, and to any constituent realised with the B-accent as CT. The accentuation is therefore taken to be a criterion for classifying a constituent as either CT or F. In English, constituents marked as CT are realised with a fall-rise accent, and constituents marked as F are realised with a fall accent.

As far as Czech is concerned, Veselá et al. (2003), who studied a relatively large sample of spontaneous speech, report that native speakers tend to realise CT with a rise accent, T with a level accent and F with a fall accent[3]. In accord with the proposal spelled out in Büring (2003), the accentuation is taken to be a criterion for classifying a constituent as either CT or F.

In addition to considering the phonetic realisation of each relevant syntactic element, it is also crucial to consider the semantic and/or pragmatic import that it has. For instance, Jackendoff (1972) notes that (1-a) cannot be answered by (2-b), and that (2-a) cannot be answered by (1-b). This observation can be straightforwardly explained by Büring's (2007, p.451) definition of question-answer congruence below.

(3)   QUESTION-ANSWER CONGRUENCE (QAC)
        A is a felicitous answer to Q only if

        a. $[\![Q]\!]^o \subseteq [\![A]\!]^f$, and
        b.   there is no alternative focusing A' of A which has less F-markings and meets (3-a).

To keep the notation transparent, I use the symbols $<\cong>$ and $<\ncong>$ to mark congruence and non-congruence, respectively. In order to see why (1-a) $\cong$ (1-b) and (2-a) $\cong$ (2-b), but (1-a) $\ncong$ (2-b) and (2-a) $\ncong$ (1-b), it first needs to be established what $[\![\ ]\!]^o$ and $[\![\ ]\!]^f$ mean. Rooth (1992), who introduced this distinction, refers to $[\![\ ]\!]^o$ as the ordinary semantic value, and to $[\![\ ]\!]^f$ as the focus semantic value. In the definition of question-answer congruence above, $[\![Q]\!]^o$ refers to the ordinary semantic value of the question, and $[\![A]\!]^f$ to the focus semantic value of the answer.

---

[3]However, a number of native speakers I consulted claimed that it is often possible to realise CT with a level accent, which is otherwise typical of elements interpreted as T. In the absence of any convincing study that would empirically support this claim, I assume, following Veselá et al. (2003), that CT is realised with a rise accent.

The ordinary semantic values of the questions in (1-a) and (2-a) are obtained by substituting the *wh*-elements by variables.

(4) $[\![(1\text{-a})]\!]^o = [\![\text{What did Fred eat?}]\!]^o = \{\text{Fred ate y} \mid y \in D_e\}$

(5) $[\![(2\text{-a})]\!]^o = [\![\text{Who ate the beans?}]\!]^o = \{\text{x ate the beans} \mid x \in D_e\}$

The focus semantic values of the answers in (1-b) and (2-b) are obtained by substituting the focused elements by variables[4].

(6) $[\![(1\text{-b})]\!]^f = [\![\text{[Fred]}_{CT}\text{ ate [the BEANS]}_F.]\!]^f = \{\text{Fred ate y} \mid y \in D_e\}$

(7) $[\![(2\text{-b})]\!]^f = [\![\text{[FRED]}_F\text{ ate [the beans]}_{CT}.]\!]^f = \{\text{x ate the beans} \mid x \in D_e\}$

Under the definition of question-answer congruence in (3), (1-a) ≅ (1-b) and (2-a) ≅ (2-b), because both (3-a) and (3-b) are satisfied in each case. Jackendoff's (1972) observation that (1-a) ≇ (2-b) and (2-a) ≇ (1-b) also follows, because (3-a) is violated in each case.

Note that CT cannot be considered an instance of F. If it were, the focus semantic values of the answers in (1-b) and (2-b) would be identical (see (8) and (9) below). This would make the two answers indistinguishable in terms of congruence, which, in turn, would lead one to predict, contrary to fact, that they could be used interchangeably.

(8) $[\![(1\text{-b})]\!]^f = [\![\text{[Fred]}_F\text{ ate [the BEANS]}_F.]\!]^f = \{\text{x ate y} \mid x, y \in D_e\}$

(9) $[\![(2\text{-b})]\!]^f = [\![\text{[FRED]}_F\text{ ate [the beans]}_F.]\!]^f = \{\text{x ate y} \mid x, y \in D_e\}$

Luckily, the very definition of question-answer congruence above prevents treating CT as F in (1-b) and (2-b). While (8) and (9) satisfy (3-a), they violate (3-b)[5].

---

[4]One could object that the definite article *the* is not focused and therefore should not be part of what gets substituted by the variable *x*. It is standardly assumed that the element of the answer that correspond to the *wh*-element of the question is marked as F. What is marked by capital letters is stress rather than focus. The article *the* is therefore part of the element marked as F. It is not stressed, because determiners are usually not accented in English.

[5]If one attempted to treat the constituent marked as CT as an instance of F, one would run into a number of additional problems. Note, for instance, the contrast between the following pairs of examples taken from Büring (to appear, p.5).

(i)   Churchmoth recorded this song in EIGHTY-THREE.
    a.   And [Muckensturm]$_{CT}$ recorded it in [SEVENTY-TWO]$_F$.
    b.   And in [SEVENTY-TWO]$_F$, [Muckensturm]$_{CT}$ recorded it.

(ii)   Churchmoth wrote this song in EIGHTY-THREE.
    a.   No, [MUCKENSTURM]$_F$ wrote it in [SEVENTY-TWO]$_F$.
    b.   #No, in [SEVENTY-TWO]$_F$, [MUCKENSTURM]$_F$ wrote it.

What remains to be explained is what licenses the accentuation of the element marked as CT. Building on his previous work, Büring (to appear) formulates the following rule to account for the distribution of the category CT (pp.3-4).

(10)   CT-INTERPRETATION RULE (CIR)
        For a sentence S^{CT+F} to be felicitous, there must be at least one question meaning in S^{CT+F}'s CT-value which is

        a.   currently pertinent, and   PERTINENCE
        b.   logically independent of $[\![S^{CT+F}]\!]_O$, and   INDEPENDENCE
        c.   identifiable.   IDENTIFIABILITY

$[\![S^{CT+F}]\!]_O$ refers to the ordinary meaning of the sentence containing CT and F. It is used in juxtaposition with $[\![S^{CT+F}]\!]_F$ and $[\![S^{CT+F}]\!]_{CT}$, which refer to the sentence's F and CT alternatives, respectively. Since the CT alternatives will be of primary importance in what is to follow, it is worth considering Jackendoff's examples in the light of Büring's proposal. The CT alternatives for (1-b) and (2-b) are sets of questions of the form *What did x eat?* and *Who ate y?*, respectively. Crucially, the variable x must be replaced by an individual other than *Fred*, and the variable y must be replaced by a dish other than *the beans*. (10-a) ensures that the alternative is relevant in the given context, (10-b) ensures that it neither entails nor contradicts the ordinary meaning of the sentence containing CT and F, and (10-c) ensures that it is recognisable by the speaker and the addressee.

In the following sections, it will be shown that Büring's (to appear) proposal is too restrictive. More concretely, the mechanism by which CT alternatives are generated precludes (1-b) and (2-b) from being alternatives, because (1-b) cannot generate CT alternatives of the form *Who ate y?*, and (2-b) cannot generate alternatives of the form *What did x eat?*. This is at odds with the fact that it is possible to conjoin answers to these questions in Czech. In order to explain this possibility, it is assumed that the presence of CT indicates that the clause that contains it is a partial answer to a question that requires a multiple-pair answer (see (11)). This definition follows from the generalisation in (12), which is itself inspired by Kuno's (1982) observation that answers to multiple-pair questions contain an expression according to which the answer is sorted (i.e., a sortal key). The last ingredient of the analysis is that it must always be possible for the element interpreted as CT to realise a rise accent (see (13)). This can be seen as a language-specific filter that is operative in Czech, but that need not be operative in other languages.

(11)   CT-INTERPRETATION

The clause that contains CT is a partial answer to a question that requires a multiple-pair answer.

(12)   CT-PRESENCE

A partial answer to a question that requires a multiple-pair answer must contain an element interpreted as CT.

(13)   CT-REALISATION

The element interpreted as CT must have the possibility to realise a rise accent.

2.3 Discourse Trees

Building on the original proposal of Roberts (2012), Büring (2003) assumes that any given DISCOURSE (D) is hierarchically organised into a DISCOURSE TREE (D-tree). Each node within a D-tree represents either a declarative or an interrogative sentence, and is referred to as MOVE (M). In turn, each M corresponds to the syntactic representation of a given sentence, less any CT and/or F marking. Any terminal node is referred to as ANSWER (A), and any non-terminal node (with the exception of the root) as either QUESTION (Q) or SUB-QUESTION (SQ) or SUB-SUB-QUESTION (SSQ), and so *ad infinitum.*

(14)   An abstract example of a D-tree.
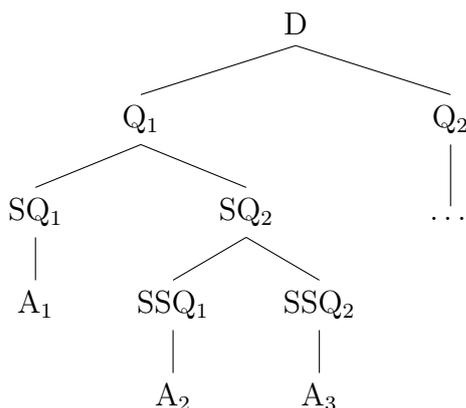
Importantly, the temporal unfolding is built into, or represented by, the hierarchical structure. The set of Ms of the D-tree above would be realised in the following sequence: Q_1-SQ_1-A_1-SQ_2-SSQ_1-A_2-SSQ_2-A_3-Q_2. To make the proposal clearer, let's consider a concrete example of D represented by (14).

(15)   A concrete example of a D-tree.

```
                                        D
                         _____
                        |                           |
              How was your day?              Will you marry me?
            _____                   |
           |                      |                   |
   Did you think about me?   How about your work?     …
           |                 _____
           |                |               |
      Yes, I did.       Were you busy?   Still there?
                            |               |
                            |               |
                        Mad busy!          Yes.
```
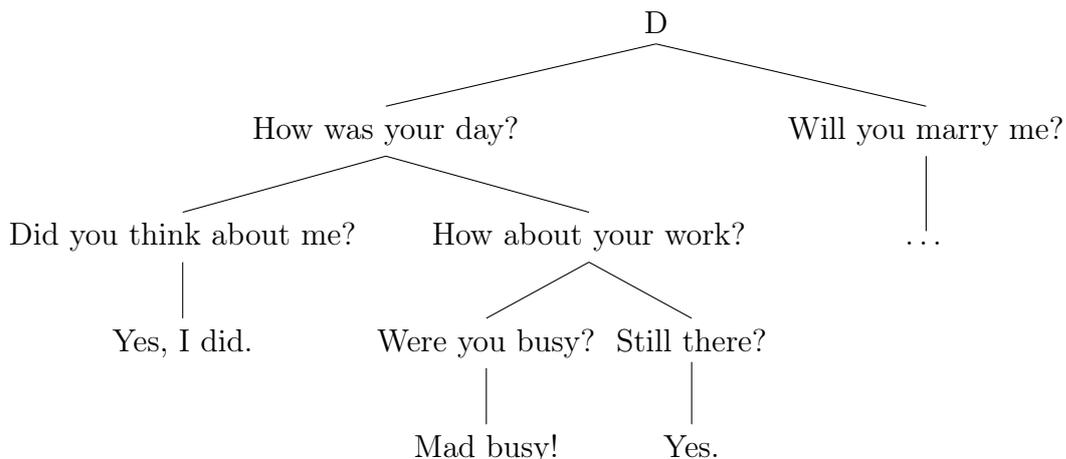
Fortunately, for the present purposes, there will be no need to build D-trees as big as the one above. However, it is very important that the reader gets familiar with the concept, which is essential for defining the category CT[6]. (For ease of exposure, the label D is left out from any and every D-tree below.)

2.4   Directionality: Anaphoric CT vs Cataphoric CT

The previous version of the rule in (10), proposed in Büring (2003), has often been wrongly interpreted as ensuring that the presence of CT indicates that there exists an alternative question that has been left unanswered.

(16) $[\text{Fred}]_{\text{CT}}$ ate $[\text{the BEANS}]_{\text{F}}$.

For (16) to be felicitous, there would have to be at least one alternative question meaning in the discourse that would meet the conditions set out in (10). However, (10) does not state whether the alternative should temporarily precede or follow the utterance of (16). Consequently, it is plausible to distinguish anaphoric and cataphoric CTs. The former are used in situations in which the alternative temporarily precedes the utterance of (16), while the latter in situations in which it temporarily follows it.

---

[6]For an in-depth overview of the notion D-tree, the reader is referred to Büring (2003).

(17)   A D-tree representing an anaphoric CT.

```
                        Who ate what?
              ┌──────────────────────────┐
        What did x eat?           What did FRED eat?
              ┊                           │
             ...                  [Fred]_CT ate [the BEANS]_F.
```
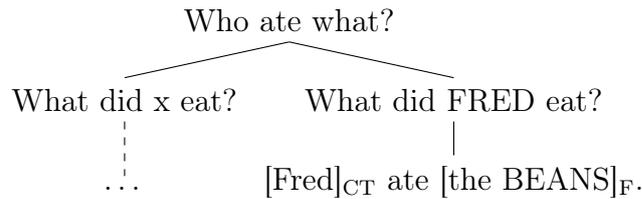
(18)   A D-tree representing a cataphoric CT.

```
                        Who ate what?
              ┌──────────────────────────┐
      What did FRED eat?            What did x eat?
              │                           ┊
     [Fred]_CT ate [the BEANS]_F.        ...
```
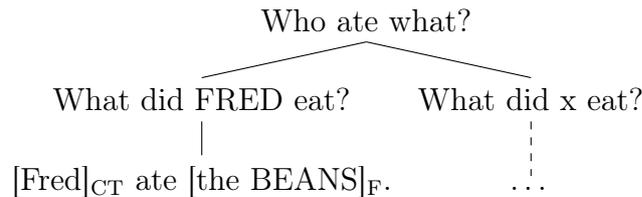
An anaphoric CT might indicate the presence of an alternative question that has already been answered. This is the case with the last answer to a multiple question. The sentence in (16) can serve as a partial answer to the question *Who ate what?* in a context in which there are two unique 'individual-dish' pairings between the set of individuals comprising, say, *Fred* and *Lucy*, and the set of dishes comprising, say, *the beans* and *the aubergine*. In such a context, the utterance of (16) may either precede or follow the utterance of *Lucy ate the aubergine*. Crucially, the use of CT does not give rise to the presupposition of incompleteness.

2.5   Basic Restrictions on Constituent Order

Czech is a language that is considered to have a very flexible constituent order. While this is generally true, the order of constituents is (often) heavily restricted by the context in which a given sentence is used. The information packaging ensures that each constituent gets interpreted in a particular way. The information structure then places certain (language-specific) restrictions on the order of these constituents. The aim of this section is to explore the nature of some of these restrictions.

2.5.1   Object-Oriented Questions and Simplex Answers

Assume that the speaker asks the following question[7]. In the answer, the subject is interpreted as CT, and the object as F[8].

(19)   A   co   PETR?   Co   snědl   TEN?
       and what Peter.NOM what.ACC eat.PST he.DEM
       And what about PETER? What did HE eat?

The following are all possible permutations of subject, verb and object that the addressee could produce in reply to the question in (19). Crucially, the subject can be interpreted as CT only in the initial position, where the rise accent can be realised.

(20)   [Petr]$_{CT}$ [snědl]$_G$ [FAZOLE]$_F$. (SVO)
       Peter.NOM eat.PST beans.ACC
       Peter ate the BEANS.

(21)   *[Petr]$_{CT}$ [FAZOLE]$_F$ [snědl]$_G$. (SOV)

(22)   *[Snědl]$_G$ [Petr]$_{CT}$ [FAZOLE]$_F$. (VSO)

(23)   *[Snědl]$_G$ [FAZOLE]$_F$ [Petr]$_{CT}$. (VOS)

(24)   *[FAZOLE]$_F$ [Petr]$_{CT}$ [snědl]$_G$. (OSV)

(25)   *[FAZOLE]$_F$ [snědl]$_G$ [Petr]$_{CT}$. (OVS)

Let's consider why (21)-(25) are not acceptable as answers to (19). (21) is not acceptable because *snědl*, which is marked as G, appears in the sentence-final position, which is typically reserved for elements marked as F. The notation indicates that stress shifted from this canonical position to *fazole*. Since elements marked as G scramble across elements marked as F, there is no motivation behind scrambling *fazole* across *snědl*, and shifting stress from *snědl* to *fazole*. (22) and (23) are not acceptable because the movement of *snědl* to the sentence-initial position is not warranted by anything. In Czech, the verb moves either when forming a *yes-no* question or in the absence of another suitable host for clitics. Neither (22) nor (23) is intended to be interpreted as a question, and neither (22)

---

[7]Unless specified otherwise, the questions used below are assumed to be uttered in the following context. Disregarding the speaker and the addressee, there were >2 individuals (Peter, Mary, Jacob) and >2 dishes (beans, spinach, aubergine), all of whom/which were familiar to the speaker and the addressee. The speaker did not know who ate what, and the addressee supplied this information. For reasons to do with simplicity, it is assumed that the relation between individuals and dishes is one-to-one. In other words, it is assumed that each individual is linked with exactly one dish. Unless stated otherwise, this is also the case in subsequent examples.

[8]This is due to the fact that *Petr* is being contrasted with the other individuals in the context, and that *fazole* corresponds to the *wh*-element in the question.

nor (23) contains any clitics that could license the movement of the verb. It is worth noting at this point that examples (21), (22) and (23) are not acceptable regardless of the accent that the subject realises. Example (24) can be derived via rightward backgrounding, and example (25) via a combination of rightward backgrounding and scrambling. In both cases, the subject could be realised with a level accent and interpreted as T[9]. However, interpreting *Petr* as T is at odds with what the context requires, so this alternative realisation is not ideal.

2.5.2 Subject-Oriented Questions and Simplex Answers

To check the reverse, assume that the speaker asks the following question. In the answer, the object is interpreted as CT, and the subject as F[10].

(26) A   co   FAZOLE? Kdo   snědl   TY?
      and what beans.NOM who.NOM eat.PST them.DEM
      And what about the BEANS? Who ate THEM?

The following are possible permutations of subject, verb and object that the addressee could produce in reply to the question in (26). Crucially, the object can be interpreted as CT only in the initial position, where the rise accent can be realised.

(27) *[PETR]$_F$ [snědl]$_G$ [fazole]$_{CT}$. (SVO)
      Peter.NOM eat.PST beans.ACC
      Peter ate the beans.

(28) *[PETR]$_F$ [fazole]$_{CT}$ [snědl]$_G$. (SOV)

(29) *[Snědl]$_G$ [PETR]$_F$ [fazole]$_{CT}$. (VSO)

(30) *[Snědl]$_G$ [fazole]$_{CT}$ [PETR]$_F$. (VOS)

(31) $^{??}$[Fazole]$_{CT}$ [PETR]$_F$ [snědl]$_G$. (OSV)

(32)   [Fazole]$_{CT}$ [snědl]$_G$ [PETR]$_F$. (OVS)

Let's consider why (27)-(31) are not acceptable as answers to (26). Example (27) can be derived via rightward backgrounding. The object could be realised with a level accent and interpreted as T[11]. However, interpreting *fazole* as T is at odds

---

[9]The subject does not have to be interpreted as CT because (24) and (25) both count as complete answers to the question in (19). Crucially, when the answer contains an element interpreted as CT, the question in (19) must be part of a strategy to answer some super-question (e.g., *Who ate what?*).

[10]This is due to the fact that *fazole* is being contrasted with the other dishes in the context, and that *Petr* corresponds to the *wh*-element in the question.

[11]The object does not have to be interpreted as CT because (27) counts as a complete answer to the question in (26). Crucially, when the answer contains an element interpreted as CT, the

with what the context requires, so this alternative realisation is not ideal. (28) can be derived via a combination of rightward backgrounding and scrambling. However, there is no motivation for scrambling *fazole* across *snědl*. (29) and (30) are not acceptable because the movement of *snědl* to the sentence-initial position is not warranted by anything. As already mentioned above, the verb moves either when forming a *yes-no* question or in the absence of another suitable host for clitics. Neither (29) nor (30) is intended to be interpreted as a question, and neither (29) nor (30) contains any clitics that could licence the movement of the verb. It is worth noting at this point that examples (28), (29) and (30) are not acceptable regardless of the accent that the object realises. (31) is degraded because *snědl*, which is interpreted as G, appears in the sentence-final position, which is typically reserved for elements marked as F. The notation indicates that stress shifted from this canonical position to *Petr*. Since elements marked as G scramble across elements marked as F, there is no motivation behind shifting stress from *snědl* to *Petr*, instead of scrambling *snědl* across *Petr*.

In principle, the speaker could select from three types of constituent order (i.e., SVO, OSV, OVS) when answering the question in (19), and from two types of constituent order (i.e., SVO, OVS) when answering the question in (26). However, the context requires the questions in (19) and (26) to be interpreted as sub-questions (i.e., *What did Peter eat?* and *Who ate the beans?*) of a common super-question (i.e., *Who ate what?*). Hence, an answer to (19) or (26) counts as a partial answer. According to (12), the answer must contain an element interpreted as CT. The rise accent typical of CTs is restricted to appear in an SVO configuration (see (20)) in the answer to (19), and in an OVS configuration (see (32)) in the answer to (26). The fact that the element interpreted as CT must precede the element interpreted as F is in line with similar observations made by other authors for other languages[12].

2.5.3 Object-Oriented Questions and Complex Answers

Apart from requesting information about a single individual or a single dish, it is also plausible to request information about multiple individuals or dishes at the same time. This can be achieved by coordinating two subjects or two objects in the question. Assume that the speaker asks the following question. In the answer, the subject is interpreted as CT, and the object as F.

---

question in (26) must be part of a strategy to answer some super-question (e.g., *Who ate what?*).

[12]Among others, Büring (1995) for German, and Wagner (2012) for German, Italian and English.

(33)   A   co   Petr   a   MARIE?   Co   snědli   TI?
        and what Peter.NOM and Mary.NOM what.ACC eat.PST they.DEM
        And what about Peter and MARY? What did THEY eat?

Interestingly, the only permissible constituent order within each partial answer
(i.e., conjunct) is one in which the subject can be realised with the rise accent.

(34)   [Petr]$_{CT}$ [snědl]$_G$ [FAZOLE]$_F$ (a   [Marie]$_{CT}$ [snědla]$_G$
        Peter.NOM eat.PST beans.ACC and Mary.NOM eat.PST
        [ŠPENÁT]$_F$). (SVO-SVO)
        spinach.ACC
        Peter ate the BEANS (and Mary ate the SPINACH).

The other two constituent orders (i.e., OSV and OVS) that could in principle be
used in an answer to the object-oriented question in (19) may not be used in either
a partial or a complete answer to (33). This is predicted by the combination of
(12) and (13) in the context of (33). More precisely, the following examples are
unacceptable, because the subject *Petr* cannot be interpreted as T.

(35)   *[FAZOLE]$_F$ [Petr]$_T$ [snědl]$_G$ (a [ŠPENÁT]$_F$ [Marie]$_T$ [snědla]$_F$). (OSV-
        OSV)

(36)   *[FAZOLE]$_F$ [snědl]$_G$ [Petr]$_T$ (a [ŠPENÁT]$_F$ [snědla]$_G$ [Marie]$_T$). (OVS-
        OVS)

2.5.4   Subject-Oriented Questions and Complex Answers

To check the reverse, assume that the speaker asks the following question. In the
answer, the object is interpreted as CT, and the subject as F.

(37)   A   co   fazole   a   ŠPENÁT?   Kdo   snědl
        and what beans.NOM and spinach.NOM who.NOM eat.PST
        TY?
        them.DEM
        And what about the beans and the SPINACH? Who ate THEM?

Interestingly, the only permissible constituent order within each partial answer
(i.e., conjunct) is one in which the object can be realised with the rise accent.

(38)   [Fazole]$_{CT}$ [snědl]$_G$ [PETR]$_F$ (a   [špenát]$_{CT}$ [snědla]$_G$
        beans.ACC eat.PST Peter.NOM and spinach.ACC eat.PST
        [MARIE]$_F$).
        Mary.NOM
        PETER ate the beans (and MARIE ate the spinach).

The other constituent order (i.e., SVO) that could in principle be used in an answer to the subject-oriented question in (26) may not be used in either a partial or a complete answer to (37)[13]. This is predicted by the combination of (12) and (13) in the context of (37). More precisely, the following examples are unacceptable, because the object *fazole* cannot be interpreted as T.

(39)   *[PETR]$_F$ [snědl]$_G$ [fazole]$_T$ (a [MARIE]$_F$ [snědla]$_G$ [špenát]$_T$).

Crucially, the overt realisation of either *Petr* and *Marie* or *fazole* and *špenát* in the question does not *per se* restrict the interpretation of the (corresponding) elements in the answer. While (38) would be infelicitous as an answer to (33), (34) would be felicitous as an answer to (37). Thus, the constituent of the answer that corresponds to the *wh*-element of the question does not have to be always interpreted as F, and the constituent of the answer that is primed by the question to be interpreted as CT does not always have to be interpreted as such.

There are many factors that need to be controlled when considering question-answer pairs such as (37)-(34). First, subjects tend to be better topics than objects. Second, the subject in each conjunct in (34) is animate, while the object is inanimate. Third, first names might be more easily associated with their referents than definite nouns. Given the complex interplay of these various factors, finding an explanation for the asymmetry mentioned in the previous paragraph is beyond the scope of this text. However, the mere observation that (34) is a felicitous answer to (37) casts doubt on the definition of question-answer congruence in (3). This is so, because there apparently is no direct mapping between the ordinary meaning of the question in (37) and the focus value of the answer in (34): the constituent marked as F in the answer does not correspond to the *wh*-element in the question. According to (3), (37) $\ncong$ (34), yet (34) is a felicitous answer to (37).

Perhaps, if one assumed that discourse could be restructured, then one could maintain the definition of question-answer congruence in (3). In essence, the answer in (34) would be congruent to (33). Of course, any restructuring would have to be restricted by the given discourse context.

2.5.5   Subordination

The acceptability judgments observed above for simplex and complex answers are preserved under subordination. The answers to questions in (19), (26), (33) and (37) can be embedded. Depending on its complexity (i.e., single-pair or multiple-

---

[13]The sentence in (39) is perfectly grammatical, and it would be acceptable if the interpretation of the subject and the object within each conjunct were reversed (see (34)). The possibility of restructuring the discourse to accommodate such reversals is considered further below.

pair), the answer could be inserted into the empty slot(s) in one of the following two templates. A single-pair answer could be inserted into the template in (40), and each conjunct of a multiple-pair answer could be inserted into the template in (41).

(40)   No. Jakub   řekl,   že _____ .
        well Jacob.NOM say.PST že
        Well. Jacob said that _____ .

(41)   No. Jakub   řekl,   že _____ , a   že _____ .
        well Jacob.NOM say.PST že   and že
        Well. Jacob said that _____ and that _____ .

What is crucial is that subordination admits only those constituent orders in which the element interpreted as CT (realised with either a level or a rise accent) precedes the constituent interpreted as F (realised with a fall accent). As a consequence, rightward backgrounding is not licensed under subordination. This means that an element interpreted as F may never appear in the left periphery of a subordinate clause. Subordination therefore places further restrictions on the order of constituents within the answer.

2.5.6   Multiple Questions: Single-Pair vs Multiple-Pair Answers

Czech is a language in which all *wh*-elements are typically fronted. For present purposes, it suffices to consider multiple questions with two *wh*-elements. Interestingly, the questions in (19) and (26) can form different strategies to answer the common super-questions below[14].

(42)   Kdo   co   snědl?
        who.NOM what.ACC eat.PST
        Who ate what?

(43)   Co   kdo   snědl?
        what.ACC who.NOM eat.PST
        Who ate what?

In an attempt to answer one of the above questions, the speaker can select from two different sortal keys. The answer might be ordered by 'individual' (see (44)) or by 'dish' (see (45)). Note that, in each case, one constituent of the answer that corresponds to a *wh*-element of the question is interpreted as CT. This follows

---

[14]Czech does not obey superiority, which means that *co* could precede *kdo* in the super-question. While this reversed order seems slightly preferred when the object is interpreted as CT in the answer, native speakers generally accept both orders.

from the generalisation in (12).

(44) $[\text{Petr}]_{\text{CT}}$ $[\text{snědl}]_{\text{G}}$ $[\text{FAZOLE}]_{\text{F}}$. (SVO)
       Peter.NOM eat.PST beans.ACC
       Peter ate the BEANS.

(45) $[\text{Fazole}]_{\text{CT}}$ $[\text{snědl}]_{\text{G}}$ $[\text{PETR}]_{\text{F}}$.   (OVS)
       beans.ACC eat.PST Peter.NOM
       PETER ate the beans.

The two different strategies are schematised below. The implicit sub-questions corresponding to (19), (26) and their alternatives are enclosed in brackets.

(46)   A discourse tree for a 'by-individual' strategy.

<div align="center">

Who ate what?

(What did PETER eat?)   (What did x eat?)

$[\text{Petr}]_{\text{CT}}$ $[\text{snědl}]_{\text{G}}$ $[\text{FAZOLE}]_{\text{F}}$.   . . .

</div>

(47)   A discourse tree for a 'by-dish' strategy.

<div align="center">

Who ate what?

(Who ate the BEANS?)   (Who ate y?)

$[\text{Fazole}]_{\text{CT}}$ $[\text{snědl}]_{\text{G}}$ $[\text{PETR}]_{\text{F}}$.   . . .

</div>

It is also plausible to provide one of the following sentences as an answer to (42) or (43).

(48) $[\text{Petr}]_{\text{CT}}$ $[\text{snědl}]_{\text{G}}$ $[\text{FAZOLE}]_{\text{F}}$ a $[\text{Marie}]_{\text{CT}}$ $[\text{snědla}]_{\text{G}}$
       Peter.NOM eat.PST beans.ACC and Mary.NOM eat.PST
       $[\text{ŠPENÁT}]_{\text{F}}$. (SVO-SVO)
       spinach.ACC
       Peter ate the BEANS and Mary ate the SPINACH.

(49) $[\text{Fazole}]_{\text{CT}}$ $[\text{snědl}]_{\text{G}}$ $[\text{PETR}]_{\text{F}}$ a $[\text{špenát}]_{\text{CT}}$ $[\text{snědla}]_{\text{G}}$
       beans.ACC eat.PST Peter.NOM and spinach.ACC eat.PST
       $[\text{MARIE}]_{\text{F}}$. (OVS-OVS)
       Mary.NOM
       PETER ate the beans and MARY ate the spinach.

The above can be neatly depicted in the form of D-trees. (The question mark next to the dashed line leading to one of the alternative questions indicates that its

presence is context-dependent.)

(50)   A discourse tree for a 'by-individual' strategy.
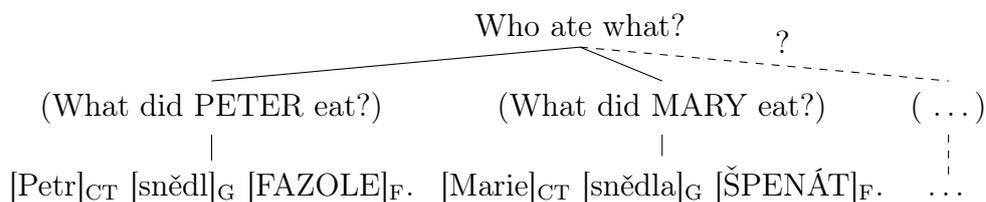
(51)   A discourse tree for a 'by-dish' strategy.
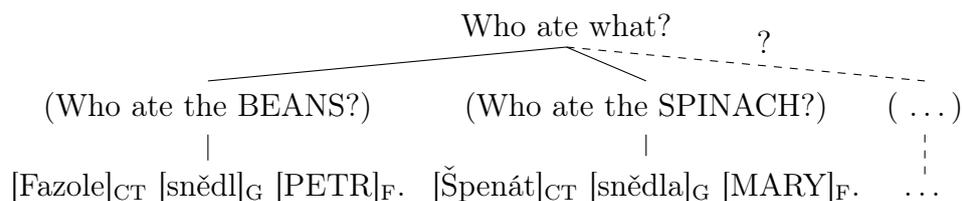
Regardless of which strategy is selected, the initial element of each partial answer may not be realised with a fall accent. This is a good indication that it is not interpreted as F.

2.5.7   Sortal Key Switching

It was shown above that a complex answer to (42) or (43) might be ordered by 'individual' (see (48)) or by 'dish' (see (49)). In addition, it is also possible to answer (42) or (43) by conjoining partial answers with different sortal keys (see (52) and (53))[15].

(52)   [Petr]$_{CT}$ [snědl]$_G$ [FAZOLE]$_F$ a   [špenát]$_{CT}$ [snědla]$_G$
         Peter.NOM eat.PST beans.ACC and spinach.ACC eat.PST
         [MARIE]$_F$.
         Mary.NOM
         Peter ate the BEANS and MARY ate the spinach.

(53)   [Fazole]$_{CT}$ [snědl]$_G$ [PETR]$_F$ a   [Marie]$_{CT}$ [snědla]$_G$ [ŠPENÁT]$_F$.
         beans.ACC eat.PST Peter.NOM and Mary.NOM eat.PST spinach.ACC
         PETER ate the beans and Mary ate the SPINACH.

Again, the use of the above examples can be captured in the form of D-trees.

---

[15]Given the combination of (12) and (13), the sortal key has to be different in each conjunct in (52) and (53).

(54)   A discourse tree for a 'by-individual-by-dish' strategy.

Who ate what?   ?

(What did PETER eat?)   (Who ate the SPINACH?)   (...)

[Petr]$_{CT}$ [snědl]$_G$ [FAZOLE]$_F$.   [Špenát]$_{CT}$ [snědla]$_G$ [MARY]$_F$.   ...

(55)   A discourse tree for a 'by-dish-by-individual' strategy.

Who ate what?   ?

(Who ate the BEANS?)   (What did MARY eat?)   (...)

[Fazole]$_{CT}$ [snědl]$_G$ [PETR]$_F$.   [Marie]$_{CT}$ [snědla]$_G$ [ŠPENÁT]$_F$.   ...

The possibility of changing the sortal key was noted in Wagner (2012), who argued against the analysis of parallel examples by Neeleman & van de Koot (2008) as involving a switch in the relative ordering of constituents interpreted as CT and F. The sortal key can be switched most easily if the question is general enough not to force the answer to follow the 'by-individual' or the 'by-dish' strategy. In other words, it would not be ideal to use (52) or (53) as an answer to (33) or (37)[16].

The fact that the sortal key can be switched poses problems for any analysis that imposes strict interpretive correspondence between the relevant elements of each partial answer. According to Büring (to appear), for instance, the CT alternatives for the two conjuncts in (52) would be *What did x eat?* and *Who ate y?*, respectively. However, the meanings of *What did x eat?* and *Who ate y?* are not compatible in the sense that the former cannot be taken to be an alternative for the latter, and *vice versa*. In other words, the additional layer of semantic embedding (that turns a set of propositions into a set of simple questions) prevents the propositions expressed by the two conjuncts in (52) from being alternatives. However, (11) does not place any such restrictions on the alternatives; on the contrary, the acceptability of (52) in the context of (42)/(43) is predicted. This is so, because the complete answers to *What did x eat?* and *Who ate y?* count as partial answers to (42)/(43).

Given the possibility of switching the sortal key, it could be proposed that the elements interpreted as CT must be 'given' in the sense of Schwarzschild (1999). This formulation of givenness encompasses clauses (10-a) and (10-c): only those elements that are in the consciousness of the addressee at the time of the utterance

---

[16]This cannot be due to the answer being incongruent to the question: as was noted above, (34) counts as a felicitous answer to (37).

count as 'pertinent' and 'identifiable'[17]. In the light of this, consider the following question[18].

(56)   Kdo se   kdy narodil?
       who REFL.CL when born
       Who was born when?

The addressee does not consider the two dates to be in the consciousness of the speaker at the time the reply is uttered. This explains why (57) can, and (58) cannot, serve as a felicitous answer to (56).

(57)   [Petr]$_{CT}$ [se   narodil]$_G$ [1. ŘÍJNA]$_F$ a   [Marie]$_{CT}$
       Peter.NOM REFL.CL born   1st October and Mary.NOM
       [se   narodila]$_G$ [31. LEDNA]$_F$.
       REFL.CL born   31st January
       Peter was born on the 1st of OCTOBER, and Mary was born on the 31st
       of JANUARY.

(58)   #[1. října]$_{CT}$ [se   narodil]$_G$ [PETR]$_F$ a   [31.   ledna]$_{CT}$
       1st October REFL.CL born   Peter.NOM and Mary.NOM REFL.CL
       [se   narodila]$_G$ [MARIE]$_F$.
       born 31st   January
       PETER was born on the 1st of October, and MARY was born on the 31st
       of January.

However, if the context comprised (a mention of) the possible dates of birth of the relevant individuals, then (58) would be a felicitous answer to (56). Büring (to appear) uses a similar example to show that there exists an asymmetry between CT and F. Nevertheless, it seems that what needs to be 'pertinent' and 'identifiable' is the element interpreted as CT rather than the alternative of the sentence that contains it. Thus, only the elements that might be interpreted as G might be interpreted as CT[19].

Clearly, Büring's (to appear) definition of CT is too restrictive, as it rules out some of the Czech examples that involve the sortal key switch. Equally clearly, my definition of CT is too permissive, as it does not rule out the English examples

---

[17](10-b) is an independent property of question-answer pairs. A partial answer to a question must neither entail nor contradict any other partial answer to that question. This holds independently of the analysis proposed here. It could, perhaps, be formalised using Grice's maxims.

[18]The question used below is assumed to be uttered in the following context. Disregarding the speaker and the addressee, there were >2 individuals (Peter, Mary, Jacob), all of whom were familiar to the speaker and the addressee. The speaker did not know who was born when, and the addressee supplied this information.

[19]Interestingly, syntactic elements such as time and manner adverbs, and sentential subjects cannot be (easily) interpreted as T. However, given the right context, all these elements can be interpreted as CT in Czech, because they can be easily contrasted with other similar elements.

that cannot involve the sortal key switch. Finding a unified analysis that would explain the data patterns in the two languages is beyond the scope of this text.

2.5.8   Position of *Že*

As was briefly mentioned above, topicalisation of the element interpreted as CT is readily available in subordinate clauses introduced by *že*. In general, the topicalised constituent may precede the particle in the second conjunct of an embedded coordinate structure. Either of the following two sentences can be used to answer (42) or (43)[20].

(59)   Jakub   řekl,   že [Petr]$_{CT}$ [snědl]$_G$ [FAZOLE]$_F$, a
       Jacob.NOM say.PST že Peter.NOM eat.PST beans.ACC   and
       [Marie]$_{CT}$ že [snědla]$_G$ [ŠPENÁT]$_F$. (SVO-SVO)
       Mary.NOM že eat.PST spinach.ACC
       Jacob said that Peter ate the BEANS, and that Mary ate the SPINACH.

(60)   Jakub   řekl,   že [fazole]$_{CT}$ [snědl]$_G$ [PETR]$_F$,   a
       Jacob.NOM say.PST že beans.ACC eat.PST Peter.NOM and
       [špenát]$_{CT}$ že [snědla]$_G$ [MARIE]$_F$. (OVS-OVS)
       spinach.ACC že eat.PST Mary.NOM
       Jacob said that PETER ate the beans, and that MARY ate the spinach.

The possibility of switching the sortal key remains available under subordination, and the element interpreted as CT in the second conjunct may precede *že*.

(61)   Jakub   řekl,   že [Petr]$_{CT}$ [snědl]$_G$ [FAZOLE]$_F$, a
       Jacob.NOM say.PST že Peter.NOM eat.PST beans.ACC   and
       [špenát]$_{CT}$ že [snědla]$_G$ [MARIE]$_F$.   (SVO-OVS)
       Mary.NOM že eat.PST spinach.ACC
       Jacob said that Peter ate the BEANS, and that MARY ate the spinach.

(62)   Jakub   řekl,   že [fazole]$_{CT}$ [snědl]$_G$ [PETR]$_F$,   a
       Jacob.NOM say.PST že beans.ACC eat.PST Peter.NOM and
       [Marie]$_{CT}$ že [snědla]$_G$ [ŠPENÁT]$_F$. (OVS-SVO)
       Mary.NOM že eat.PST spinach.ACC
       Jacob said that PETER ate the beans, and that Mary ate the SPINACH.

However, the element interpreted as CT is generally banned from appearing before *že* in the first conjunct, regardless of whether the element interpreted as CT in the second conjunct precedes or follows *že*. As Manfred Krifka (p.c.) points out, this

---

[20]It is assumed that the interaction takes place in the same context. Disregarding the speaker and the addressee, there were >2 individuals (Peter, Mary, Jacob) and >2 dishes (beans, spinach, aubergine), all of whom/which were familiar to the speaker and the addressee. The speaker did not know who ate what, and the addressee supplied this information.

might be due a requirement that *že* be immediately adjacent to the predicate that selects it[21].

(63) *Jakub   řekl,   [Petr]$_{CT}$ že [snědl]$_G$ [FAZOLE]$_F$, a
      Jacob.NOM say.PST Peter.NOM že eat.PST beans.ACC   and
      ([Marie]$_{CT}$) že ([Marie]$_{CT}$) [snědla]$_G$ [ŠPENÁT]$_F$. (SVO-SVO)
      Mary.NOM že Mary.NOM eat.PST   spinach.ACC
      Jacob said that Peter ate the BEANS, and that Mary ate the SPINACH.

(64) *Jakub   řekl,   [fazole]$_{CT}$ že [snědl]$_G$ [PETR]$_F$,   a
      Jacob.NOM say.PST beans.ACC že eat.PST Peter.NOM and
      ([špenát]$_{CT}$) že ([špenát]$_{CT}$) [snědla]$_G$ [MARIE]$_F$. (OSV-OSV)
      spinach.ACC že spinach.ACC eat.PST   Mary.NOM
      Jacob said that PETER ate the beans, and that MARY ate the spinach.

Interestingly, certain predicates that express some sort of emphasis are marginally compatible with the topicalised constituent in the first conjunct preceding *že*[22].

(65) ??Jakub   si   stěžoval,   [Marie]$_{CT}$ že [ho]$_G$
      Jacob.NOM REFL.CL complain.PST Mary.NOM že him.ACC.CL
      [NEMILUJE]$_F$, a   ([Lucie]$_{CT}$) že ([Lucie]$_{CT}$) [ho]$_G$
      not-love.PRS   and Lucy.NOM že Lucy.NOM him.DAT.CL
      [IGNORUJE]$_F$.
      ignore.PRS
      Jacob complained that Mary does not LOVE him, and that Lucy IGNORES
      him.

This type of topicalisation can also be found in sentences containing other left-peripheral particles. The following pair of examples shows this for the particle *aby*.

---

[21]This formulation of the restriction is not entirely correct. In certain cases, the embedding predicate is followed by additional syntactically and phonologically non-null material within its containing clause.

(i)   Jakub   řekl   Lucii,   (*[Petr]$_{CT}$) že ([Petr]$_{CT}$) [snědl]$_G$ [FAZOLE]$_F$ ...
      Jacob.NOM tell.PST Lucy.DAT Peter.NOM že Peter.NOM eat.PST beans.ACC
      Jacob told Lucy that Peter ate the BEANS ...

The restriction must reflect this. It seems that *že* must be immediately adjacent to the clause containing the predicate that selects it.

[22]Native speakers differ in the degree to which they accept the topicalised element to intervene between the subordinating predicate and the particle *že*. This movement operation results in strong markedness, which might explain why, even though not completely unacceptable, examples such as (65) are not productive in contemporary Czech.

(66) A   co   Petr   a   MARIE? Co   chtěl   Jakub,   aby
      and what Peter.NOM and Mary.NOM what want.PST Jacob.NOM aby
      snědli   TI?
      eat.PST they.DEM
      And what about Peter and MARY? What did Jacob want THEM to eat?

(67) Jakub   chtěl,   aby [Petr]_CT   [snědl]_G [FAZOLE]_F, a
      Jacob.NOM want.PST aby Peter.NOM eat.PST beans.ACC   and
      [Marie]_CT   aby [snědla]_G [ŠPENÁT]_F.
      Mary.NOM aby eat.PST spinach.ACC
   Jacob wanted Peter to eat the BEANS, and Mary to eat the SPINACH.

From the original set of examples, it might be tempting to conclude that *že* forces the topicalised constituent to be interpreted as CT. However, given the fact that *aby* can also be preceded by an element interpreted as CT, it remains unclear whether *že* places any semantic restrictions on the element that precedes it.

2.6   Formalism

2.6.1   Contrastive Topicalisation

Sturgeon (2008) assumes that the rise accent, typically associated with elements marked as CT, can be realised in SpecIP. While it remains an open question whether this is the only position in which it can be realised, the analysis proposed above is fully compatible with this assumption. Constant (2012; 2014) proposes that there is a functional projection high in the left periphery of the clause that is associated with elements interpreted as CT. At some point in the derivation, these elements must move (either overtly or covertly) into the specifier of this functional projection. Constant refers to this movement operation as 'topic abstraction'. In the absence of the evidence to the contrary, the present analysis assumes that the elements interpreted as CT move to SpecIP, where they have the possibility to realise the rise accent. Whether there are cases where this movement is covert remains to be seen.

(68)   An abstract representation of contrastive topicalisation.
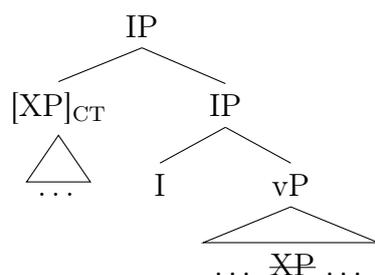
2.6.2 Coordination

Munn (1993) assumes that coordinate structures are hierarchical adjunct structures, and that, in case of subordination, only the first conjunct is selected by the subordinating predicate. Using this analysis, coordination in matrix clauses is considered here as an adjunction of the second conjunct to the first conjunct. Given the assumption that the element interpreted as CT moves to SpecIP, the adjunction may take place at the level of IP. This is schematised below[23].

(69)   An abstract representation of coordinate structure.
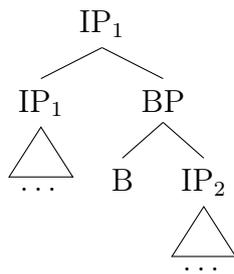
2.6.3 Contrastive Topicalisation under Subordination

Assuming the coordination structure above seems necessary, for, as was shown above, the topicalised constituent may precede *že* only in the second conjunct, provided that certain requirements to do with the licensing of CT are met. This asymmetry is expected if it is only the first conjunct that is selected by the subordinating predicate. The clause containing the subordinating predicate has to be immediately adjacent to the head of the phrase that it selects (i.e., *že*)[24]. The emerging picture is one where the subordinated coordinate structures such as (59) and (60) are represented as follows.

---

[23]BP stands for Bolean Phrase; B may host conjunctions (e.g., *a* ('and'), *ale* ('but')).
[24]As will be argued in the next section, the particle *že* can appear in C or I.

(70) A more detailed abstract representation of the coordinate structure with contrastive topicalisation.

```
                    CP
              ┌──────┴──────┐
              C            IP₁
              │        ┌────┴────┐
             že      IP₁        BP
                    ╱△╲      ┌───┴───┐
                    …        B      IP₂
                             │    ┌──┴──┐
                             a  [XP]_CT  IP₂
                               ╱△╲    ┌──┴──┐
                                …     I₂    vP
                                      │    ╱─┴──╲
                                     že   … XP …
```
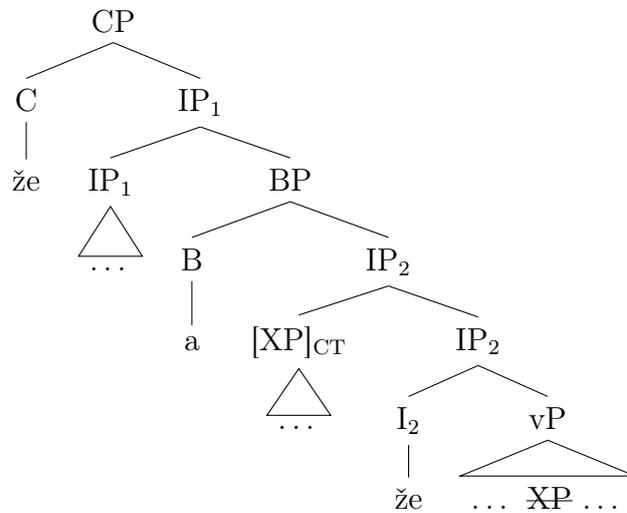
2.7 Conclusion

A number of conclusions can be drawn from the relatively large data sample presented above. Perhaps the most crucial one is that the distribution of the information-structural category CT is restricted by the following rules (repeated from above).

(11) CT-INTERPRETATION
      The expression that contains CT is a partial answer to a question that requires a multiple-pair answer.

(12) CT-PRESENCE
      A partial answer to a question that requires a multiple-pair answer must contain an element interpreted as CT.

(13) CT-REALISATION
      The element interpreted as CT must have the possibility to realise a rise accent.

The rule in (11), which is itself rooted in the rule in (12), is motivated by the need to explain a number of apparent mismatches in question-answer congruence (e.g., (37)-(34)), and the possibility of switching the sortal key from conjunct to conjunct (e.g., (52) and (53)). The rule in (13) is motivated by the restricted distribution of CT and F in coordinated and subordinated structures, and by the observation that CTs are typically realised a rise accent. In addition, the element interpreted as CT must count as 'given' in the sense of Schwarzschild (1999). The definition

from the first chapter is repeated below.

(28)   DEFINITION OF GIVEN
        An utterance U counts as given iff it has a salient antecedent A and

        a.   if U is type e, then A and U corefer;
        b.   otherwise: modulo $\exists$-type shifting, A entails the existential F-closure
              of U.

Crucially, topicalisation in coordination under subordination can target elements
interpreted as CT, but not elements interpreted as F. Modulo the syntactic and
semantic restrictions, topicalisation is allowed in both the first and the second
conjunct. The element that is interpreted as CT moves to SpecIP, which is a
position that has been independently argued to allow the realisation of the rise
accent. However, the topicalised phrase can appear in front of the particle *že*
only in the second conjunct. The syntactic structure must reflect this fact. The
most convenient solution is to treat coordination as hierarchical adjunction. In
order to maintain that the element interpreted as CT moves to SpecIP, and at the
same time explain the possibility of it preceding and following the particle *že*, it is
assumed that *že* can appear in either I or C. This conclusion will be corroborated
by further data in the next section. The next section will also provide evidence
showing that *že* can be preceded by an element interpreted as CF. Consequently,
the particle *že* can be preceded by elements interpreted as CT or CF, but not by
elements interpreted as F.

Chapter 3

Contrastive Left Dislocation

3.1 Introduction

The consensus in the more recent literature on the Czech left periphery is that postulating only one functional projection above the highest head of the I-domain is not sufficient to account for certain data. The available analyses differ in detail, but most of them assume that particles such as *že* ('that') and *aby* ('for') are complementisers that are located in the highest head of the C-domain. Another assumption that is often made is that clitics such as *se* ('oneself') appear in either the highest head of the I-domain or the lowest head of the C-domain. Under certain unspecified conditions, clitics can optionally move to some higher functional head. In the account of the Czech left periphery proposed below, the positional (in)flexibility of complementisers and clitics is reversed. More concretely, the twofold claim advocated below is that clitics appear in I, and that complementisers might appear in either C or I. It follows that *že* and *aby* should not always be labeled 'complementisers'. While the above assumptions account for most of the attested distributional patterns, it is necessary to further assume that *že*, but not *aby*, might lexicalise a functional head immediately above CP. This is necessary to explain why contrastive left dislocation can take place below *že*, but not *aby*. As far as contrastive left dislocation is concerned, the fully articulated structure of the Czech left periphery and the distribution of the relevant elements within it is taken to be the following: [$_{CP2}$ [$_{C2}$ že/*aby] [$_{CP1}$ XP [$_{CP1}$ [$_{C1}$ Ø] [$_{IP}$ ~~XP~~ → RES [$_{IP}$ [$_I$ CL] [$_{vP}$ ... ~~XP~~ ...]]]]]]].

3.2 Matrix Clauses

In Czech, the unmarked order of constituents in a declarative sentence is SVO. For expository ease, the examples used in this chapter are mostly limited to sentences

involving only the subject and the verb. To control the interpretation, it is assumed that the declarative sentences below are answers to the following question[1].

(1)   Co   ta   DÍVKA?
       what that girl.NOM
       What about that GIRL?

The following two examples demonstrate that, regardless of its Case, the subject precedes the verb[2,3].

(2)   [Ta dívka]$_{\text{CT/T}}$ se   [USMÁLA]$_{\text{F}}$.
       that girl.NOM REFL.CL smile.PST
       That girl SMILED.

(3)   [Té dívce]$_{\text{CT/T}}$ se   [DAŘILO]$_{\text{F}}$.
       that girl.DAT   REFL.CL do-well.PST
       That girl did WELL.

Changing the order from SV to VS results in markedness[4]. The subject in the following two examples appears in the sentence-final position, which is usually where the focus exponent is located. The leftward movement of the verb should make it possible for the subject to be deaccented[5]. However, for some reason, realising the main stress on the fronted verb is not possible.

(4)   #[USMÁLA]$_{\text{F}}$ se   [ta   dívka]$_{\text{CT/T}}$.
       smile.PST   REFL.CL that girl.NOM
       Intended: That girl SMILED.

(5)   #[DAŘILO]$_{\text{F}}$ se   [té   dívce]$_{\text{CT/T}}$.
       do-well.PST REFL.CL that girl.DAT
       Intended: That girl did WELL.

In (2) and (3) above, the verb is inherently reflexive. This forces the presence of the reflexive clitic *se*. Without it, the two sentences would be ungrammatical. It has long been noted that Czech clitics are restricted to appear in the second

---

[1] It is assumed that the interaction took place in the following context. Disregarding the speaker and the addressee, there were 2 individuals (Peter, Mary), both of whom were familiar to the speaker and the addressee. Peter and Mary recently sat an exam. The speaker wanted to know either how Mary reacted to the results of the exam (this applies to cases where the subject is NOM) or how she did in the exam (this applies to cases where the subject is DAT), and the addressee supplied this information.

[2] Similarly to German and Polish, for instance, Czech allows both NOM and DAT subjects.

[3] For ease of exposition, the information-structural status of clitics is ignored. Though, it is worth pointing out that some authors consider clitics to be interpreted as G by default.

[4] If the verb appeared in the initial position, the two sentences would no longer constitute felicitous answers to the question in (1). However, with a different information-structural marking, they would be perfectly acceptable answers to *Who smiled?* and *Who did well?*, respectively.

[5] This was referred to above as rightward backgrounding.

position within their containing clause[6]. As the following examples show, placing *se* in the clause-initial position is not an option.

(6) *Se   [ta   dívka]$_{CT/T}$ [USMÁLA]$_F$.
     REFL.CL that girl.NOM   smile.PST
     Intended: That girl SMILED.

(7) *Se   [té   dívce]$_{CT/T}$ [DAŘILO]$_F$.
     REFL.CL that girl.DAT   do-well.PST
     Intended: That girl did WELL.

Changing the respective positions of the subject and the verb does not improve the acceptability of the two sentences above.

(8) *Se   [USMÁLA]$_F$ [ta   dívka]$_{CT/T}$.
     REFL.CL smile.PST   that girl.NOM
     Intended: That girl SMILED.

(9) *Se   [DAŘILO]$_F$ [té   dívce]$_{CT/T}$.
     REFL.CL do-well.PST that girl.DAT
     Intended: That girl did WELL.

Following Sturgeon (2008), let us assume that clitics are located in I, that the subject moves to SpecIP and that its lower copy is deleted. An abstract structural representation detailing the relevant parts of the sentences in (2) and (3) is assumed to be the following. Note that this structure is identical to that introduced at the end of the previous chapter. The only difference is the addition of the assumption that clitics are located in I.

(10)   An abstract representation of (contrastive) topicalisation.

_____
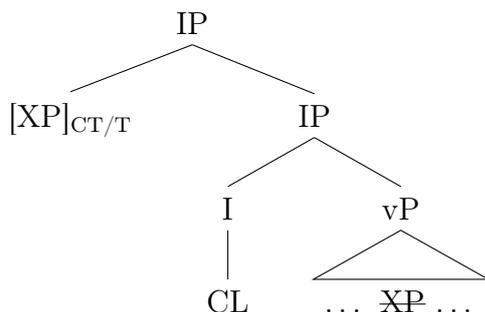
[6]Admittedly, this is a convenient oversimplification: Lenertová (2001) presents examples of clause-initial clitics and of third-position clitics; Dotlačil (2007) provides examples involving clitic climbing. What is crucial is that the relevant examples are compatible with an analysis that assumes that clitics appear in a fixed position. Indeed, Lenertová (2001) herself notes this possibility.

Since the aim of this chapter is to explore the left periphery of the Czech clause, it is paramount to investigate the distribution of the elements that can occur within it. A possible way of making the structure of the left periphery more complex, and thereby more transparent, is by left dislocating the subject. I assume that the subject in the following examples undergoes further movement to SpecCP, and that its copy in SpecIP is spelled out as a demonstrative resumptive pronoun[7].

(11) Ta dívka | [ta]$_{CT}$ se   [USMÁLA]$_F$.
      that girl.NOM   that.NOM REFL.CL smile.PST
      That girl, she SMILED.

(12) Té dívce | [té]$_{CT}$ se   [DAŘILO]$_F$.
      that girl.DAT   that.DAT REFL.CL do-well.PST
      That girl, she did WELL.

Sufficient evidence for the movement analysis of this type of left dislocation can be found in Sturgeon (2008)[8], who also argues that the resumptive pronoun is interpreted as CT. Since it is interpreted as CT, (11) and (12) must both be partial answers to a question that requires a pair-list answer[9]. One such possible question might be the following[10].

(13) Co ta dívka   a ten CHLAPEC?
      what that girl.NOM and that boy.NOM
      What about that girl and that BOY?

As far as sentences such as (11) and (12) are concerned, the subject has to precede the resumptive pronoun. In other words, spelling out the higher copy of the subject as a resumptive pronoun results in ungrammaticality.

(14) *Ta   [ta dívka]$_{CT}$ se   [USMÁLA]$_F$.
      that.NOM that girl.NOM REFL.CL smile.PST
      Intended: That girl, she SMILED.

(15) *Té   [té dívce]$_{CT}$ se   [DAŘILO]$_F$.
      that.DAT that girl.DAT REFL.CL do-well.PST
      Intended: That girl, she did WELL.

---

[7]The vertical lines in the Czech sentences mark intonation phrase boundaries. Unless relevant, this information is omitted in the ill-formed examples.

[8]The main evidence being Case connectivity, island sensitivity and reconstruction.

[9]This follows from (11).

[10]It is assumed that the interaction took place in the following context. Disregarding the speaker and the addressee, there were 2 individuals (Peter, Mary), both of whom were familiar to the speaker and the addressee. Peter and Mary recently sat an exam. The speaker wanted to know either how they reacted to the results of the exam (this applies to cases where the subject is NOM) or how they did in the exam (this applies to cases where the subject is DAT), and the addressee supplied this information.

What is perhaps more crucial is to note that the resumptive pronoun has to precede the reflexive clitic. This is, of course, predicted if the clitic is located in I and the resumptive pronoun in SpecIP.

(16) *Ta dívka   se   [ta]$_{CT}$ [USMÁLA]$_F$.
      that girl.NOM REFL.CL that.NOM smile.PST
      Intended: That girl, she SMILED.

(17) *Té dívce   se   [té]$_{CT}$ [DAŘILO]$_F$.
      that girl.DAT REFL.CL that.DAT do-well.PST
      Intended: That girl, she did WELL.

The examples below show that any vP-adjoined adverb has to follow the reflexive clitic, which is in accord with the assumption that clitics are located in I.

(18) Ta dívka | [ta]$_{CT}$ (*dnes) se   (dnes) [USMÁLA]$_F$.
      that girl.NOM   that.NOM today   REFL.CL today smile.PST
      That girl, she SMILED today.

(19) Té dívce | [té]$_{CT}$ (*dnes) se   (dnes) [DAŘILO]$_F$.
      that girl.DAT   that.DAT today   REFL.CL today do-well.PST
      That girl, she did WELL today.

Based on the discussion so far, an abstract structural representation detailing the relevant parts of the clauses in (11) and (12) is assumed to be the following.

(20)   An abstract representation of contrastive left dislocation.
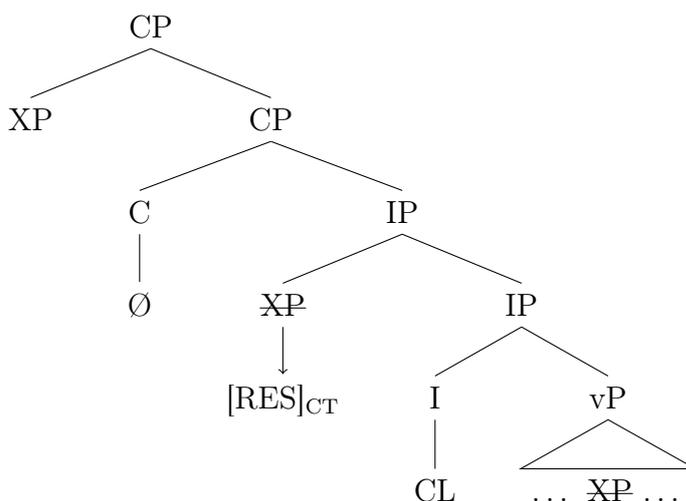
As far as the interpretation of (11) and (12) is concerned, left dislocation precludes the interpretation of the element in SpecIP as either T or F. The construction has a contrastive feel to it. While it is certainly true that the resumptive pronoun can be interpreted as CT, I am not aware of any literature that would test the

possibility of it being interpreted as CF. Consider the following sentences.

(21)   Ten chlapec   se   usmál.
       that boy.NOM REFL.CL smile.PST
       That boy smiled.

(22)   Tomu chlapci   se   dařilo.
       that   boy.DAT REFL.CL do-well.PST
       That boy did well.

One of the most widely used diagnostics for CF-hood is the correction test. The following examples are felicitous reactions to (21) and (22), respectively[11]. It is noteworthy that, in both cases, the resumptive pronoun is realised with a fall accent, which is typical of F in Czech (see Veselá et al. (2003)). Neither of the two sentences would be felicitous in the given context if the resumptive pronoun were realised with a rise accent.

(23)   Ne! Ta   dívka | [TA]$_{CF}$ se   [usmála]$_G$.
       no   that girl.NOM   that.NOM REFL.CL smile.PST
       No! That girl, SHE smiled.

(24)   Ne! Té   dívce | [TÉ]$_{CF}$ se   [dařilo]$_G$.
       no   that girl.DAT   that.DAT REFL.CL do-well.PST
       No! That girl, SHE did well.

What can be concluded from the grammaticality of (23) and (24) is that left dislocation is compatible with the resumptive pronoun being interpreted as CF. I afford to speculate that the phonological realisation of the resumptive pronoun gives the addressee a cue as to whether it is interpreted as CT (i.e., a rise accent) or CF (i.e., a fall accent). This is a mere speculation based on the perception of the left dislocate as being realised in the same way in (11), (12), (23) and (24). Concluding anything with more certainty would require acoustic analysis of the relevant examples produced in appropriate contexts. However, as the title of the present chapter suggests, the discussion will revolve around examples involving elements interpreted as CT.

---

[11]It is assumed that the interaction took place in the following context. Disregarding the speaker and the addressee, there were 3 individuals (Jacob, Peter, Mary), all of whom were familiar to the speaker and the addressee. Peter and Mary recently sat an exam. The speaker wanted to inform the addressee either about who reacted positively to the results of the exam (this applies to the case where the subject is NOM) or about who did well in the exam (this applies to the case where the subject is DAT). Knowing that the information provided by the speaker was untrue, the addressee corrected them.

3.3 Embedded Clauses

Meyer (2010) identifies three types of Mood in Czech: indicative, conditional and imperative. Indicative and conditional frequently appear in embedded clauses. These clauses are typically selected and introduced by syntactic elements that are compatible with the given Mood[12]. In his paper, Meyer implies that embedded clauses introduced by *že* ('that') are indicative, whereas those introduced by *aby* ('that') are conditional. As a necessary consequence, the corresponding embedding predicates that select such embedded clauses have to be compatible with their Mood. Indeed, Meyer notes that certain verbs favour complements of certain Mood. To take but two examples, the verb *říct* ('to say') selects either an indicative or a conditional complement, and the verb *chtít* ('to want') selects only the latter. It is informative to see whether the distinction between indicative and conditional is in any way or form manifested in the structural make-up of the clause. To this end, the grammatical sentences that have been introduced in the foregoing might be used in indicative and conditional clauses embedded under *říct* and *chtít*, respectively. Before the relevant examples are presented, a note of caution is in order. The present analysis diverges from the traditional view that *že* and *aby* are complementisers that are restricted to appear in the C-domain by assuming that these syntactic elements can also appear in the I-domain. Hence, *že* and *aby* are not referred to as complementisers.

3.3.1 Indicative Mood

The examples in this section are considered to be answers to one of the following two questions[13].

(25) A   co   ta   DÍVKA? Jak Jakub   řekl | že se   na
      and what that girl.NOM how Jacob.NOM say.PST že REFL.CL on
      výsledky   tvářila   TA?
      results.ACC react.PST she.DEM.NOM
      And what about that GIRL? How did Jacob say that SHE reacted to the
      results?

---

[12]A number of examples demonstrating that this is not always the case are presented in §3.5.

[13]It is assumed that the interaction took place in the following context. Disregarding the speaker and the addressee, there were 3 individuals (Jacob, Peter, Mary), all of whom were familiar to the speaker and the addressee. Peter and Mary recently sat an exam. The speaker wanted to know either how Jacob said that Mary reacted to the results of the exam (this applies to cases where the subject is NOM) or how he said that she did in the exam (this applies to cases where the subject is DAT), and the addressee supplied this information.

(26) A   co   ta   DÍVKA? Jak Jakub   řekl | že se   ta
      and what that girl.NOM how Jacob.NOM say.PST   že REFL.CL that
      zkouška   povedla   TÉ?
      exam.NOM   succeed.PST her.DEM.DAT
      And what about that GIRL? How did Jacob say that SHE did in the
      exam?

The following two examples show that the subject in NOM (i.e., *ta dívka*) can either precede or follow the reflexive clitic[14]. If *se* is located in I, one needs to assume that *že* lexicalises a higher functional head to explain the fact that the subject can intervene between it and the clitic in (28). In (27), *že* is realised in I. In (28), it is realised in C. The flexibility in the position of *že* ensures that the analysis of matrix clauses can be extended to embedded clauses: the embedded clause in (27) is IP, and the one in (28) is CP.

(27) Jakub   řekl | že se   [ta   dívka]$_{*CT/?T}$ [USMÁLA]$_F$.
      Jacob.NOM say.PST   že REFL.CL that girl.NOM   smile.PST
      Jacob said that that girl SMILED.

(28) Jakub   řekl | že [ta   dívka]$_{CT/?T}$ se   [USMÁLA]$_F$.
      Jacob.NOM say.PST   že that girl.NOM   REFL.CL smile.PST
      Jacob said that that girl SMILED.

What is not predicted by the analysis proposed for matrix clauses is the possibility of having left dislocation below *že*. If the subject ended up in SpecCP and the resumptive pronoun was located in SpecIP, then *že* would have to be located in a functional projection above CP. The acceptability of the following example proves that such a conclusion is necessary if the assumption that clitics are always located in I is to be maintained.

(29) Jakub   řekl | že ta   dívka | [ta]$_{CT/*T}$ se
      Jacob.NOM say.PST   že that girl.NOM   that.NOM REFL.CL
      [USMÁLA]$_F$.
      smile.PST
      Jacob said that that girl, she SMILED.

Before any analysis is proposed, let us consider the corresponding paradigm involving different Case. The following two examples show that the subject in DAT (i.e., *té dívce*) can either precede or follow the reflexive clitic. Even though the data set is not exhaustive, this is a good indication that Case does not influence

---

[14]One might wonder why the movement of the subject is optional rather than obligatory. After all, the subject of the sentence in (2) had to precede the reflexive clitic. A tentative exmplanation for this is provided in §3.4 below. Note that the subject can be interpreted as CT in (28), but not in (27).

the availability of the movement operation at hand.

(30) Jakub   řekl | že se   [té dívce]$_{*CT/?T}$ [DAŘILO]$_F$.
      Jacob.NOM say.PST   že REFL.CL that girl.DAT   do-well.PST
      Jacob said that that girl did WELL.

(31) Jakub   řekl | že [té dívce]$_{CT/?T}$ se   [DAŘILO]$_F$.
      Jacob.NOM say.PST   že that girl.DAT   REFL.CL do-well.PST
      Jacob said that that girl did WELL.

If Case does not influence the availability of movement, then it is predicted that left dislocation of the subject in DAT should be possible. This prediction is borne out.

(32) Jakub   řekl | že té dívce | [té]$_{CT/*T}$ se
      Jacob.NOM say.PST   že that girl.DAT   that.DAT REFL.CL
      [DAŘILO]$_F$.
      do-well.PST
      Jacob said that that girl, she did WELL.

Again, if the subject ends up in SpecCP and the resumptive pronoun is located in SpecIP, then *že* must be located in a functional projection above CP.

3.3.2 Conditional Mood

The examples in this section are considered to be answers to one of the following two questions[15].

(33) A   co   ta DÍVKA? Jak Jakub   chtěl | aby se   na
      and what that girl.NOM how Jacob.NOM want.PST   aby REFL.CL on
      výsledky   tvářila   TA?
      results.ACC react.PST she.DEM.NOM
      And what about that GIRL? How did Jacob want HER to react to the
      results?

(34) A   co   ta DÍVKA? Jak Jakub   chtěl | aby se
      and what that girl.NOM how Jacob.NOM want.PST   aby REFL.CL
      ta   zkouška   povedla   TÉ?
      that exam.NOM succeed.PST her.DEM.DAT
      And what about that GIRL? How did Jacob want HER to do in the exam?

---

[15]It is assumed that the interaction took place in the following context. Disregarding the speaker and the addressee, there were 3 individuals (Jacob, Peter, Mary), all of whom were familiar to the speaker and the addressee. Peter and Mary recently sat an exam. The speaker wanted to know either how Jacob wanted Mary to react to the results of the exam (this applies to cases where the subject is NOM) or how he wanted her to do in the exam (this applies to cases where the subject is DAT), and the addressee supplied this information.

Veselovská (1995) notes that *že* in embedded clauses can be followed by a maximal category that is itself followed by clitics (p.289). This was shown in (28) and (31) above. In addition, she claims that *aby* in embedded clauses does not allow this. The difference in the acceptability of the following two examples instantiates this claim. To account for the variation in the acceptability of (36), it could be assumed that native speakers differ as to whether they allow the embedded clause introduced by *aby* to be CP or not[16].

(35)   Jakub   chtěl | aby se   [ta   dívka]$_{*CT/?T}$ [USMÁLA]$_F$.
       Jacob.NOM want.PST   aby REFL.CL that girl.NOM   smile.PST
       Jacob wanted that girl to SMILE.

(36) $^{?(?)}$Jakub   chtěl | aby [ta   dívka]$_{CT/T}$ se   [USMÁLA]$_F$.
       Jacob.NOM want.PST   aby that girl.NOM   REFL.CL smile.PST
       Intended: Jacob wanted that girl to SMILE.

What is crucial is that left dislocation cannot take place below *aby*. This can be taken to indicate that the functional projection above CP that is required to account for examples such as (29) and (32) above is missing[17].

(37)   *Jakub   chtěl | aby ta   dívka | [ta]$_{CT/T}$ se
       Jacob.NOM want.PST   aby that girl.NOM   that.NOM REFL.CL
       [USMÁLA]$_F$.
       smile.PST
       Intended: Jacob wanted that girl to SMILE.

The paradigm involving subjects in DAT (see below) lines up with the paradigm involving subjects in NOM (see above).

(38)   Jakub   chtěl | aby se   [té   dívce]$_{*CT/?T}$ [DAŘILO]$_F$.
       Jacob.NOM want.PST   aby REFL.CL that girl.DAT   do-well.PST
       Jacob wanted that girl to do WELL.

(39) $^{?(?)}$Jakub   chtěl | aby [té   dívce]$_{CT/T}$ se   [DAŘILO]$_F$.
       Jacob.NOM want.PST   aby that girl.DAT   REFL.CL do-well.PST
       Intended: Jacob wanted that girl to do WELL.

---

[16]The acceptability of (36) seems to improve when *ta dívka* is marked as CT, so the difference might boil down to interpretation.

[17]It is worth noting at this point that clauses introduced by *aby* have been argued to involve CP (see Meyer (2006)). At the same time, it has been argued that infinitival clauses do not project CPs (see Dotlačil (2004)). The verb here can embed clauses introduced by *aby* as well as infinitival clauses. In the light of such observations, stipulating that various matrix verbs select complements of various sizes does not seem too *ad hoc*.

(40)   *Jakub   chtěl | aby té   dívce | [té]$_{CT/T}$ se
        Jacob.NOM want.PST   aby that girl.DAT   that.DAT REFL.CL
        [DAŘILO]$_F$.
        do-well.PST
   Intended: Jacob wanted that girl to do WELL.

3.4   Syntactic Analysis

It turns out that maintaing a single-CP analysis of the left periphery of certain
embedded clauses is untenable (see (29) and (32)). The hierarchical position of *že*
may sometimes be higher than that of *aby*. To account for the relevant data, the
following structural representations are assumed. The boxed nodes are assumed to
be selected by the embedding predicate. The dashed line at the top of these nodes
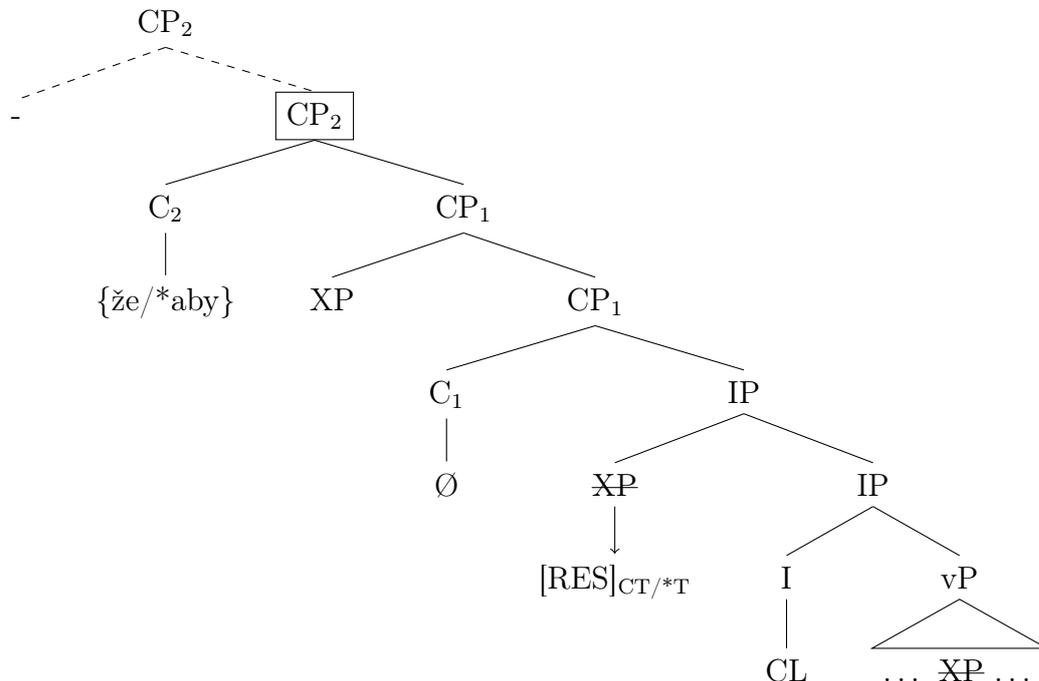indicates that adjunction is disallowed.

(41)   An abstract representation detailing the relevant parts of the embedded
        clauses in (27), (30), (35), (38).
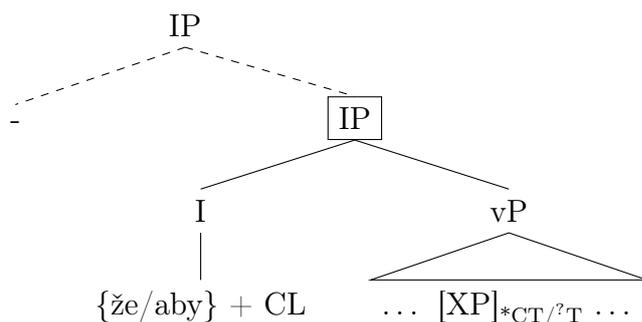
(42)   An abstract representation detailing the relevant parts of the embedded
        clauses in (28), (31), (36) and (39).
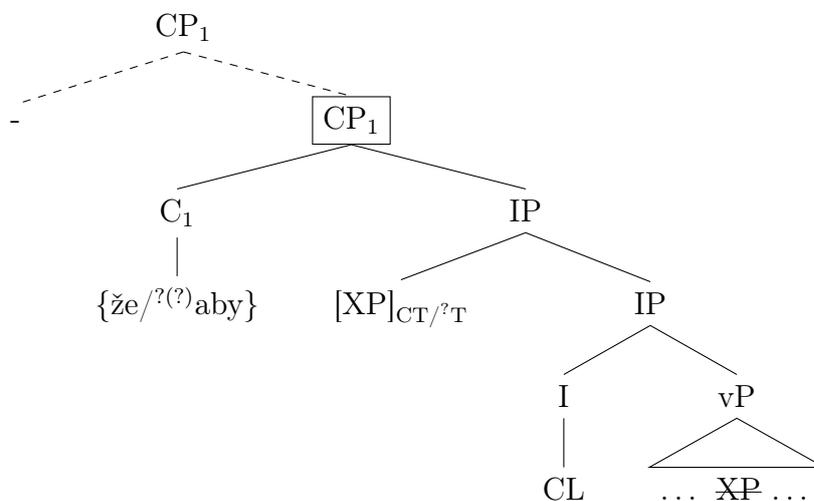
(43)   An abstract representation detailing the relevant parts of the embedded
         clauses in (29), (32), (37) and (40).

```
                        CP₂
                   ┌----´╲
                   -      CP₂
                        ╱    ╲
                      C₂      CP₁
                      │      ╱    ╲
                 {že/*aby}  XP     CP₁
                                  ╱    ╲
                                C₁      IP
                                │      ╱  ╲
                                Ø     XP    IP
                                      │    ╱  ╲
                              [RES]_{CT/*T}  I    vP
                                            │    ╱╲
                                           CL  … XP …
```

The quirky fact is that no phonologically realised syntactic material may intervene
between either *že* or *aby* and the verb that selects the phrase that they head. Even
in the simplest case (i.e., when IP is selected), the subject cannot appear in SpecIP
and precede either *že* or *aby* in I[18].

(44) $^{??/*}$Jakub   řekl | [ta dívka]$_{CT/T}$ že se   [USMÁLA]$_F$.
     Jacob.NOM say.PST   that girl.NOM že REFL.CL smile.PST
     Intended: Jacob said that that girl SMILED.

(45) $^{??/*}$Jakub   chtěl | [ta dívka]$_{CT/T}$ aby se   [USMÁLA]$_F$.
     Jacob.NOM want.PST   that girl.NOM aby REFL.CL smile.PST
     Intended: Jacob wanted that girl to SMILE.

In Minimalism, adjunction by movement subsumes adjunction by base-generation,
because the operation MOVE subsumes the operation MERGE. To disable ad-
junction to nodes selected by embedding predicates, one could try to employ the
following restriction on adjunction proposed in McCloskey (2006: p.93).

---

[18]Lenertová (2001) reports similar examples as acceptable. However, while a few native speakers
who provided acceptability judgments on (44) and (45) found the sentences just slightly degraded,
most of the informats found them either severely degraded or unacceptable. This seems to be
another manifestation of inter-speaker variation.

(46)   ADJUNCTION PROHIBITION
        Adjunction to a phrase which is s-selected by a lexical (open class) head is
        ungrammatical.

However, this condition cannot account for the impossibility of (44) and (45). The
reason for this is the following. The embedding predicate might select CP with
null C, which in turn selects IP with *že* or *aby* in I and with the subject in SpecIP.
What is needed to rule out this possibility is a condition that makes sure that *že*
or *aby* are spelled out in the topmost projection in which they can be spelled out
relative to the actual structure present. I will provide this rule further below, after
discussing data on embedded coordination.

3.5 Restrictions on CP-doubling: Mood, Modality and Polarity

A conclusion that could be drawn from the discussion so far is that CP-doubling is
licensed only in indicative complements of lexical verbs. This, combined with the
assumption that complements can be of varying sizes and the fact that there is
some inter-speaker variation, can account for the differences in the acceptability of
the data above. The picture is, however, more complicated. The difference between
the two types of Mood (i.e., indicative and conditional) cannot account for the
following data. In each case, the complement is introduced by *že*, the presence
of which signals that its containing clause is indicative[19]. When the embedding
predicate is modified by an epistemic modal, left dislocation cannot take place (see
(48) and (50)).

(47)   ?Jakub   možná řekl | že ta dívka   se   usmála.
        Jacob.NOM maybe say.PST   že that girl.NOM REFL.CL smile.PST
        Jacob maybe said that that girl smiled.

(48)   *Jakub   možná řekl | že ta dívka | ta   se
        Jacob.NOM maybe say.PST   že that girl.NOM   that.NOM REFL.CL
        usmála.
        smile.PST
        Intended: Jacob maybe said that that girl smiled.

(49)   ?Jakub   možná řekl | že té dívce   se   dařilo.
        Jacob.NOM maybe say.PST   že that girl.DAT REFL.CL do-well.PST
        Jacob maybe said that that girl did well.

---

[19]For ease of exposition, no contextual information is provided here. The examples considered
below could be used as answers to questions of the type *What did Jacob maybe say?*, *What did
Jacob not say?*, etc. For contrastive left dislocation to be licensed in the embedded clause of the
answer, the subject would have to be marked as G.

(50)   *Jakub   možná řekl | že té   dívce | té   se
       Jacob.NOM maybe say.PST   že that girl.DAT   that.DAT REFL.CL
       dařilo.
       do-well.PST
       Intended: Jacob maybe said that that girl did well.

Negating the embedding verb also prevents left dislocation from taking place (see (52) and (54)).

(51)   ?Jakub   neřekl | že ta   dívka   se   usmála.
       Jacob.NOM not-say.PST   že that girl.NOM REFL.CL smile.PST
       Jacob did not say that that girl smiled.

(52)   *Jakub   neřekl | že ta   dívka | ta   se
       Jacob.NOM not-say.PST   že that girl.NOM   that.NOM REFL.CL
       usmála.
       smile.PST
       Intended: Jacob did not say that that girl smiled.

(53)   ?Jakub   neřekl | že té   dívce   se   dařilo.
       Jacob.NOM not-say.PST   že that girl.DAT REFL.CL do-well.PST
       Jacob did not say that that girl did well.

(54)   *Jakub   neřekl | že té   dívce | té   se
       Jacob.NOM not-say.PST   že that girl.DAT   that.DAT REFL.CL
       dařilo.
       do-well.PST
       Intended: Jacob did not say that that girl did well.

Inherently negative verbs that can select complements introduced by *že* are not compatible with left dislocation either (see (56) and (58)).

(55)   ?Jakub   pochyboval | že ta   dívka   se   usmála.
       Jacob.NOM doubt.PST   že that girl.NOM REFL.CL smile.PST
       Jacob doubted that that girl smiled.

(56)   *Jakub   pochyboval | že ta   dívka | ta   se
       Jacob.NOM doubt.PST   že that girl.NOM   that.NOM REFL.CL
       usmála.
       smile.PST
       Intended: Jacob doubted that that girl smiled.

(57)   ?Jakub   pochyboval | že té   dívce   se   dařilo.
       Jacob.NOM doubt.PST   že that girl.DAT REFL.CL do-well.PST
       Jacob doubted that that girl did well.

(58) *Jakub   pochyboval | že té   dívce | té   se
    Jacob.NOM doubt.PST   že that girl.DAT   that.DAT REFL.CL
    dařilo.
    do-well.PST
    Intended: Jacob doubted that that girl did well.

If indicative clauses always licensed left dislocation, then (48), (50), (52), (54), (56) and (58) should be acceptable, which they are not. The generalisation that can be drawn from the above can be formulated in the following way[20].

(59)   CP-Doubling
        CP-doubling is licensed in embedded clauses selected by non-irrealis, non-negative bridge verbs.

There has been some disagreement in the literature as to whether the realis-irrealis distinction should be maintained along the indicative-conditional distinction. The data presented in the foregoing show that neither of the two distinctions can be reduced to the other. If 'indicative' always corresponded to 'realis', then the ungrammaticality of (52) and (54), for instance, would remain unexplained. The data above therefore show that both distinctions are necessary.

3.6 Crosslinguistic Parallelism

In Germanic languages, finite verbs in matrix clauses are often restricted to appear in the second position within their containing clause. The traditional analysis of the verb-second (V2) phenomenon assumes that some eligible phrase moves to SpecCP, and that the finite verb moves to C. The impossibility of V2 in embedded clauses introduced by overt complementisers follows: the finite verb cannot move to C, because C is already occupied by the complementiser. However, Frisian, which is a verb-final language, allows V2 even in a subset of embedded clauses introduced by overt complementisers. Interestingly, deHaan & Weerman (1985) argue that the availability of V2 is dependent on the availability of CP-doubling[21], which, in turn, follows from the generalisation in (59). When CP-doubling is licensed, the complementiser that introduces the embedded clause is located in the higher C, and the finite verb can move to the lower C.

Let us consider some Frisian data taken from deHaan & Weerman (1985: pp.84-85). The verb *leau* ('believe') can select a complement introduced by the complementiser *dat* ('that').

---

[20]A slightly different version of this generalisation was originally proposed in Iatridou & Kroch (1992), p.7.

[21]The authors use a different type of notation.

(60)   Ik leau   dat hy him wol rêde kin. (Frisian)
         I   believe that he him   save can
         I believe that he can save him.

In the embedded clause, the finite verb *kin* ('can') may undergo movement to the left periphery.

(61)   Ik leau   dat hy kin him wol rêde. (Frisian)
         I   believe that he can him   save
         I believe that he can save him.

Taken together, the following two examples show that, when the embedding verb is negated, V2 is no longer licensed in the embedded clause.

(62)   Ik leau   net dat hy him wol rêde kin. (Frisian)
         I   believe not that he him   save can
         I do not believe that he can save him.

(63)   *Ik leau   net dat hy kin him wol rêde. (Frisian)
         I   believe not that he can him   save
         I do not believe that he can save him.

The verb *sei* ('say') can also select a complement introduced by the complementiser *dat* ('that').

(64)   Pyt sei   dat hy my sjoen hie. (Frisian)
         Pyt said that he me seen   had
         Pyt said that he had seen me.

In the embedded clause, the finite verb *hie* ('had') may undergo movement to the left periphery.

(65)   Pyt sei   dat hy hie my sjoen. (Frisian)
         Pyt said that he had me seen
         Pyt said that he had seen me.

Taken together, the following two examples show that, when the embedding verb is modified by an epistemic modal, V2 is no longer licensed in the embedded clause.

(66)   Pyt woe   sizze dat er my sjoen hie. (Frisian)
         Pyt wanted to-say that he me seen   had
         Pyt wanted to say that he had seen me.

(67)   *Pyt woe   sizze dat hy hie my sjoen. (Frisian)
         Pyt wanted to-say that he had me seen
         Pyt wanted to say that he had seen me.

The availability of CP-doubling in Czech embedded clauses therefore coincides with the availability of CP-doubling in Frisian embedded clauses. It is worth noting that Iatridou & Kroch (1992) argue that Danish behaves like Frisian. The analysis of CP-doubling in Czech, manifested in the form of contrastive left dislocation, thus receives independent empirical support.

Since the generalisation in (59) mentions bridge verbs, one might wonder whether the ungrammaticality of (48), (50), (52), (54), (56) and (58) above could be due to the embedding predicate being a 'bad' bridge verb. Featherston (2004) refuses to divide embedding predicates into 'bridge verbs' and 'non-bridge verbs', and argues instead that the 'bridgeness' of verbs is a continuum. On this view, verbs vary as to how easily they allow extraction out of their complements. A plausible explanation of the unavailability of contrastive left dislocation in the relevant examples above might therefore be that CP-doubling is licensed only in complements of very 'bridgy' verbs. Since contrastive left dislocation is not allowed if the embedding predicate is modified by an epistemic modal, or negated, or inherently negative, extraction out of its complement should not be allowed either. This prediction is borne out.

(68)   Kdy$_{?1/*2}$ Jakub   ~~kdy$_T$~~ možná řekl | že se   ta
       when   Jacob.NOM when maybe say.PST   že REFL.CL that
       dívka   ~~kdy$_2$~~ usmála?
       girl.NOM   when smile.PST
       When did Jacob maybe say that that girl smiled?

(69)   Kdy$_{1/*2}$ Jakub   ~~kdy$_T$~~ neřekl | že se   ta   dívka
       when   Jacob.NOM when not-say.PST   že REFL.CL that girl.NOM
       ~~kdy$_2$~~ usmála?
       when smile.PST
       When did Jacob not say that that girl smiled?

(70)   Kdy$_{1/*2}$ Jakub   ~~kdy$_T$~~ pochyboval | že se   ta   dívka
       when   Jacob.NOM when doubt.PST   že REFL.CL that girl.NOM
       ~~kdy$_2$~~ usmála?
       when smile.PST
       When did Jacob doubt that that girl smiled?

Note that, if the embedding verb meets the requirements set out in (59), it is possible for *kdy* to undergo long-distance A'-movement from the embedded clause to the matrix clause. This is shown in the example below.

(71) Kdy$_{1/2}$ Jakub ~~kdy$_T$~~ řekl | že se ta dívka ~~kdy$_2$~~
  when Jacob.NOM when say.PST   že REFL.CL that girl.NOM when
  usmála?
  smile.PST
  When did Jacob say that that girl smiled?

The parallelism between the possibility to extract out of an embedded clause and to use contrastive left dislocation in it does not seem at all accidental. While I offer nothing beyond the observation that this parallelism exists in Czech, this observation bears a strong resemblance to similar observations made by other authors for other languages. Featherston (2004), for instance, reports the results of two experimental studies on German which show a close relation between the possibility to extract out of an embedded clause and to use V2 in it.

3.7   Embedded Coordinated Clauses

The coordination data introduced below show that the structure of the Czech left periphery would have to be complicated if one did not allow *že* and *aby* to lexicalise different functional heads. It happens to be the case that looking at matrix and embedded clauses is of little help when investigating the syntactic distribution of the two left-peripheral particles: matrix clauses do not permit *že* or *aby* when left dislocation takes place[22], and embedded clauses permit them only immediately after the verb that selects them[23]. Fortunately, the proposed analysis of the distribution of the two particles can be tested by investigating embedded clauses involving coordination. In this section, the declarative sentences involving embedded clauses with *že* are considered to be answers to (72), and those involving embedded clauses with *aby* are considered to be answers to (73)[24].

---

[22]It might be possible to topicalise an embedded clause in which contrastive left dislocation took place.

[23]There is some inter-speaker variation with respect to the possibility of the left dislocate to intervene between the embedding predicate and *že* or *aby*: a small proportion of speakers seem to find such constructions marginally acceptable. Interestingly, Lenertová (2001) reports similar examples as perfectly acceptable.

[24]It is assumed that the interaction took place in the following context. Disregarding the speaker and the addressee, there were 4 individuals (Jacob, Peter, Mary, Lucy), all of whom were familiar to the speaker and the addressee. Peter, Mary and Lucy recently sat an exam. The speaker wanted to know either how well Jacob said that Peter and Mary did in the exam (this applies to cases with *že*) or how well Jacob wanted them to do in the exam (this applies to cases with *aby*), and the addressee supplied this information.

(72)   A   co   ten chlapec   a   ta   DÍVKA? Jak Jakub   řekl |
       and what that boy.NOM and that girl.NOM how Jacob.NOM say.PST
       že se   ta zkouška   povedla   TĚM?
       že REFL.CL that exam.NOM succeed.PST that.DAT
       And what about that boy and that GIRL? How well did Jacob say that
       THEY did in the exam?

(73)   A   co   ten chlapec   a   ta   DÍVKA? Jak Jakub
       and what that boy.NOM and that girl.NOM how Jacob.NOM
       chtěl | aby se   ta zkouška   povedla   TĚM?
       want.PST   aby REFL.CL that exam.NOM succeed.PST that.DAT
       And what about that boy and that GIRL? How well did Jacob want THEM
       to do in the exam?

What was argued in the previous chapter in relation to possible answers to similar
questions was that only the initial conjunct of the embedded coordinated clause is
selected by the embedding predicate. If this observation is correct, then at least
some of the requirements usually placed on complements should be lifted in any
and every non-initial conjunct. One of these requirements prevented adjunction
to the phrase selected by the embedding predicate (see (46)). The two examples
below show that this requirement does not hold in non-initial conjuncts: *té dívce*
precedes *že* in (74) and *aby* in (75)[25].

(74)   Jakub   řekl | že [tomu chlapci]$_{CT}$ [ta   zkouška]$_G$
       Jacob.NOM say.PST   že that   boy.DAT   that exam.NOM
       [NEVYŠLA]$_F$ | ale [té   dívce]$_{CT}$ že se   [DAŘILO]$_F$.
       fail.PST   but that girl.DAT že REFL.CL do-well.PST
       Jacob said that that boy FAILED the exam, but that that girl did WELL.

(75)   Jakub   chtěl | aby [tomu chlapci]$_{CT}$ [ta   zkouška]$_G$
       Jacob.NOM want.PST   aby that   boy.DAT   that exam.NOM
       [NEVYŠLA]$_F$ | ale [té   dívce]$_{CT}$ aby se   [DAŘILO]$_F$.
       fail.PST   but that girl.DAT aby REFL.CL do-well.PST
       Jacob wanted that boy to FAIL the exam, but that girl to do WELL.

What is more striking is that left dislocation can take place in the second conjunct
of the embedded clause introduced by *aby* (see (77)). Hence, left dislocation is
licensed not only in indicative clauses (i.e., clauses introduced by *že*), but also in
conditional clauses (i.e., clauses introduced by *aby*).

---

[25]For expository ease, the embedded clauses of the relevant examples presented in this section
involve only subjects in DAT. In addition, the adversative conjunction *ale* ('but') is used instead
of *a* ('and') to force the contrastive reading.

(76)   Jakub   řekl | že [tomu chlapci]$_{CT}$ [ta   zkouška]$_G$
        Jacob.NOM say.PST   že that   boy.DAT   that exam.NOM
        [NEVYŠLA]$_F$ | ale té   dívce | [té]$_{CT}$ že se   [DAŘILO]$_F$.
        fail.PST   but that girl.DAT   that.DAT že REFL.CL do-well.PST
        Jacob said that that boy FAILED the exam, but that that girl did WELL.

(77)   Jakub   chtěl | aby [tomu chlapci]$_{CT}$ [ta   zkouška]$_G$
        Jacob.NOM want.PST   aby that   boy.DAT   that exam.NOM
        [NEVYŠLA]$_F$ | ale té   dívce | [té]$_{CT}$ aby se
        fail.PST   but that girl.DAT   that.DAT aby REFL.CL
        [DAŘILO]$_F$.
        do-well.PST
        Jacob wanted that boy to FAIL the exam, but that girl to do WELL.

If left dislocation in the second conjunct of the embedded clause proceeds in the
same way as left dislocation in matrix clauses and uncoordinated embedded clauses,
then the analysis which assumes that *že* and *aby* are restricted to appear in the
C-domain fails to account for the acceptability of (76) and (77). Apparently, *že* and
aby must sometimes be allowed to appear in I. The following examples show that
the two particles can also intervene between the left dislocate and the resumptive
pronoun. In structural terms, they must be allowed to appear in C.

(78)   Jakub   řekl | že [tomu chlapci]$_{CT}$ [ta   zkouška]$_G$
        Jacob.NOM say.PST   že that   boy.DAT   that exam.NOM
        [NEVYŠLA]$_F$ | ale té   dívce | že [té]$_{CT}$ se   [DAŘILO]$_F$.
        fail.PST   but that girl.DAT   že that.DAT REFL.CL do-well.PST
        Jacob said that that boy FAILED the exam, but that that girl did WELL.

(79)   Jakub   chtěl | aby [tomu chlapci]$_{CT}$ [ta   zkouška]$_G$
        Jacob.NOM want.PST   aby that   boy.DAT   that exam.NOM
        [NEVYŠLA]$_F$ | ale té   dívce | aby [té]$_{CT}$ se
        fail.PST   but that girl.DAT   aby that.DAT REFL.CL
        [DAŘILO]$_F$.
        do-well.PST
   Jacob wanted that boy to FAIL the exam, but that girl to do WELL.

Interestingly, the impossibility of left dislocation below *aby* in uncoordinated
embedded clauses carries over to coordinated embedded clauses.

(80)   Jakub   řekl | že [tomu chlapci]$_{CT}$ [ta   zkouška]$_G$
        Jacob.NOM say.PST   že that   boy.DAT   that exam.NOM
        [NEVYŠLA]$_F$ | ale že té   dívce | [té]$_{CT}$ se   [DAŘILO]$_F$.
        fail.PST   but že that girl.DAT   that.DAT REFL.CL do-well.PST
        Jacob said that that boy FAILED the exam, but that that girl did WELL.

(81)   *Jakub   chtěl | aby [tomu chlapci]$_{CT}$ [ta   zkouška]$_{G}$
       Jacob.NOM want.PST   aby that   boy.DAT   that exam.NOM
       [NEVYŠLA]$_{F}$ | ale aby té   dívce | [té]$_{CT}$ se
       fail.PST   but aby that girl.DAT   that.DAT REFL.CL
       [DAŘILO]$_{F}$.
       do-well.PST
       Intended: Jacob wanted that boy to FAIL the exam, but that girl to do
       WELL.

The following is the proposed structural representation of contrastive left dislocation
in the second conjunct of an embedded coordinated clause.

(82)   An abstract structural representation detailing the second conjunct of an
          embedded coordinate clause.
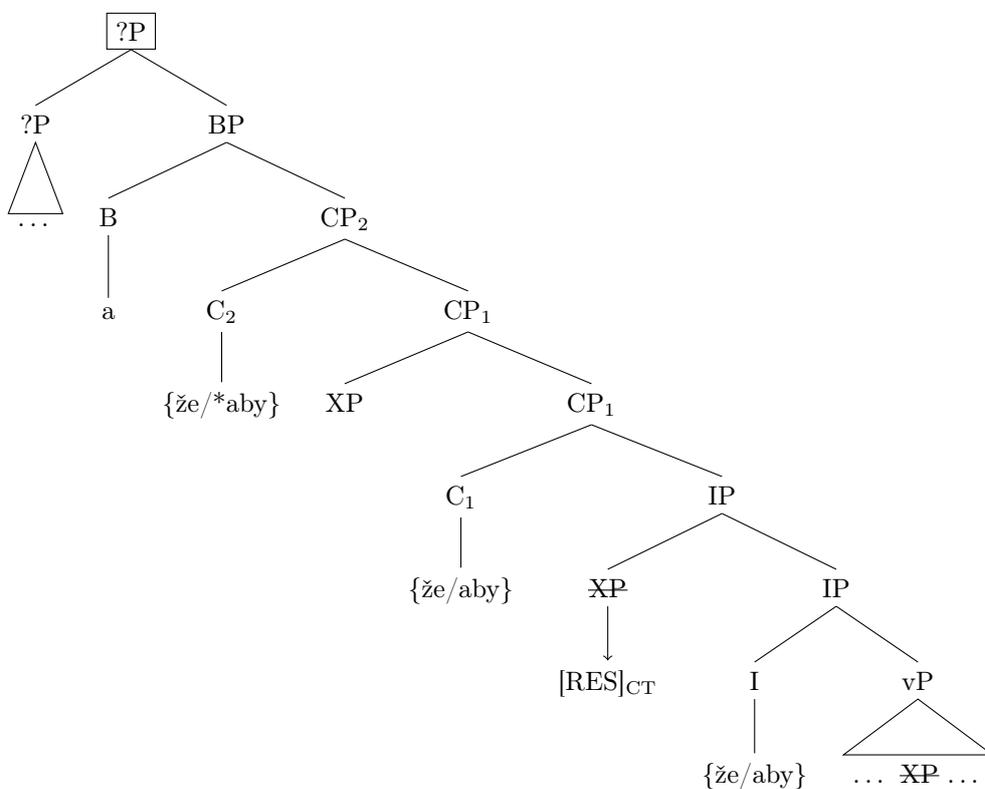
One might wonder whether the unacceptability of (81) is due to the unavailability
of CP-doubling or the inability of *aby* to appear in the higher C. Since the second
conjunct is not selected by the embedding predicate, CP-doubling should be, in
principle, possible, just like it is possible in matrix clauses, which are themselves
unselected. Recall that the claim that the second conjunct is not selected was
based on the observation that the restriction on adjunction (see (46)) did not seem
to apply to it. What is problematic is the fact that in order for the second conjunct

not to be selected, it must be treated as an adjunct to the first conjunct[26]. Since (46) is a general ban on adjunction to phrases selected by lexical heads, it also rules out adjunction to the highest phrase of the first conjunct. The problem thus becomes circular. A possible solution might be to assume that the second conjunct is adjoined not to the selected phrase, but rather to some lower phrase[27]. This solution is at odds with the assumption that the selected phrase might be as small as IP. If only IP were selected (see (27), (30), (35) and (38)), adjunction would have to be to vP or VP. Since this seems highly unlikely, it might be better to change the structural restriction in (46) to the linear one below[28].

(83) ADJACENCY RESTRICTION
      A head of a phrase s-selected by a lexical head has to be phonologically
      realised and linearly adjacent to the phrase headed by that lexical head.

The restriction above, while uniquely designed to account for the Czech data, finds independent empirical support. Philip (2013) argues for the existence of a class of functional heads called 'linkers'. The author defines this class as follows: 'A linker is a syntactically independent, semantically vacuous word serving only to mark the presence of an independently existing relationship - modification or θ-role assignment - between a head in one extended projection and a distinct dependent projection, the dependent being sister to (a projection of) the head' (pp.167-168). The relevant part of Philip's (2013) proposal concerns the distribution of linkers, which the author assumes to be restricted by the following set of violable word order constraints. The ranking between the constraints is assumed to be the following: HEAD-PROXIMATE FILTER » FINAL-CLAUSE CONSTRAINT, HEAD UNIFORMITY CONSTRAINT.

(84) HEAD-PROXIMATE FILTER
      The highest head in a complete extended projection must be contiguous
      with the lexical head of its superordinate extended projection.

(85) FINAL-CLAUSE CONSTRAINT
      A clausal dependent must follow the lexical head of its superordinate
      extended projection.

(86) HEAD UNIFORMITY CONSTRAINT
      A functional head must match the lexical head of its extended projection
      in the direction of headedness.

---

[26]This is argued for in Munn (1993), whose analysis of coordination is adapted here.
[27]This is in accord with the analysis proposed in McCloskey (2006), whose restriction on adjunction in (46) does not block adjunction to IP if it is selected by C (i.e., a functional head).
[28]This possibility was first pointed out to me by Manfred Krifka (p.c.).

As far as complementisers are concerned, the analysis above correctly predicts that the only order found cross-linguistically in VO languages is V[CVO]. This follows from the assumption that linkers, being semantically vacuous, do not encode any ordering restrictions. Therefore, they cannot influence the choice of an optimal candidate selected by the combination of the three constraints above. Lexical heads and functional heads that are not semantically vacuous might introduce additional constraints that are ranked differently with respect to the three constraints above, depending on their nature. Importantly, Philip (2013) argues that linkers, such as English *that*, must be linearly contiguous with their superordinate head. The restrictions in (84), (85) and (86) could be applied to the Czech data. However, similarly to McCloskey's (2006) restriction in (46), this would not prevent *že* or *aby* from appearing in a lower functional head. In addition, the strict interpretation of Philip's (2013) analysis would wrongly rule out examples where the embedding predicate is followed by another constituent within the matrix clause itself.

(87) Jakub   řekl   Petrovi | že té   dívce | té   se
      Jacob.NOM say.PST Peter.DAT   že that girl.DAT   that REFL.CL
      dařilo.
      do-well.PST
      Jacob told Peter that that girl did well.

That the embedded clause is indeed selected by *řekl* is clear from the fact that omitting it renders the otherwise acceptable sentence unacceptable. This is exemplified below.

(88) *Jakub   řekl   Petrovi.
      Jacob.NOM say.PST Peter.DAT
      Jacob told Peter.

While the analysis proposed by Philip (2013) might not be directly applied to the Czech data presented above, I would like to borrow one of its insights, namely that linkers are semantically empty. I would like to argue that the particle *že* is a linker when functioning as a subordinating conjunction. This also explains why it never appears in matrix clauses: matrix clauses are not linked to any superordinate heads.

3.8 Embedded Clauses and Clitic Climbing

There is one additional caveat which concerns the functionality of the proposed analysis. Recall that *že* and *aby* were claimed to be able to lexicalise C or I, depending on, among other things, whether the complement selected by the

embedding predicate was CP or IP. However, the consensus in the literature is that embedded clauses introduced and headed by either *že* or *aby* are CPs. Dotlačil (2007) notes that, while *wh*-movement out of such clauses is possible (see (89) and (90)), clitic climbing is blocked (see (91) and (92))[29]. The unacceptability of (91) and (92) is supposed to stem from the impossibility of clitics to move past CP.

(89) Komu   si   myslel | že to   Petr   řekl
      who.DAT REFL.CL think.PST   že it.ACC.CL Peter.NOM say.PST
      ~~komu~~?
      who.DAT
      Whom did you think that Peter said it to?

(90) Komu   si   chtěl | aby to   Petr   řekl
      who.DAT REFL.CL want.PST   aby it.ACC.CL Peter.NOM say.PST
      ~~komu~~?
      who.DAT
      Whom did you want Peter to say it to?

(91) *Ty   si   mu   myslel | že ~~mu~~
      you.NOM REFL.CL him.DAT.CL think.PST   že him.DAT.CL
      to   Petr   řekl.
      it.ACC.CL Peter.NOM say.PST
      Intended: You thought that Peter said it to him.

(92) *Ty   si   mu   chtěl | aby ~~mu~~
      you.NOM REFL.CL him.DAT.CL want.PST   aby him.DAT.CL
      to   Petr   řekl.
      it.ACC.CL Peter.NOM say.PST
      Intended: You wanted Peter to say it to him.

The counterparts of (91) and (92) with the clitic *mu* left in the embedded clause are grammatical.

(93) Ty   si   myslel | že mu   to
      you.NOM REFL.CL think.PST   že him.DAT.CL it.ACC.CL
      Petr   řekl.
      Peter.NOM say.PST
      You thought that Peter said it to him.

(94) Ty   si   chtěl | aby mu   to
      you.NOM REFL.CL want.PST   aby him.DAT.CL it.ACC.CL
      Petr   řekl.
      Peter.NOM say.PST
      You wanted Peter to say it to him.

---

[29]Clitic climbing is meant to be understood as movement of clitics from an embedded clause to a matrix clause.

Dotlačil (2007) claims that infinitival clauses are IPs, and he attributes the possibility of clitic climbing out of infinitival clauses to the lack of the CP projection. The following example shows this[30].

(95)   Petr   si   mu   to   přál   ~~mu~~
        Peter.NOM REFL.CL him.DAT.CL it.ACC.CL wish.PST him.DAT.CL
        ~~to~~   říct.
        it.ACC.CL say.INF
        Peter wished to say it to him.

My analysis of the Czech left periphery, which assumes a degree of flexibility in the positioning of *že* and *aby*, could be saved by allowing various functional heads to encode finiteness[31]. The C of CPs or the I of IPs headed by *že* or *aby* would be specified as [+finite], and the I of infinitival IPs would be specified as [−finite]. Clitic climbing would then be blocked by a functional head specified as [+finite]. A similar restriction on clitic climbing is proposed in Lenertová (2001).

Alternatively, one might assume, following Veselovská (1995), that infinitival clauses are smaller than IPs, and that clitics may not move out of IPs. On this account, the impossibility of clitic climbing can be explained as follows. Since the presence of either *že* or *aby* signals the presence of IP, clitic climbing out of embedded clauses introduced and headed by either *že* or *aby* is disallowed. Clitic climbing out of infinitival clauses is allowed, because they are smaller than IPs. Crucially, which one of the two possibilities outlined above happens to be empirically or theoretically more adequate is orthogonal to the present discussion. What is important is that the proposed analysis is compatible with at least some of the existing analyses of clitic climbing in Czech (i.e., Lenertová (2001) and Veselovská (1995)).

3.9   Embedded Root Phenomena and Speech Acts

At first sight, it might seem rather remarkable that CP-doubling in Czech (which is a VO language) should pattern with CP-doubling in Frisian (which is an OV language). However, on a closer look, it could be observed that both contrastive left dislocation in Czech and embedded verb second in Frisian are instances of embedded root phenomena. In Czech, left dislocation is almost exclusively found in root

---

[30]Note that the ungrammaticality of (91) and (92) cannot be due to the presence of the reflexive clitic *si* in the matrix clause. The following example, in which the embedding predicate selects an infinitival complement, involves clitic climbing. In other words, clitics that are base-generated in the embedded clause can be adjoined to clitics that are base-generated in the matrix clause.

[31]This line of reasoning conforms to the analysis of English and Serbo-Croatian proposed in Wurmbrand (2015). Among other things, the author assumes that functional heads in both the C-domain and the I-domain might be either [+finite] or [−finite].

clauses. In Frisian, verb second is typically operative in root clauses. Consequently, it can be proposed that what was labelled above as CP-doubling is in fact an addition of a functional projection encoding illocutionary force. On this view, the highest functional projection in (43) is taken to encode assertoric force. The licensing conditions on 'CP-doubling' in Czech follow from the (in)compatibility of the embedding predicate with the embedded speech act: ASSERT>ASSERT (see (29) and (32)); *MAYBE-ASSERT>ASSERT (see (48) and (50)); *NOT-ASSERT>ASSERT (see (52) and (54)); *DOUBT>ASSERT (see (56) and (58)). As noted in de Haan (2001), the distribution of embedded verb second in Frisian might be explained in a similar vein.

Krifka (2001) argues that, 'while coordination is a well-formed operation for speech acts, disjunction is not' (p.16). The author claims that 'syntactic forms that look like disjunction of two speech acts ... are interpreted in special ways, for example, by lowering the disjunction to the propositional level' (*ibid.*). The example he gives to show this is the following.

(96)   Al made the pasta, or Bill made the salad.

    a.   'I assert: Al made the pasta, or Bill made the salad.'
    b. #'I assert: Al made the pasta, or I assert: Bill made the salad.'

In the sentence in (96), disjunction is interpreted as disjunction of the asserted propositions, and not as disjunction of the acts of assertion. This can, of course, be derived syntactically by assuming that the constituents corresponding to the two propositions (i.e., *Al made the pasta* and *Bill made the salad*) are disjoined below the projection that encodes the speech act of assertion. In addition to the above, Krifka (2001) argues that it is possible to embed speech acts. The twofold claim that speech acts may be embedded and that they may not be disjoined can be used as a diagnostic for the presence of a speech act in contrastive left dislocation under embedding. The following example demonstrates that two embedded clauses in which contrastive left dislocation took place can be conjoined, but not disjoined[32].

(97)   Petr   řekl | že tomu chlapci | tomu   ta   zkouška
        Peter.NOM said   že that boy.DAT   that.DAT that exam.NOM
        nevyšla | {a/*nebo} té   dívce | té   že nevyšel ten
        fail.PST   and/or   that girl.DAT   that.DAT že fail.PST that
        pohovor.
        interview.NOM
        Peter said that that boy did not pass the exam, and that that girl did not
        succeed in the interview.

---

[32]The curly brackets signal a set of mutually exclusive alternatives.

Admittedly, for the above argument to be valid, it needs to be shown that disjunction is a licit syntactic operation. Unlike (97) above, (98) below does not involve contrastive left dislocation.

(98)   Petr   řekl | že tomu chlapci   ta   zkouška   nevyšla |
           Peter.NOM said   že that boy.DAT that exam.NOM fail.PST
           {a/nebo} že   té   dívce   nevyšel ten   pohovor.
           and/or   that that girl.DAT fail.PST interview.NOM
           Peter said that that boy did not do-well in the exam and/or that that girl did well in the interview.

Given the above, I conclude that contrastive left dislocation needs a speech act.

3.10   Alternative Syntactic Analysis

One of the claims made above was that (modulo the restrictions placed on the embedded clause by the matrix clause) matrix and embedded clauses come in different sizes. Another claim was that *že* and *aby* should not be labelled complementisers, because they can appear not only in C, but also in I. There is, however, an alternative analysis which maintains that both particles are restricted to appear in C. On this analysis, *že*, but not *aby*, has the option of lexicalising a higher C in clauses in which CP-doubling is permitted. One potential problem with such an analysis relates to the position of clitics. If *aby* were always located in C, and if clitics were always located in I, then the subject in SpecIP could intervene between the two elements. In order to account for the inter-speaker variation noted above (see the discussion of (36)), it would have to be assumed that native speakers differ as to whether they allow movement of clitics from I to C. Assuming this, however, would fail to account for the contrast between, for instance, (28) and (36). If *že* and *aby* were always in C, and if the movement of clitics from I to C was either allowed or disallowed for a given speaker, then there should be no contrast between the acceptability of (28) and (36). Stipulating that clitics were always located in C, one could propose to explain the aforementioned contrast by allowing *že* to move to a higher C. This would amount to saying that *že*, but not *aby*, could lexicalise different functional heads. As a consequence, the alternative analysis would begin to bear a strong resemblance to the analysis proposed here. However, while allowing *že* to lexicalise or move to a higher C would solve the issue related to the contrast between (28) and (36), it would require an unnecessary complication of syntax. In order to explain the contrast between the two examples, at least two CPs would be necessary. This would create another problem. In clauses that do not license CP-doubling, there is only a single C position. The alternative analysis,

which would allow *že*, and possibly *aby*, to appear in a higher C, would therefore fail to account for the contrast between (51) and (36), for instance. Note that both examples were argued to disallow CP-doubling. Of course, the analysis could be 'saved' by further complicating syntax. One could assume that two CPs are the bare minimum, and that doubling of the higher CP is allowed in embedded clauses that meet the requirements in (59). In making the alternative analysis work, one would therefore end up adapting the analysis proposed in the foregoing, only a syntactic level higher. Since the alternative analysis is theoretically less elegant than the one proposed here, it is not considered further.

3.11   Conclusion

As far as the distribution of subordinating conjunctions is concerned, the structural analysis proposed above improves upon the existing ones by having a broader empirical coverage and by requiring fewer functional projections. The subordinating conjunctions *že* and *aby* were argued to have the possibility to lexicalise different functional heads (i.e., C or I), depending on, among other things, whether their containing clause was selected by the embedding predicate or not. In the light of the data presented above, I proposed that Czech be added to the existing list of languages exhibiting CP-doubling. In addition, I demonstrated that the licensing conditions on CP-doubling in Czech are similar, if not identical, to the licensing conditions on CP-doubling in West Frisian. The (im)possibility of CP-doubling in certain contexts combined with the varying sizes of complements and varying positions of the subordinating conjunctions *že* and *aby* allowed me to maintain a relatively small amount of functional projections while deriving sentences involving contrastive topicalisation and contrastive left dislocation. Similarly to English *that*, Czech *že* and *aby* can be doubled. What remains to be investigated, then, is whether such doubling has any interpretive effect. Since *že* and *aby* were argued to be semantically vacuous linkers, their doubling is predicted to have no direct interpretive effect. However, the addition of an additional phonologically non-null element might have an indirect interpretive effect.

Chapter 4

Tag Questions

4.1 Introduction

Tag questions are generally understood as comprising two parts: an anchor and a tag. The anchor is that part of the question to which the tag is attached. In the least marked use, the anchor has declarative force and the tag has interrogative force[1]. In English, the polarity of the tag is typically the opposite of the polarity of the anchor. If the polarity of the anchor is positive, the polarity of the tag is negative, and *vice versa*. The two questions below exemplify this.

(1)   Peter lied to you, didn't he?

(2)   Peter didn't lie to you, did he?

Czech has a number of elements that can be used as question tags, and the particle *že* is one of them. The vast majority of tag questions involving *že* are formed by inserting the particle at the end of a declarative clause or sentence. This use is almost identical to the use of question tags in English. However, there are certain differences. For instance, *že* can appear on its own (i.e., without any auxiliary verbs, polarity markers etc.). As a consequence, the particle has been claimed to have neither positive nor negative polarity (see Urešová (2008)). The following two examples illustrate the neutral polarity of *že*. In (3), it is combined with a positive anchor. In (4), it is combined with a negative anchor[2].

---

[1]Of course, this is not the only possibility. For instance, it is not hard to encounter tag questions that function as commands. Consider the following two examples. In each case, the the anchor has the form of a command rather than a declarative.

(i)   Clean your room, will you?

(ii)   Go to the shop, could you?

[2]Whether an anchor is positive or negative is determined by the polarity of the lexical verb in the root clause. If the anchor involved embedding, the polarity of the anchor would be determined

(3) Petr   ti   lhal | že?
     Peter.NOM you.DAT.CL lie.PST   že
     Peter lied to you, right?

(4) Petr   ti   nelhal | že?
     Peter.NOM you.DAT.CL not-lie.PST   že
     Peter did not lie to you, right?

As has been hinted, there exists a possibility of using *že* in conjunction with other syntactic categories. One such category is polarity markers, the use of which seems to 'override' the neutral polarity of *že*. Consider the following pair of examples.

(5) Petr   ti   lhal | že ano?
     Peter.NOM you.DAT.CL lie.PST   že yes
     Peter lied to you, right?

(6) Petr   ti   nelhal | že ne?
     Peter.NOM you.DAT.CL not-lie.PST   že no
     Peter did not lie to you, right?

Apparently, when the particle is used together with a polarity marker, the polarity of the tag has to be identical to the polarity of the anchor. However, the picture is more complicated. While a positive anchor cannot be combined with a negative tag (see (7)), a negative anchor can be combined with a positive tag (see (8)).

(7) *Petr   ti   lhal | že ne?
      Peter.NOM you.DAT.CL lie.PST   že no
      Intended: Peter lied to you, right?

(8) Petr   ti   nelhal | že ano?
     Peter.NOM you.DAT.CL not-lie.PST   že yes
     Peter did not lie to you, right?

The interim conclusion is the following: a positive anchor can combine only with a positive tag (see (5)); a negative anchor can combine with either a negative tag (see (6)) or a positive tag (see (8)). In certain cases, it is also possible for the tag to comprise syntactic material that is identical to the syntactic material in the anchor[3].

(9) Petr   ti   lhal | že (ti)   lhal   (*ti)?
     Peter.NOM you.DAT.CL lie.PST   že you.DAT.CL lie-PST you.DAT.CL
     Peter lied to you, right?

---

by the polarity of the matrix verb, not the embedded one.

[3]In this particular case, the clitic *ti* in the tag may be dropped. However, the questions sound better with it. In addition, as (9) and (10) show, *ti* has to precede the verb in the tag.

(10)   Petr   ti   nelhal | že (ti)   nelhal
       Peter.NOM you.DAT.CL not-lie.PST   že you.DAT.CL not-lie-PST
       (*ti)?
       you.DAT.CL
       Peter did not lie to you, right?

In (9) and (10), the polarity of the verb in the tag has to match the polarity of the verb in the anchor. As the examples below show, any mismatch in polarity of the two verbs results in unacceptability.

(11)   *Petr   ti   lhal | že (ti)   nelhal?
        Peter.NOM you.DAT.CL lie.PST   že you.DAT.CL not-lie-PST
        Intended: Peter lied to you, right?

(12)   *Petr   ti   nelhal | že (ti)   lhal?
        Peter.NOM you.DAT.CL not-lie.PST že you.DAT.CL lie-PST
        Intended: Peter did not lie to you, right?

Interestingly, even when the restriction on matching polarity is satisfied, it is impossible for the tag to comprise a polarity marker and a verb at the same time.

(13)   *Petr   ti   lhal | že ano (ti)   lhal?
        Peter.NOM you.DAT.CL lie.PST   že yes you.DAT.CL lie-PST
        Intended: Peter lied to you, right?

(14)   *Petr   ti   nelhal | že ne (ti)   nelhal?
        Peter.NOM you.DAT.CL not-lie.PST že ne you.DAT.CL not-lie-PST
        Intended: Peter did not lie to you, right?

Since the polarity marker and the verb are in complementary distribution, I conclude that they represent the same category. Recall that, in the discussion of *že* in declaratives, I proposed that the particle might lexicalise two different functional heads (i.e., C or I). In both cases, *že* selects a clausal complement (i.e., CP/IP or vP). In (9) and (10), this is directly deducible from the presence of *ti (ne)lhal* in the tag. However, the clausal status of *ano* in (5) and *ne* in (6) can be deduced only indirectly, as the two polarity markers are in complementary distribution with *ti lhal* and *ti nelhal*, respectively. While this analysis explains the use of *ano* in (5) and *ne* in (6), it fails to explain the use of *ano* in (8). To account for this use, I first need to introduce some of the properties of minimal answers to *yes-no* questions in Czech.

4.2 Minimal Answers to *Yes-No* Questions

Gruet-Skrabalova (2015) considers minimal answers to *yes-no* questions of various types. In Czech, *yes-no* questions typically have the following form.

(15) Lhal   ti   Petr?
      lie.PST you.DAT.CL Peter.NOM
      Did Peter lie to you?

(16) Nelhal   ti   Petr?
      not-lie.PST you.DAT.CL Peter.NOM
      Did Peter not lie to you?

Notice that the verb in the examples above does not appear in its canonical position, but rather clause-initially. Importantly, *yes-no* questions can also take the form of a rising declarative. As the name suggests, when they do, they are distinguished from their declarative counterparts prosodically. In these *yes-no* questions, the verb remains in its canonical position.

(17) Petr   ti   lhal?
      Peter.NOM you.DAT.CL lie.PST
      Peter lied to you?

The question above can be answered by either *ano* or *ne* (see below). The polarity markers can optionally be followed by a verb that is identical to the verb in the question they answer. In such cases, the polarity of the polarity marker must match the polarity of the verb. The fact that the polarity marker and the verb can co-occur suggests that the two elements do not represent the same category. The question in (17) might carry a positive presupposition, which can be can be interpreted as a positive assertion (i.e., [P]). The presupposition can be confirmed by *ano* (i.e., *ano* = [P]) and denied by *ne* (i.e., *ne* = [not P]). Alternatively, polarity markers in minimal answers to positive rising declaratives might be thought of not as presuppositional confirmations and denials, but rather as answers to polar questions (i.e., [P or not P]). In such cases, *ano* corresponds to a positive proposition (i.e., *ano* = [P]) and *ne* to a negative one (i.e., *ne* = [not P]). Accidentally, the difference between the two possible interpretations of the question is not reflected in the distribution of polarity markers in the answer.

(18) Ano (lhal).   [confirmation]
      yes lie.PST
      Yes (, he lied to me).

(19)   Ne (nelhal).   [denial]
        no not-lie.PST
        No (, he did not lie to me).

(20)   Ano (*nelhal).
        yes not-lie.PST
        Yes (, he did not lie to me).

(21)   Ne (*lhal).
        no lie-.PST
        No (, he lied to me).

When the rising declarative comprises negation, the distinction between the two possible interpretations becomes visible. Consider the question below.

(22)   Petr ti   nelhal?
        Peter you.DAT.CL not-lie.PST
        Peter did not lie to you?

The proposition expressed by the question above can be confirmed by either *ano* or *ne* and denied by either *ale ano* or *ne*. Apparently, the polarity of the polarity marker does not have to agree with the polarity of the verb that might follow it (see (25) and (26)).

(23)   Ale ano (lhal).   [denial]
        but yes lie.PST
        Yes (, he lied to me).

(24)   Ne (nelhal).   [confirmation]
        no not-lie.PST
        No (, he did not lie to me).

(25)   Ano (nelhal).   [confirmation]
        yes not-lie.PST
        Yes (, he did not lie to me).

(26)   Ne (lhal). [denial]
        no lie.PST
        No (, he lied to me).

Gruet-Skrabalova (2015) argues that 'the use of particles depends on whether sentential negation is interpreted as true or as expletive' (p. 205). Questions that involve true negation carry a negative presupposition, which can be interpreted as a negative assertion (i.e., [not P]). The presupposition can be confirmed by *ano* (i.e., *ano* = [not P]) and denied by *ne* (i.e., *ne* = [not not P]). Questions that involve expletive negation do not carry a negative presupposition, and are interpreted as polar questions (i.e., [P or not P]). In such cases, *ano* corresponds to a positive

proposition (i.e., *ano* = [P]) and *ne* to a negative one (i.e., *ne* = [not P]).

According to Gruet-Skrabalova (2015), polarity markers express either absolute polarity or relative polarity. Absolute polarity can be either positive (i.e., [+]) or negative (i.e., [−]). In the examples above, *ano* is [+] and *ne* is [−]. Similarly, *lhal* is [+] and *nelhal* is [−]. Relative polarity, on the other hand, has the following two values: same ([Q+,A+] or [Q−,A−]) and reverse ([Q+,A−] or [Q−,A+]). It specifies the relation between the absolute polarity of the question and the absolute polarity of the answer. Here, I consider only the absolute polarity.

I distinguish between the polarity and interpretation of questions. Polarity expresses the absolute polarity of a question (i.e., [+] = [P] or [−] = [not P]), and interpretation expresses its meaning (i.e., [P], [not P] or [P or not P]). For reasons that will become apparent in the next section, the focus here is only on answers that are confirmations. A 'confirmation' is taken to mean 'the interpretation of the answer is identical to the polarity of the question'. Since the interpretation of the answer is dependent on the interpretation of the question, it is important to consider all possible scenarios.

Positive questions and negative questions with expletive negation ask [P or not P]. In the answer, *ano* = [P] and *ne* = [not P]. Positive rising declaratives ask for a confirmation of [P]. In the answer, *ano* = [P]. Negative questions with true negation ask for a confirmation of [not P]. In the answer, *ano* = [not P]. The following are some consequences of the analysis: a positive question or a positive rising declarative can be confirmed only by *ano*; a negative question with expletive negation can be confirmed only by *ne*; a negative question with true negation can be confirmed only by *ano*. The discussion so far is summarised below.

(27)   Positive rising declaratives

     a.   Question: has positive polarity and asks for a confirmation of [P].

     b.   Answer: *ano* = [P], *ne* = [not P].

     c.   Question-Answer Correspondence: *ano* = true, *ne* = false.

     d.   Correlate: these can correspond to tags with [+] verbs.

(28)   Positive questions

     a.   Question: has positive polarity and asks [P or not P].

     b.   Answer: *ano* = [P], *ne* = [not P].

     c.   Question-Answer Correspondence: *ano* = true, *ne* = false.

     d.   Correlate: these can correspond to tags with [+] verbs.

(29)   Negative rising declaratives

    a.   Question: has negative polarity and asks for a confirmation of [not P].

    b.   Answer: *ano* = [not P], *ne* = [not not P].

    c.   Question-Answer Correspondence: *ano* = true, *ne* = false.

    d.   Correlate: these can correspond to tags with [−] verbs.

(30)   Negative questions (with expletive negation)

    a.   Question: has negative polarity and asks [P or not P].

    b.   Answer: *ano* = [P], *ne* = [not P].

    c.   Question-Answer Correspondence: *ano* = false, *ne* = true.

    d.   Correlate: these do not correspond to any tags with verbs.

4.3   Analysis of Tag Questions

Semantically, tag questions with [+] anchor are either positive rising declaratives (see (27)) or positive questions (see (28)), and tag questions with [−] anchor are either negative rising declaratives (see (29)) or negative questions with expletive negation (see (30)).

The apparently odd distribution of polarity markers in tag questions involving *že* can be straightforwardly explained by assuming that the tag itself functions as a kind of minimal answer. Above, I mentioned that Gruet-Skrabalova (2015) argues that a positive *yes-no* question (see (17)) can be confirmed only by a minimal answer comprising *ano* (see (18)), and that a negative *yes-no* question (see (22)) can be confirmed by a minimal answer comprising either *ne* (see (24)) or *ano* (see (25)). As far as tag questions involving *že* are concerned, I have demonstrated above that a positive anchor is compatible only with a tag that comprises *ano* (see (5)), and that a negative anchor is compatible with a tag that comprises either *ne* (see (6)) or *ano* (see (8)). Disregarding bias, the licensing conditions on the use of polarity markers in minimal answers to *yes-no* questions that take the form of rising declaratives are therefore identical to the licensing conditions on the use of polarity markers in tag questions with *že*.

The analysis of tag questions proposed above predicts a certain degree of ambiguity. More specifically, tag questions with *že* that are not coupled with any polarity markers might be two-way ambiguous, regardless of whether they combine with a [+] anchor or a [−] anchor. This is, however, a desirable result, because tag questions are ambiguous between the interpretation of polarity questions (i.e., [P or not P]), positive assertions (i.e., [P]) and negative assertions (i.e., [not P]). Consider the following two questions.

(31)   Petr   jel   na dovolenou | že?
        Peter.NOM go.PST on holiday.PREP   že
        Peter went on holiday, didn't he?

(32)   Petr   nejel   na dovolenou | že?
        Peter.NOM not-go.PST on holiday.PREP   že
        Peter didn't go on holiday, did he?

In (31), the speaker has reasonable grounds to assume that *Petr* is going on holiday.
In (32), on the other hand, he has reasonable grounds to assume that he is not
going on holiday. However, in each case, the question asks whether the relevant
assumption is valid or not[4]. This is not necessarily the case in the following pair of
examples.

(33)   K Vánocům mi   koupíte   počítač | že?
        for Christmas me.DAT buy.FUT computer.ACC   že
        You will buy me a computer for Christmas, right?

(34)   K Vánocům mi   nekoupíte ponožky | že?
        for Christmas me.DAT not-buy.FUT computer.ACC   že
        You will not buy me socks for Christmas, right?

The sentences above are perhaps best imagined to be uttered by a demanding
child. In both (33) and (34), the speaker seeks a confirmation of the proposition
expressed in the anchor. Tag questions of this type are biased in that the speaker
expects the hearer to express agreement with the presupposed proposition. The
properties of the four types of tag questions are summarised below.

(35)   Tag questions with strong positive bias (e.g., (33))

        a.   Anchor: has positive polarity.
        b.   Question: asks for a confirmation of [P].
        c.   Tag: *ano* = [P], *ne* = [not P].
        d.   Anchor-Tag Correspondence: *ano* = true, *ne* = false.
        e.   Correlate: positive rising declaratives.

(36)   Tag questions with weak positive bias (e.g., (31))

        a.   Anchor: has positive polarity.
        b.   Question: asks [P or not P].
        c.   Tag: *ano* = [P], *ne* = [not P].
        d.   Anchor-Tag Correspondence: *ano* = true, *ne* = false.
        e.   Correlate: positive questions.

---

[4]Of course, given some different context, these examples would be compatible with the
interpretation of the anchor as a positive assertion and a negative assertion, respectively. What
is crucial here is that they do not have to be interpreted as such.

(37) Tag questions with strong negative bias (e.g., (34))

    a. Anchor: has negative polarity.

    b. Question: asks for a confirmation of [not P].

    c. Tag: *ano* = [not P], *ne* [not not P].

    d. Anchor-Tag Correspondence: *ano* = true, *ne* = false.

    e. Correlate: negative rising declaratives.

(38) Tag questions with weak negative bias (e.g., (32))

    a. Anchor: has negative polarity.

    b. Question: asks [P or not P].

    c. Tag: *ano* = [P], *ne* = [not P].

    d. Anchor-Tag Correspondence: *ano* = false, *ne* = true.

    e. Correlate: negative questions (with expletive negation).

The tag questions with strong bias can be taken to be polite requests. On this view, (33) and (34) can be paraphrased as 'I want you to buy me a computer for Christmas.' and 'I do not want you to buy me socks for Christmas.', respectively. The tag questions with weak bias would be interpreted differently: (31) and (32) can be paraphrased as 'I think that Peter went on holiday. What do you think?' and 'I think that Peter did not go on holiday. What do you think?', respectively. In relation to this, it is important to note that some authors argued that positive *yes-no* questions are, or at least can be, unbiased. Admittedly, tag questions behave differently, in that they are always biased (cf. Ladd (1981)).

4.4 Other Properties of Tag Questions

The correspondence between a minimal answer to a *yes-no* question and a tag in a tag question is not a direct one. More concretely, tags comprising *že* are not minimal answers, but rather embedded minimal answers. Gruet-Skrabalova (2015) notes that minimal answers to *yes-no* questions can be embedded under *že*. Consider the following question, repeated from above.

(15) Lhal ti Petr?
    lie.PST you.DAT.CL Peter.NOM
    Did Peter lie to you?

In the answer, either the polarity marker or the verb might be embedded under *myslím* along with *že*. In this case, *že* functions as a linker, connecting the matrix clause with the embedded clause. (The curly brackets indicate a set of mutually exclusive alternatives. The alternatives are separated by a slash.)

(39) Myslím | že {ano/lhal}.
      think.PRS   že yes/lie.PST
      I think that he did/lied.

What is crucial is that, unlike in unembedded minimal answers, in embedded minimal answers, the polarity marker and the verb cannot co-occur. Compare the example below to (18).

(40) *Myslím | že ano lhal.
       think.PRS   že yes lie.PST
       Intended: I think that he did/lied.

Since the polarity marker and the verb cannot co-occur in question tags either (see (13) and (14)), the tag questions with *že* introduced above should be analysed as questions with embedded minimal answers attached to them. Since such tag questions are either positively or negatively biased, the polarity marker or the verb may be dropped, as they can be recovered by the addressee from the polarity of the anchor.

Based on the fact that question tags and minimal answer under embedding may not contain any clitics, it is likely that IP is elided. This follows from the assumption that clitics are located in I. In such constructions, *že* is located in C. However, in cases where clitics survive along with the rest of the syntactic material, it is possible that *že* is as low as I.

4.5   Conclusion

Above, I argued for a relation between tag questions and *yes-no* questions. More precisely, anchors were argued to be interpreted as various types of *yes-no* questions, and tags as embedded answers to such questions. The distribution of the polarity markers in question tags was derived via the assumption that the tag must confirm the polarity of the anchor with respect to the interpretation of the question associated with that anchor. The relationship between the anchor and the tag was argued to be mediated by *že*. Assuming that *že* in tag questions is a linker allows us to derive the impossibility of having certain polarity markers in the tag. Remember that the anchor was claimed to express a certain degree of bias, depending on the interpretation of the question associated with it. The anchor is an expression associated with the speaker. Since the anchor is linked to the tag, the tag itself is also associated with the speaker. Given this, it has to conform to the bias expressed by the speaker in the anchor. As a consequence, the polarity markers in the tag must not contradict this bias. Therefore, tags comprising

polarity items that deny the proposition expressed by the anchor are not licensed (see (7)). This is, of course, not the case in minimal answers to polar questions, where denial is possible (see (17)-(19)), because it is not issued by the speaker, but rather by the addressee. However, it remains an open issue whether tag questions without *že* behave differently to tag questions with *že*. Similarly, to say anything precise about the syntactic position of the particle in tag questions would require a thorough analysis of the Czech left periphery with respect to polarity, ellipsis etc. Unfortunately, this is beyond the scope of this thesis.

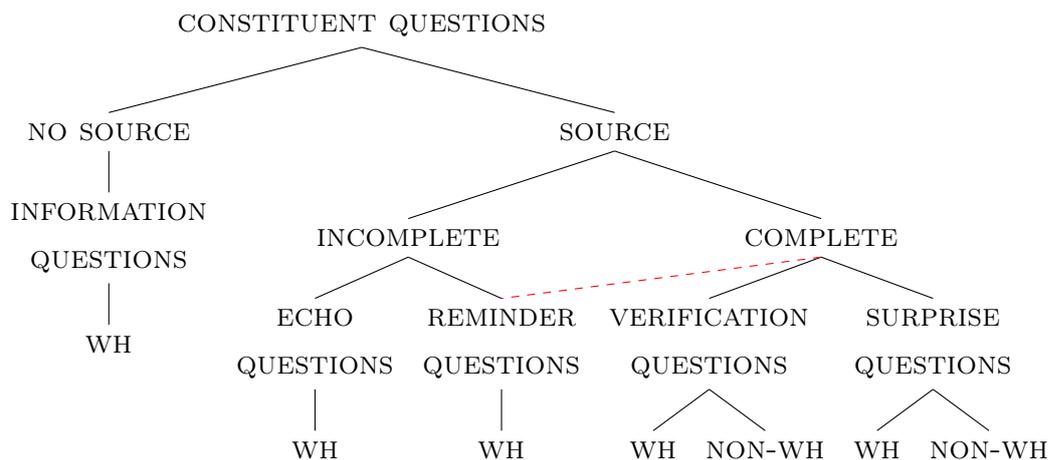# Chapter 5

Constituent Questions

5.1   Introduction

In this chapter, I consider the linguistic properties of *že* in constituent questions. The term 'constituent questions' is usually taken to be synonymous with the term '*wh*-questions'. However, here it is used as a cover term for questions involving *wh*-elements and corresponding questions not involving *wh*-elements. Let us consider the diagram in (1) below. Constituent questions are subdivided into questions that either have or do not have a discourse antecedent[1]. Those that do not have a discourse antecedent function as information questions. Such questions are used to request new information from the addressee. Constituent questions that do have a discourse antecedent are further divided according to the nature of the antecedent. Those questions that have a discourse antecedent that is incomplete function either as echo questions or as reminder questions. The antecedent of an echo question is incomplete by virtue of being partially or fully inaudible. Echo questions are used to request a repetition of the inaudible expression. The antecedent of a reminder question is incomplete by virtue of being partially or fully forgotten. Reminder questions are used to request a repetition of the forgotten expression. Reminder questions formally belong to the class of constituent questions that have a discourse antecedent that is fully complete. However, due to the fact that at least some part of the antecedent is not accessible to the speaker, they behave essentially like echo questions, the antecedent of which is fully or partially incomplete. Constituent questions that have a discourse antecedent that is fully complete and accessible to the speaker function either as verification questions or as surprise questions. Verification questions are used to request a verification of some part of the antecedent expression. Surprise questions are used to express surprise over some part of the antecedent expression. The licensing conditions on

---

[1]The antecedent is taken to be an utterance.

the use of *wh*-elements and non-*wh*-elements in the given question boil down to whether there is a discourse antecedent and, if there is, whether it is fully accessible to the speaker: only constituent questions that have a discourse antecedent that is fully accessible to the speaker may contain non-*wh*-elements. The above typology of constituent questions is schematised below[2].

(1)   A typology of constituent questions.

```
                        CONSTITUENT QUESTIONS


        NO SOURCE                            SOURCE
            |
       INFORMATION
                             INCOMPLETE              COMPLETE
        QUESTIONS

            |          ECHO      REMINDER  VERIFICATION   SURPRISE
           WH
                     QUESTIONS  QUESTIONS   QUESTIONS     QUESTIONS

                         |          |
                        WH         WH      WH  NON-WH   WH  NON-WH
```

The typology proposed above is meant to capture the properties of constituent questions in Czech[3]. Since the particle *že* cannot appear in information questions, the focus here will be only on questions that have a discourse antecedent. Let us first consider the following set of examples. The questions in (3), (4), (5) and (6) are all taken to be linked to the discourse antecedent in (2)[4].

(2)   Petr   si   přečetl   [Vojnu a   MÍR]$_\text{F}$.
        Peter.NOM REFL.CL read.PST War   and Peace.ACC
        Peter read War and Peace.

(3)   Petr   si   přečetl   [CO]$_\text{F}$?
        Peter.NOM REFL.CL read.PST what.ACC
        Peter read what?

---

[2]Of course, constituent questions can also function as rhetorical questions, matching questions etc. Since the main aim of this chapter is not to provide an exhaustive account of constituent questions, only the most relevant types are considered here.

[3]The various types of constituent questions can be syntactically ambiguous, as they can all take the form of a question involving *wh*-movement. However, I believe that the different types of constituent questions can be distinguished by a combination of pragmatics and prosody. In other words, the context within which the question is used and the way it is realised disambiguate.

[4]It is assumed that the interaction took place in the following context. Disregarding the speaker and the addressee, there was 1 individual (Peter), who was familiar to both the speaker and the addressee. Peter recently read a book, and the speaker informed the addressee about it. The rest of the contextual information is supplied in the main text.

(4) [CO]_F si Petr přečetl?
what.ACC REFL.CL Peter.NOM read.PST
What did Peter read?

(5) Petr si přečetl [Vojnu a MÍR]_F?
Peter.NOM REFL.CL read.PST War and Peace.ACC
Peter read War and Peace?

(6) [Vojnu a MÍR]_F si Petr přečetl?
War and Peace.ACC REFL.CL Peter.NOM read.PST
War and Peace Peter read?

An echo question, which can take the form of (3) or (4), would be used if the addressee did not hear that it was *Vojnu a Mír* that Peter read. A reminder question, which can take the form of (4), would be used if the addressee heard that it was *Vojnu a Mír* that Peter read, but forgot it. A verification question, which can take the form of (3), (4), (5) or (6), would be used if the addressee was not certain whether he heard that it was *Vojnu a Mír* that Peter read. A surprise question, which can take the form of (3), (4), (5) or (6), would be used if the addressee heard that it was *Vojnu a Mír* that Peter read, but found it surprising.

Interestingly, *že* is licensed in (4) and (6), but not in (3) and (5). Consequently, the particle can appear only in questions that have a discourse antecedent and that involve dislocation of a (non-)*wh*-element. The consensus in the literature seems to be that the particle *že* forces its containing utterance to be linked to the immediately preceding discourse. Veselovská (1993) and Gruet-Skrabalova (2012*a*), for instance, note that *wh*-questions with *že* are interpreted as echo questions. However, the two authors differ as to what the syntactic position of the particle is: Veselovská (1993) assumes that *že* is located inside the fronted phrase, and Gruet-Skrabalova (2012*a*) that it is a complementiser[5].

Below, I use novel data to demonstrate that the two assumptions about the particle's distribution are not mutually exclusive. I claim that there are two types of *že*: the phrase-internal *že* (which is restricted to occur within SpecCP) and the phrase-external *že* (which is restricted to occur in C). Combining this dual analysis of *že* with some further assumptions makes it possible to account for the distribution of the particle in a variety of constituent questions. In the remaining sections of this chapter, I focus exclusively on constituent questions involving *wh*-elements. As opposed to constituent questions involving non-*wh*-elements, when a complex phrase is fronted, these questions sometimes allow *že* to appear inside it. The examples presented below are taken to be echo questions. No contextual

---

[5]Working in the cartographic framework, Gruet-Skrabalova (2012*a*) claims that *že* is located in Foc. The important point is that she treats the particle as a functional head located in the left periphery of the clause.

information is provided: in each case, it is assumed that the inaudible expression being questioned corresponds to the *wh*-element of the question, and that the rest of the question, apart from any occurrences of *že*, is marked as 'given' in the sense of Schwarzschild (1999).

5.2   Distributional Puzzle

Wh-questions are usually formed by fronting all *wh*-elements[6]. At least since Rudin (1988), the landing site of the first *wh*-element is assumed to be SpecCP[7]. In general, *wh*-movement can target complements and adjuncts. When *že* appears in the structure[8], it must come immediately after the first fronted *wh*-element. Consequently, clitics have to follow the particle[9].

(7)   Co   že si   Petr   přečetl?
       what.ACC že REFL.CL Peter.NOM read.PST
       Again, what did Peter read?

It could be assumed either that *že* forms a constituent with the *wh*-element or that clitics may not always demarcate the first major constituent of their containing clause. As was mentioned above, Veselovská (1995) assumes that clitics cannot go past IP. Given this, the question of whether *že* forms a constituent with the *wh*-element (see (8)), or not (see (9)), cannot be answered by clitic placement.
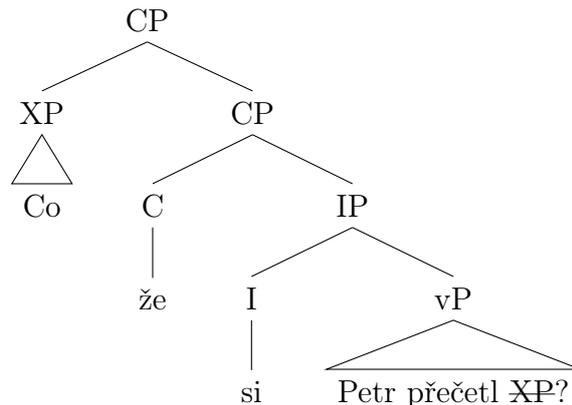
(8)   Phrase-internal *že*.

---
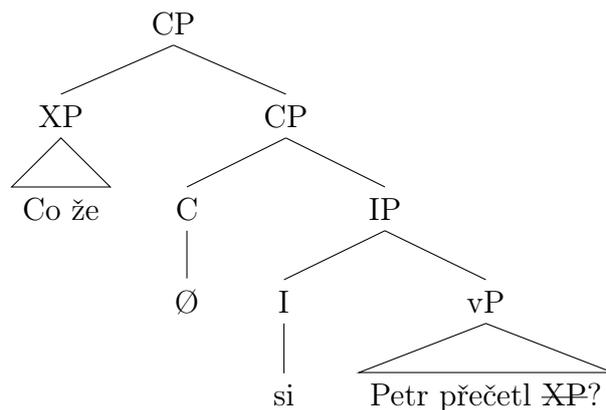
[6]Certain constituent question allow the *wh*-element(s) to remain *in situ*. Information questions with *in-situ wh*-element(s) are typically introduced by the conjunction *a* ('and').

[7]Not everyone agrees with this. Sturgeon (2008), for instance, considers SpecIP to be the landing site of the first *wh*-element. This assumption would is compatible with my assumption that *že* could appear as low as I.

[8]For ease of exposure, unless indicated otherwise, *wh*-questions with *že* presented in the reminder of this chapter are taken to be used as reminder questions.

[9]Since *že* can bear stress and host clitics, it cannot be considered to be part of the clitic cluster.

(9)   Phrase-external *že*.

```
                         CP
                  _____/_____
                 /              \
               XP               CP
              /\          _____/_____
             /__\        /              \
             Co         C                IP
                        |           _____/\_____
                        |          /            \
                       že         I              vP
                                  |          ___/\___
                                  |         /        \
                                  si     Petr přečetl X̶P̶?
```

5.3   Pied-Piping and Left Branch Extraction

Gruet-Skrabalova (2011, 2012*a*) assumes that *že* in *wh*-questions is a complementiser. One prediction that follows from such an assumption is that *že* should always follow the (first) fronted constituent comprising the *wh*-element, regardless of its complexity. In (10), the phrase containing the *wh*-element is allowed to undergo pied-piping. In (11), the *wh*-element itself is allowed to undergo left branch extraction.

(10)   Jakou knihu   si   Petr   přečetl?
        which book.ACC REFL.CL Peter.NOM read.PST
        Which book did Peter read?

(11)   Jakou si   Petr   přečetl knihu?
        which REFL.CL Peter.NOM read.PST book.ACC
        Which book did Peter read?

As predicted, *že* is allowed to follow the fronted constituent, regardless of its complexity.

(12)   Jakou knihu   že si   Petr   přečetl?
        which book.ACC že REFL.CL Peter.NOM read.PST
        Which book did Peter read?

(13)   Jakou že si   Petr   přečetl knihu?
        which že REFL.CL Peter.NOM read.PST book.ACC
        Which book did Peter read?

However, *že* can appear not only in a position immediately following the fronted constituent (see (12) and (13)), but also in a position within it (see (14)).

(14) Jakou že knihu   si   Petr   přečetl?
      which že book.ACC REFL.CL Peter.NOM read.PST
      Which book did Peter read?

Maintaining the assumption that *že* is in C and that clitics are in I, *knihu* has to be located in SpecIP. Remember that one of the claims made in the previous chapters was that *že* could also be located in I. The fact that the particle can follow both *jakou* and *knihu* at the same time (see (15)) lends support to this claim[10].

(15) ?Jakou že knihu   že si   Petr   přečetl?
      which že book.ACC že REFL.CL Peter.NOM read.PST
      Which book did Peter read?

The sentence in (14) might therefore be derived by a combination of movement of *jakou knihu* to SpecIP and sub-extraction of *jakou* to SpecCP (a solution following Gruet-Skrabalova's (2012*a*) assumption that *že* is a complementiser). Remember that this sort of compound movement operation was independently proposed by Sturgeon (2008) to account for contrastive left dislocation. The example in (14) could also be derived by moving *jakou knihu* to SpecCP directly and by suffixing *že* to *jakou* (a solution following Veselovská's (1993) assumption that *že* is located inside the fronted phrase).

5.4 Towards Feature Percolation

To test the validity of the two hypotheses about the distribution of *že*, it is important to consider cases of pied-piping and left branch extraction that are more complex than those presented above. The following examples show that movement of both *o jak dlouhé knize* and *o jaké dlouhé knize* is possible.

(16) O   jak dlouhé knize   si   Petr   přečetl?
      about how long   book.LOC REFL.CL Peter.NOM read.PST
      How long a book did Peter read about?

(17) O   jaké dlouhé knize   si   Petr   přečetl?
      about which long   book.LOC REFL.CL Peter.NOM read.PST
      Which long book did Peter read about?

As predicted, *že* is allowed to follow the fronted constituent in both cases.

---

[10]The remarks made above and below would also apply to other, alternative analyses. For instance, an advocate of the cartographic approach would have to assume the possibility of *že* being realised in two different positions along the clausal spine and above clitics.

(18) O   jak dlouhé knize   že si   Petr   přečetl?
      about how long   book.LOC že REFL.CL Peter.NOM read.PST
      How long a book did Peter read about?

(19) O   jaké dlouhé knize   že si   Petr   přečetl?
      about which long   book.LOC že REFL.CL Peter.NOM read.PST
      Which long book did Peter read about?

It might also be possible for *o jak dlouhé* and *o jaké dlouhé* to strand *knihu*. However, only the former is grammatical in Czech (cf. Bašić's (2004) account of left branch extraction in Serbian)[11].

(20) O   jak dlouhé si   Petr   přečetl   knize?
      about how long   REFL.CL Peter.NOM read.PST book.LOC
      How long a book did Peter read about?

(21) *O   jaké dlouhé si   Petr   přečetl   knize?
      about which long   REFL.CL Peter.NOM read.PST book.LOC
      Intended: Which long book did Peter read about?

Consequently, *že* should be allowed to appear between *o jak dlouhé* and *knize*, but not between *o jaké dlouhé* and *knize*. This prediction is borne out.

(22) O   jak dlouhé že knize   si   Petr   přečetl?
      about how long   že book.LOC REFL.CL Peter.NOM read.PST
      How long a book did Peter read about?

(23) *O   jaké dlouhé že knize   si   Petr   přečetl?
      about which long   že book.LOC REFL.CL Peter.NOM read.PST
      Intended: Which long book did Peter read about?

Interestingly, it is not possible to move either *o jak* or *o jaké* on their own.

(24) *O   jak si   Petr   přečetl   dlouhé knize?
      about how REFL.CL Peter.NOM read.PST long   book.LOC
      Intended: How long a book did Peter read about?

(25) *O   jaké si   Petr   přečetl   dlouhé knize?
      about which REFL.CL Peter.NOM read.PST long   book.LOC
      Intended: Which long book did Peter read about?

Given the above, *že* should be allowed to appear neither between *o jak* and *dlouhé knize*, nor between *o jaké* and *dlouhé knize*. However, this prediction is contradicted by the following data.

---

[11]Some native speakers consider (20) to be ungrammatical. In addition, some native speakers do not find (21) completely ungrammatical. In other words, there is some inter-speaker variation.
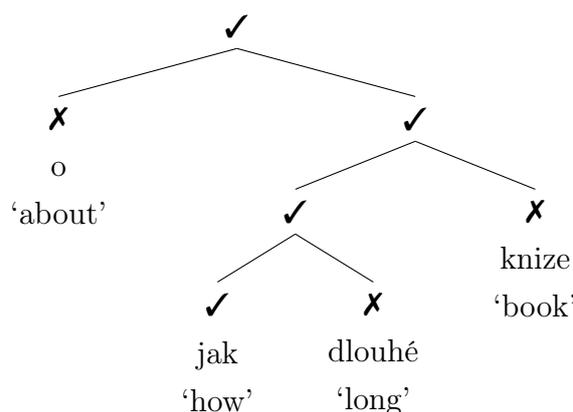
(26)   O   jak že dlouhé knize   si   Petr   přečetl?
        about how že long   book.LOC REFL.CL Peter.NOM read.PST
        How long a book did Peter read about?

(27)   O   jaké že dlouhé knize   si   Petr   přečetl?
        about which že long   book.LOC REFL.CL Peter.NOM read.PST
        Which long book did Peter read about?

Crucially, the sentences in (26) and (27) cannot be derived by means of sub-extraction. This follows from the fact that neither *o jak* in (24) nor *o jaké* in (25) can strand *dlouhé knize*[12].

5.5 Analysis

The now classic view is that feature percolation is what allows pied-piping. In *wh*-questions, it is the *wh*-feature that allows pied-piping. It is possible to assume that *že* in *wh*-questions is suffixed to nodes that bear the *wh*-feature. On this view, *že* should be allowed to appear along the path of feature percolation leading from the *wh*-element to the root node of the moved constituent containing it. The trees in (28) and (29) represent the relevant parts of sentences in (16) and (17), respectively. While nodes marked as ✓ bear the *wh*-feature, nodes marked as ✗ do not. The particle *že* can be suffixed to the former, but not the latter. This analysis makes correct predictions about the distribution of *že* in the relevant examples above.

(28)   A partial structural representation of (16).

---

[12]Gisbert Fanselow (p.c.) points out the possibility of analysing the contrast between (24)/(25) and (26)/(27) in terms of a ban on long movement.

(29)   A partial structural representation of (17).

```
                        ✓
             ┌──────────┴──────────┐
             ✗                     ✓
             o              ┌──────┴──────┐
          'about'           ✓             ✗
                          jaké      ┌─────┴─────┐
                         'which'    ✗           ✗
                                  dlouhé       knize
                                  'long'       'book'
```
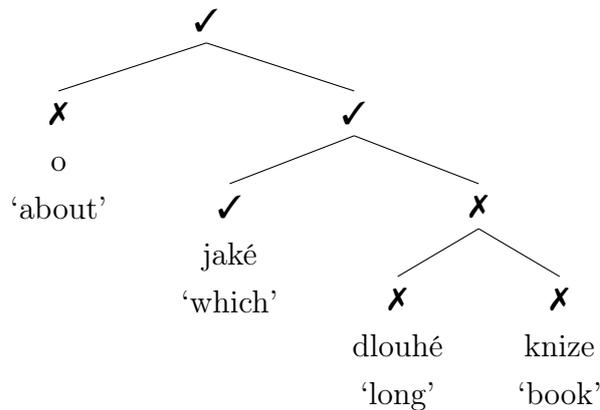
The syntactic structure of *o jak dlouhé knize* differs from that of *o jaké dlouhé knize*. In (16), *jak* modifies AP. In (17), *jaké* modifies NP. The following two examples, which together form a minimal pair, bear testimony to this. While *dlouhé* can be omitted in (30), it cannot be omitted in (31). The reason for this is that *jak* is dependent on the presence of AP.

(30)   O   jak *(dlouhé) knize   si   Petr   přečetl?
        about how long   book.LOC REFL.CL Peter.NOM read.PST
        How (long) a book did Peter read about?

(31)   O   jaké (dlouhé) knize   si   Petr   přečetl?
        about which long   book.LOC REFL.CL Peter.NOM read.PST
        Which (long) book did Peter read about?

A perceptive reader would have noticed that the analysis proposed above needs to be refined to disallow structures of the following type. It is plausible to assume that one of the adjacent instances of *že* gets deleted under haplology (see Neeleman & van de Koot (2006), and Nevins (2010)). The same mechanism can be utilised to (iteratively) reduce any two adjacent instances of *že* into one.

(32)   *O   jak dlouhé knize   že že si   Petr   přečetl?
         about how long   book.LOC že že REFL.CL Peter.NOM read.PST
         Peter read about how long a book?

(33)   *O   jaké dlouhé knize   že že si   Petr   přečetl?
         about which long   book.LOC že že REFL.CL Peter.NOM read.PST
         Peter read about which long book?

5.6 Further Data

5.6.1 *In-Situ Wh*-Phrases

It is sometimes possible to leave the phrase containing the *wh*-element *in situ*. This option is readily available to echo questions. The following examples show that both *o jak dlouhé knize* and *o jaké dlouhé knize* can remain *in situ*. Crucially, *že* cannot appear in any of the positions in which it is allowed to appear when the relevant phrase undergoes movement[13]. This observation is also true for the other paradigms presented above and below.

(34)   Petr   si   přečetl   o   jak (*že) dlouhé (*že)
       Peter.NOM REFL.CL read.PST about how že   long   že
       knize   (*že)?
       book.LOC že
       Peter read about how long a book?

(35)   Petr   si   přečetl   o   jaké (*že) dlouhé (*že)
       Peter.NOM REFL.CL read.PST about which že   long   že
       knize   (*že)?
       book.LOC že
       Peter read about which long book?

Maintaining the assumption that the *wh*-feature is what allows pied-piping in *wh*-questions, one could argue that the *wh*-feature does not percolate in the absence of *wh*-movement. While this argument seems to be a reasonable one[14], it does not provide any explanation as to why *že* in (34) and (35) above cannot follow the *wh*-element itself. After all, the *wh*-element comes from the lexicon with the *wh*-feature. Hence, the impossibility of having *že* immediately following the *wh*-element remains unexplained by the feature percolation analysis.

5.6.2 *Wh*-Phrases under Sluicing

It is possible for *wh*-questions with *že* to be embedded. As is going to be shown in the next chapter, such questions are interpreted as echo questions. One of the constructions that allows embedded *wh*-questions with *že* is sluicing. I follow Merchant (2001) in assuming that '... no non-operator material may appear in COMP ...' in sluicing, and that the fronted constituent containing a *wh*-element moves in exactly the same way as it does in non-sluicing constructions involving *wh*-fronting (p.62). If it were true that *že* was always located inside SpecCP, then

---

[13]As was noted at the outset of this chapter, *že* cannot appear in the left periphery of the clause either.

[14]Truckenbrodt (2013) proposes a similar restriction.

it should always be possible to retain the particle under sluicing. However, in (36) and (37), the rightmost *že* cannot be retained.

(36)   Jakub   i   Marie   věděli,   že   si   Petr
       Jacob.NOM and Mary.NOM know.PST that REFL.CL Peter.NOM
       přečetl   o   nějaké dlouhé knize,   ale Jakub   chtěl
       read.PST about some   long   book.LOC but Jacob.NOM want.PST
       vědět   o   jak (že) dlouhé (že) knize   (*že).
       know.INF about how že   long   že   book.LOC že
       Jacob and Mary knew that Peter read about some long book, but Jacob
       wanted to know about how long a book.

(37)   Jakub   i   Marie   věděli,   že   si   Petr
       Jacob.NOM and Mary.NOM know.PST that REFL.CL Peter.NOM
       přečetl   o   nějaké dlouhé knize,   ale Jakub   chtěl
       read.PST about some   long   book.LOC but Jacob.NOM want.PST
       vědět   o   jaké (že) dlouhé (*že) knize   (*že).
       know.INF about which že   long   že   book.LOC že
       Jacob and Mary knew that Peter read about some long book, but Jacob
       wanted to know about which long book.

One could argue that the rightmost *že* is a proclitic that is unable to be hosted by a phonologically null syntactic element[15]. However, while the particle might sometimes behave like an enclitic, the fact that the peripheral *že* in the non-sluicing counterparts of (36) and (37) can host clitics (see (18) and (19)) means that it does not behave like a proclitic. Hence, the impossibility of retaining the rightmost instance of *že* under sluicing remains unexplained by the feature-percolation analysis.

5.7   Conjoined *Wh*-Phrases

Apart from conjoining two *wh*-elements, it is also possible to conjoin two constituents containing *wh*-elements. The conjuncts might be D-linked. The syntactic structure of coordinated *wh*-questions has been analysed as either mono-clausal or bi-clausal or both (see Citko & Gračanin-Yuksek (2013)). In order to force a mono-clausal structure, and thereby avoid unnecessary complications pertaining to bi-clausal syntactic analysis of coordinated *wh*-questions, it is plausible to use a reciprocal predicate[16]. The distribution of *že* is restricted to the following. Importantly, the particle can appear twice in the sentence: it can follow either *jaký* and *jaká* or *jaký* and *dívka*.

---

[15]I would like to thank Jason Merchant for pointing out this possibility.

[16]Even though it is available, I ignore here the odd reading on which the boy kissed only himself and the girl kissed only herself.

(38) Jaký (že) chlapec   (*že) a   jaká (že) dívka   (že) se
      which že boy.NOM že   and which že girl.NOM že   REFL.CL
      políbili?
      kiss.PST
      Which boy and which girl kissed each other?

Crucially, according to the feature-percolation analysis, *že* should be allowed to appear between *chlapec* and *a*, which it cannot. Since the analysis proposed above does not make correct predictions about the distribution of *že* in a number of constituent questions, it needs to be refined.

5.8   Refinements

The analysis presented above needs to be refined to account for the distribution of the particle *že* in all of the relevant examples above. The amended version of the analysis comprises the following components.

(39) FEATURE PERCOLATION[17]
      A given feature of a node Y can percolate to a node X iff X immediately dominates Y.

(40) FEATURE DOMAIN
      The highest node bearing a given feature marks the domain of that feature.

(41) PERCOLATION BOUNDARY[18]
      In the presence of movement, the *wh*-feature percolates to the root node of the moved constituent containing an element bearing that feature. In the absence of movement, no such percolation takes place.

(42) ADJUNCTION RULE
      The particle *že* can be suffixed to any node that bears the *wh*-feature and that is followed by another phonologically non-null node within the domain of that feature.

The amended analysis proposed above has the capacity to account for the relevant *in-situ* and sluicing constructions. However, it does not have the capacity to fully account for the relevant coordination constructions. Consequently, further

---

[17]Some authors use the term 'percolation' to refer to the projection of a given feature across phrasal boundaries. However, I use it to refer to the same process both within and across phrasal boundaries.

[18]Truckenbrodt (2013), who considers percolation of the *F*-feature in constructions involving *wh*-elements, makes a distinction between *in-situ* information questions and *in-situ* echo questions. In case of the former, the *F*-feature percolates to the *wh*-word. In case of the later, it stays on the *wh*-morpheme.

assumptions need to be made in order for the amended analysis to work. Before these assumptions are outlined, let us consider how the refined analysis accounts for most of the distributional possibilities of the particle.
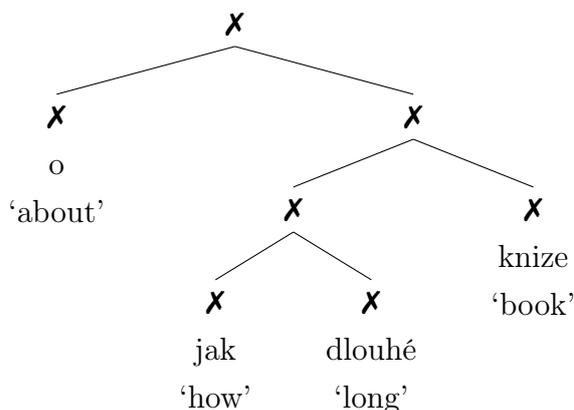
5.9   Applying the Amended Analysis

In the structural representations below, *že* can be adjoined to nodes marked as ✓, but not to nodes marked as ✗. The marking thus no longer indicates whether the given node bears the *wh*-feature or not. The decision to change the way the notation works was motivated by a desire to show the output(s) of applying all the components of the refined analyses.

5.9.1   *In-Situ Wh*-Phrases

In (34) and (35), the *wh*-feature does not percolate. As a consequence, the feature domain is the *wh*-element itself. Given that no phonologically non-null node follows the node bearing the *wh*-feature within the domain of that feature, adjunction of *že* is not licensed. This reasoning can be extended to all instances of *in-situ wh*-phrases. The trees in (43) and (44) are representations of the relevant parts of sentences in (34) and (35), respectively.

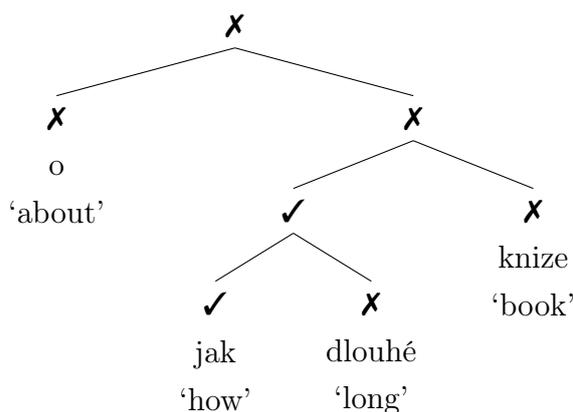(43)   A partial structural representation of (34).

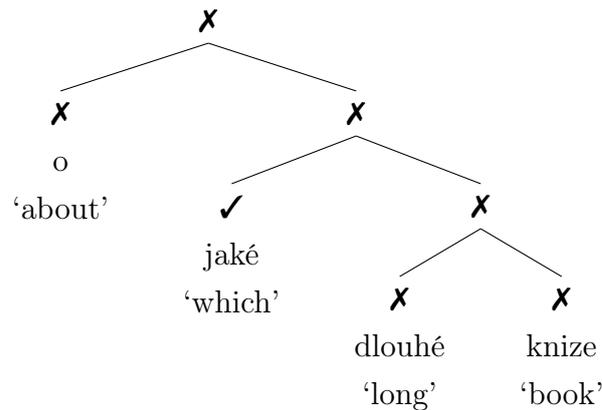(44) A partial structural representation of (35).

```
                        ✗
                   ┌────┴────┐
                   ✗         ✗
                   o    ┌────┴────┐
                'about' ✗         ✗
                 jaké       ┌─────┴─────┐
                'which'     ✗           ✗
                         dlouhé       knize
                          'long'      'book'
```
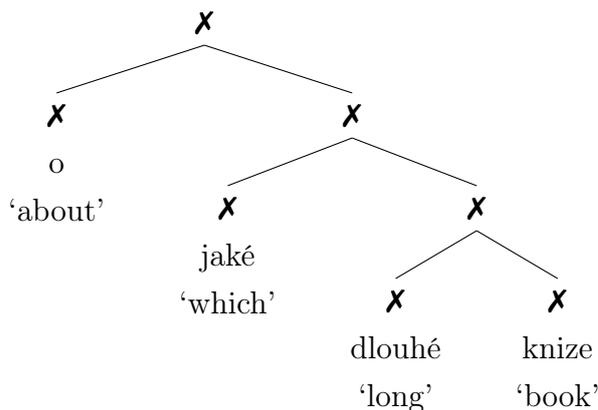
5.9.2 Sluicing *Wh*-Constructions

In (36) and (37), it is possible for *že* to appear only inside the fronted constituent. In both cases, the *wh*-feature percolates to the root node of the fronted constituent containing the *wh*-element. However, *že* is banned from appearing at the right edge within the domain of the *wh*-feature. The only option to have the particle linearly follow the moved phrase would be to phonologically realise it in C. However, as Merchant (2001) notes, '... no non-operator material may appear in C ...' in sluicing (p.62). Consequently, *že* may never immediately follow the fronted constituent in sluicing *wh*-constructions. The trees in (45) and (46) are representations of the relevant parts of the constituents in (36) and (37), respectively.

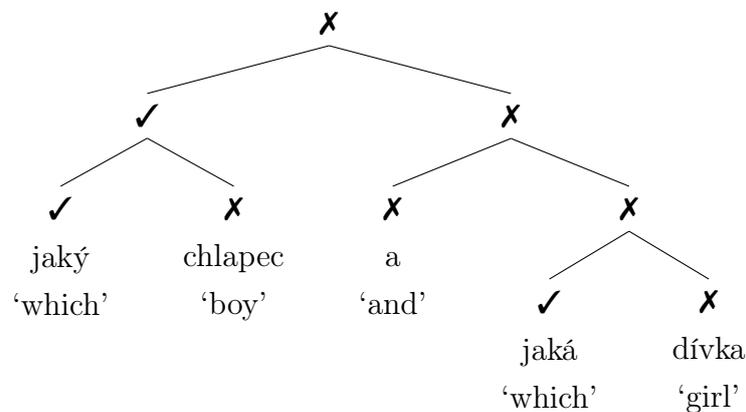(45) A partial structural representation of (36).

```
                     ✗
               ┌─────┴─────┐
               ✗           ✗
               o      ┌────┴────┐
            'about'   ✓         ✗
                   ┌──┴──┐    knize
                   ✓     ✗    'book'
                  jak  dlouhé
                 'how'  'long'
```

(46)   A partial structural representation of (37).

```
                        ✗
                 ┌──────┴──────┐
                 ✗             ✗
                 o         ┌───┴────┐
              'about'      ✓        ✗
                          jaké   ┌──┴──┐
                         'which' ✗     ✗
                              dlouhé  knize
                              'long' 'book'
```

5.9.3   Conjoined *Wh*-Phrases

Accounting for the distribution of *že* in conjoined *wh*-phrases is not as straight-forward as accounting for its distribution in *in-situ* and sluicing *wh*-constructions. In (38), the particle can linearly follow the entire coordination (i.e., *jaký chlapec a jaká dívka*). In order not to wrongly rule out this acceptable configuration, it has to be assumed that the rightmost instance of *že* is located in a position that is outside of the fronted phrase. Following the logic of the argument outlined at the beginning of this text, I assume that the rightmost instance of *že* is located in C. I refer to *že* that can appear within SpecCP as the phrase-internal *že*, and to *že* that can appear in C as the phrase-external *že*. It is necessary to make this distinction to account for all the cases where the particle follows the fronted *wh*-element or a constituent containing the *wh*-element (e.g., (12) and (13)). This follows from the assumption that *že* cannot be suffixed to a node bearing a *wh*-feature if it is not followed by a phonologically non-null node within the domain of that feature.

(47)   A partial structural representation of (38).

```
                          ✗
                 ┌────────┴────────┐
                 ✓                 ✗
            ┌────┴────┐       ┌────┴────┐
            ✓         ✗       ✗         ✗
          jaký     chlapec    a      ┌──┴──┐
         'which'    'boy'   'and'    ✓     ✗
                                    jaká  dívka
                                   'which' 'girl'
```

What is left to account for is the impossibility of the particle to occur between the first conjunct (i.e., *jaký chlapec*) and the conjunction (i.e., *a*). At this point, I would like to claim that the phrase-internal *že* and the phrase-external *že* do not differ only in their distribution. It seems that the phrase-internal *že* is a focus marker. This can be concluded from the fact that everything in its scope or domain of application has to be prosodically prominent[19]. Consider the example below. When the phrase-internal *že* c-commands both *jak* and *dlouhé*, making only *jak* prosodically prominent is not an option. On the other hand, the phrase-external *že* does not preclude such a realisation. Importantly, neither the realisation of the phrase-internal *že* nor the realisation of the phrase-external *že* is obligatory[20].

(48) O   [jak]$_F$ dlouhé ($^{\#}$že) knize   (že) si   Petr
      about how   long   že   book.LOC že   REFL.CL Peter.NOM
      přečetl?
      read.PST
      Peter read about how long a book?

(49) O   [jak dlouhé]$_F$ (že) knize   (že) si   Petr
      about how long   že   book.LOC že   REFL.CL Peter.NOM
      přečetl?
      read.PST
      Peter read about how long a book?

There are also cases where prosodic prominence is required. In the following examples, *muž* ought to be made prosodically prominent. This is due to the fact that it is not plausible for a male human being to get pregnant. If the speaker did not make the word prosodically prominent, he or she would be acknowledging, contrary to the fact, that it is plausible.

(50) $^{\#}$[Jaký]$_F$ (že) muž   (že) otěhotněl?
      which   že   man.NOM že   get-pregnant.PST
      Which man got pregnant?

(51) [Jaký ($^{\#}$že) muž]$_F$ (že) otěhotněl?
      which že   man.NOM že   get-pregnant.PST
      Which man got pregnant?

It could be argued that the phrase-internal *že* and the phrase-external *že* differ in terms of the optionality of prosodic prominence marking. However, such a claim alone would fail to explain why realising *že* between *jaký* and *muž* in (38) results in ill-formedness. After all, the phrase-internal *že* marks the prosodic prominence

---

[19]I assume this to be achieved by emphatic stress. Prepositions and other syntactic elements that usually do not carry additional stress may, perhaps, be excluded.

[20]The very same interpretive effect that the presence of the particle forces can be achieved by prosody alone.

of *jaký* in its scope, and the phrase-external *že* should not prevent *muž* from being prosodically prominent, too.

Given the above, it seems reasonable to suggest that the phrase-internal *že* is an optional focus marker, which, when present, marks the scope of focus. If focus spans over both *jaký* and *muž*, as in (51), the phrase-internal *že* is barred from appearing in between the two words. Hence, the presence of *že* in between *jaký chlapec* and *a* in (38) would force *jaký chlapec* to be prosodically prominent. It is hard, if not impossible, to imagine a context in which the question in (38) could be realised in such a way. That this analysis is along the right track is supported by the fact that the same realisation of the sentence is equally unacceptable without *že*.

5.10   Embedded Constituent Questions

The *wh*-questions with *že* that were presented above could be embedded under predicates that are semantically compatible. Since such questions are realised with a typical declarative prosody, the semantics of the embedding predicate and the context in which the given sentence is used are the only cues to their interpretation. Above, we already saw one such construction (i.e., sluicing). Below, I provide a number of other examples of embedded *wh*-questions with *že*.

(52)   Jakub   chtěl   vědět | kdo   že byl   na
       Jacob.NOM want.PST know.INF   who.NOM že be.PST at
       oslavě.
       party.PREP
       Jacob wanted to know who was at the party.

(53)   Jakub   chtěl   znovu vědět | kdo   že byl   na
       Jacob.NOM want.PST again know.INF   who.NOM že be.PST at
       oslavě.
       party.PREP
       Jacob wanted to know again who was at the party.

(54)   Jakub   byl   překvapený | kdo   že byl   na oslavě.
       Jacob.NOM be.PST surprised   who.NOM že be.PST at party.PREP
       Jacob was surprised who was at the party.

Crucially, *že* remains optional in such cases. This indicates that it is not selected by the matrix clause. Interestingly, constituent questions involving fronted non-*wh*-elements cannot be embedded. This is perhaps due to the fact that in such cases, *že* functions as a semantically empty linker. Since no element may appear between the matrix clause and the linker, movement is not licensed. In these cases, the presence of *že* is compulsory.

(55)   *Jakub   chtěl   vědět | Petr   že byl   na
        Jacob.NOM want.PST know.INF   Peter.NOM že be.PST at
        oslavě.
        party.PREP
        Jacob wanted to know that Peter was at the party.

(56)   *Jakub   chtěl   znovu vědět | Petr   že byl   na
        Jacob.NOM want.PST again   know.INF   Peter.NOM že be.PST at
        oslavě.
        party.PREP
        Jacob wanted to know again that Peter was at the party.

(57)   *Jakub   byl   překvapený | Petr   že byl   na
        Jacob.NOM be.PST surprised   Peter.NOM že be.PST at
        oslavě.
        party.PREP
        Intended: Jacob was surprised that Peter was at the party.

Admittedly, the first two examples above are not only ungrammatical, but also semantically ill-formed. The third one, however, is not. If the movement of *Petr* is behind its ungrammaticality, then the acceptability should improve in its absence. This is exactly what is observed.

(58)   Jakub   byl   překvapený | že Petr   byl   na
        Jacob.NOM be.PST surprised   že Peter.NOM be.PST at
        oslavě.
        party.PREP
        Jacob was surprised that Peter was at the party.

5.11   Conclusion

The data above show that the particle *že* in Czech *wh*-questions cannot always be analysed as a complementiser. The main theoretical claim is that the syntactic behaviour of the particle *že* in Czech *wh*-questions cannot be easily captured without postulating a mechanism akin to feature percolation. This poses serious problems for any theory that disputes the need to postulate feature percolation as a theoretical primitive (cf. Cable, (2010*a*, 2010*b*) and Heck (2008, 2009). The questions in which the particle occurs all require the presence of a discourse antecedent. The relevant questions can receive the desired interpretation without *že*. In the absence of the particle, prosody and context help the addressee to arrive at the interpretation intended by the speaker. Under embedding, the prosodic cue is no longer available. In such cases, it is the semantics of the embedding predicate and context that help the addressee to arrive at the interpretation intended by the speaker. However, when present, the particle forces its containing question to be

interpreted with respect to some discourse antecedent. This is a good indication
that it is not semantically empty. Therefore, *že* in *wh*-questions does not function as
a linker. How to formalise the relation between the given utterance and preceding
discourse remains to be seen.

Chapter 6

Experimenting with Constituent Questions

6.1 Introduction

The categorial distinction between 'given' and 'new' is frequently made in the literature on information structure. While the notions of givenness and/or newness have been extensively debated (Schwarzschild (1999), Kučerová (2007), Šimík & Wierzba (2015)), their relation to the more general notion of information status has received less attention (Prince (1981), Nissim et al. (2004), Götze et al. (2007), Riester (2008)). In this chapter, I report the results of two experimental studies on the syntax-semantics interface, which show that not all information can be classified as exclusively given or new. The results are compatible with a classification system that makes either sub-categorial distinctions (i.e., (given - new (implied - brand new))) or further categorial distinctions (i.e., (given - implied - new)). The main claim is that the representation of discourse needs to be more granular than usually assumed.

The experiments examine the semantics of the particle *že* in matrix (see (1)) and embedded (see (2)) *wh*-questions.

(1) Kdo   že zabil   prezidenta?
    who.NOM že kill.PST president.ACC
    Who killed the president? ('Who' realised with a high rising tone.)

(2) Petr   chtěl   vědět,   kdo   že zabil   prezidenta.
    Peter.NOM want.PST know.INF who.NOM že kill.PST president.ACC
    Peter wanted to know who killed the president.

More specifically, they test the possibility of *že* being sensitive to the level of givenness of the expression that corresponds to the *wh*-element of the question. It

could be argued that givenness is a scale, with 'given' and 'new' being its extremes. On this view, it should be possible to have not only propositions that are given and propositions that are new, but also propositions that are only partially given or new.

(3)   Scale of Givenness

        given $\xleftrightarrow{\text{PARTIALLY GIVEN/NEW}}$ NEW

The two experimental studies reported below make use of this idea. On the one end, there are contexts in which the expression has a discourse antecedent. On the other end, there are contexts in which the expression does not have a discourse antecedent. The former are contexts which invite the use of echo questions, and the latter are contexts which invite the use of information questions. In between these two extremes, there are contexts in which the expression is implied. These should, perhaps, invite the use of information questions, as such contexts do not *per se* introduce a discourse antecedent, but rather only imply its existence.

As was mentioned above, *wh*-questions with *že* can be interpreted as echo questions. Since *že* is always optional in such questions, one might wonder whether this particular interpretation stems from the semantics of the particle. If it did, one would expect native speakers of Czech to prefer to use *že* in *wh*-questions in truly echoic contexts. Furthermore, if the semantics of *že* were identical for both its use in matrix *wh*-questions and its use in embedded *wh*-questions, then one would expect the bias for the use of *že* in echoic contexts to occur in both matrix and embedded *wh*-questions. Since echo questions require the underlying proposition to be part of the common ground, there is also the possibility that the particle itself is sensitive to the degree of givenness. If this were true, then one would expect the native speakers' preference for the use of the particle to decrease with the decrease in the degree of givenness. This is precisely what the two experimental studies reported below tested.

6.2   Experiment I

As Veselovská (1993) and Gruet-Skrabalova (2012*a*) acknowledge, matrix *wh*-questions with *že* are used in truly echoic contexts. The presence of the particle thus seems to be what differentiates echo questions from information questions. While this observation might hold for matrix clauses, it is plausible that it does not hold for embedded clauses. If the semantics of the particle is what causes matrix *wh*-questions to be interpreted as echo questions, then there is a possibility that it might also cause embedded *wh*-questions to be interpreted as echo questions.

Assuming the existence of indirect echo questions would perhaps be inadequate, especially in the absence of any strong intuitions about the validity of such an assumption. The present experimental study therefore seeks to test the possibility of *že* in embedded *wh*-questions being sensitive to the degree of givenness. The results show that native speakers have bias for the use of embedded *wh*-questions with *že* in truly echoic environments. Hence, a distinction should be made not only between direct and reported information questions, but also between direct and reported echo questions.

6.2.1 Materials

The questionnaire used in the present study comprised test items, fillers and practice items. Each of these consisted of a short story followed by a question and two sentences. The individual stories were designed in such a way so as to help the informants contextualise the situation in which the two sentences could have been used. The stories always included some social interaction that was easy enough for the informants to envisage. Even though all the stories were realistic, a small proportion of them presented characters and situations that were rather movie-like. In total, 32 different stories were designed for the experiment. Of these, 18 stories were associated with the test items, 12 with the fillers and 2 with the practice items. To give the reader an idea of how these stories looked, a translation of one of the stories in its three incarnations is included below.

(4)   A sample story from Experiment I (Version 1).
        In the times of recession, it is almost inevitable to reduce the number of employees. Yesterday, Mr Stehno, who is the sole owner of our company, asked Mr Šídlo, who is the head of HR, to come to his office. After a few minutes' dialogue, Mr Stehno told Mr Šídlo that, for economical reasons, it was inevitable for him to fire someone, and then he told him whom he would fire. Just when he was saying whom it would be, the fire alarm went off, and Mr Šídlo heard only: 'I will fire ...'. Mr Šídlo was interested in whom Mr Stehno would fire, so he asked him about that once the fire alarm stopped.

(5)   A sample story from Experiment I (Version 2).
        In the times of recession, it is almost inevitable to reduce the number of employees. Yesterday, Mr Stehno, who is the sole owner of our company, asked Mr Šídlo, who is the head of HR, to come to his office. After a few minutes' dialogue, Mr Stehno told Mr Šídlo that, for economical reasons, it was inevitable for him to fire someone, but he did not tell him whom he would fire. Mr Šídlo was interested in whom Mr Stehno would fire, so he asked him about that.

(6)   A sample story from Experiment I (Version 3).

In the times of recession, it is almost inevitable to reduce the number of employees. Yesterday, Mr Stehno, who is the sole owner of our company, asked Mr Šídlo, who is the head of HR, to come to his office. Mr Šídlo was convinced that Mr Stehno was intending to fire someone, and because he was interested in whom Mr Stehno would fire, he asked him about that.

As has been mentioned above, each story was followed by a pair of sentences. Each pair of sentences presented some kind of minimal contrast. The variation was restricted to one of the following: the presence or absence of a particular word (see (7) and (8)), the choice of a particular word (see (9) and (10)) and the linear position of a particular word (see (11) and (12)). In sum, one sentence could differ from the other sentence within a pair in that it had one more word (see (8)) or one different word (see (10)) or one word in a different position (see (12)).

(7)   Pavel   chtěl   vědět,   kdo   se   vyspal   s
Pavel.NOM want.PST know.INF who.NOM REFL.CL sleep.PST with
Martinem.
Martin.INS
Pavel wanted to know who slept with Martin.

(8)   Pavel   chtěl   vědět,   kdo   že se   vyspal
Pavel.NOM want.PST know.INF who.NOM ŽE REFL.CL sleep.PST
s   Martinem.
with Martin.INS
Pavel wanted to know who slept with Martin.

(9)   Markéta   chtěla   vědět,   jaké   je   v
Markéta.NOM want.PST know.INF what-kind-of be.PRS in
Thajsku   ubytování.
Thailand.LOC accommodation.NOM
Markéta wanted to know what kind of accommodation there is in Thailand.

(10)   Markéta   chtěla   vědět,   jaké   je   v
Markéta.NOM want.PST know.INF what-kind-of be.PRS in
Thajsku   jídlo.
Thailand.LOC food.NOM
Markéta wanted to know what kind of food there is in Thailand.

(11)   Standa   chtěl   vědět,   kam   odletěl míč.
Standa.NOM want.PST know.INF where.LOC fly.PST ball.NOM
Standa wanted to know where the ball flew.

(12)   Standa   chtěl   vědět,   kam   míč   odletěl.
Standa.NOM want.PST know.INF where.LOC ball.NOM fly.PST
Standa wanted to know where the ball flew.

Each test item featured a pair of sentences that presented a contrast between embedded *wh*-questions with and without *že* (see (7) and (8) above). The moved *wh*-expression was either a subject or an object or an adjunct. Similarly, each filler featured a pair of sentences. However, none of the fillers comprised an embedded *wh*-question with *že* (see (11) and (12) above). The moved *wh*-expression was either an object or an adjunct. The minimal contrast was achieved either by inserting a word or by moving a word. Four of the sentences, each in a different filler, were ungrammatical. The ungrammaticality always stemmed from a severe word order violation. The purpose of these sentences was to check whether the informants were paying attention in the course of the test. Finally, the practice items, which were not part of the test *per se*, allowed the informants to get to grips with the answering mechanism. Again, each practice item featured a pair of sentences. One of the sentences included an error that was based on the incompatibility of the answer with the given story, rather than on ungrammaticality. This was meant to make the informants aware of the importance of understanding the stories, as opposed to solely focusing on the syntax of the sentences.

The stories took the form of text. Perhaps the biggest advantage of presenting the stories as text was that the informants could re-read a given story, or its part, at any point of answering the question associated with it. In the test, the stories remained visible for the duration of answering the related questions. In contrast to stories, the sentences took the form of audio recordings. Given that the purpose of the experiment was to reveal something about embedded *wh*-question with *že*, and given that this kind of construction appears only in spoken language, audio recordings offered themselves as an ideal solution. It is very likely that very many informants would have refused the embedded *wh*-questions with *že* had the sentences been presented as text. In other words, drawing on their knowledge of the prescriptive rules of the Czech grammar, the majority of informants would have ruled out these constructions as ungrammatical. The recordings were made in the UCL's phonetic laboratory in Chandler House, 2 Wakefield Street, London, WC1N 1PF. The audio recordings were identical for all three versions of a given story associated with test items.

6.2.2 Design

The experiment assumed a one-way design with three levels of independent variable. The phenomenon studied was the sensitivity of *že* in embedded *wh*-questions to the level of givenness. The closest one can get to embedded *wh*-questions with *že* are embedded *wh*-questions without *že*. Therefore, each test item comprised two sentences that formed a minimal pair consisting of a particular *wh*-question with

že and a corresponding *wh*-question without *že*. To find out about the particle's sensitivity to the level of givenness, the 18 stories associated with the test items were a product of creating three versions of six distinct stories. The first version forced the informant to interpret the embedded clause as a request for repetition by involving an interaction between two persons during which some part of the utterance by the speaker was missed by the listener. The other versions did not force the request-for-repetition reading, and were mere requests for new information. The second version involved the speaker implying that such and such was the case, without explicitly revealing any more relevant details. The third version did not involve any explicit or implicit mention of anything by the speaker, but only a thought on part of the listener, who presumed that such and such was the case. The labels given to the three conditions were Given, Implied and New, respectively. Each of the three conditions formed one level of the single factor of the one-way design.

(13)   An overview of the one-way design with three levels of the independent variable.

Given	Implied	New
že	*že*	*že*
∅	∅	∅

Every informant judged each of the three conditions (i.e., Given, Implied and New) twice, and saw exactly one version of each of the six stories. In addition, every informant saw all twelve fillers and both practice items along with their associated stories. In every script, the order of test items, fillers and practice items was the following.

(14) An overview of the distribution of test items, fillers and practice items within the script.

SCRIPT	
Item Number	Item Type
1	Practice 1
2	Practice 2
3	Filler
4	Story X (Version X)
5	Filler (error)
6	Filler
7	Story X (Version X)
8	Filler
9	Filler (error)
10	Story X (Version X)
11	Filler
12	Filler
13	Story X (Version X)
14	Filler (error)
15	Filler
16	Story X (Version X)
17	Filler
18	Filler (error)
19	Story X (Version X)
20	Filler

Using the above template, three scripts were created, which were then randomised by reversing the order of test items. Together, then, there were six scripts, each falling within one of the three types that were originally created. To ensure that each of the three types of script comprised exactly one version of each story and two exemplars of each experimental condition, the selection method depicted in (15) below was used. Fillers (with and without errors) were randomly inserted into fixed, prefabricated positions. The two practice items appeared in a fixed order, preceding all test items and fillers.

(15)   Experiment I - selection method.

	Version 1 (Given)	Version 2 (Implied)	Version 3 (New)
Story 1	Script 1	Script 2	Script 3
Story 2	Script 3	Script 1	Script 2
Story 3	Script 2	Script 3	Script 1
Story 4	Script 1	Script 2	Script 3
Story 5	Script 3	Script 1	Script 2
Story 6	Script 2	Script 3	Script 1

Instructions for informants appeared at the beginning of the questionnaire. The end was reserved for demographic and other information. More precisely, informants were asked to specify their native language, dialect, age, gender, highest achieved education, occupation and name. The last of these was required only to enable the informants' identification in potential follow-up experimental studies.

6.2.3 Methodology

The elicitation method used in the experiment was Forced Choice. The informants were asked to select one of the two sentences that were offered as answers to the following question: 'Which of the two sentences is, according to you, a better summary of the end of the above story?' The informants were not asked to select the sentence that they thought more grammatical.

It is important to note that, in a Forced Choice task, the choice of one sentence by the informant does not rule out the possibility of him also accepting, even though to a lesser degree, the other sentence. Hence, the choice made reflects the informant's preference of one sentence over some other sentence in a given trial. Crucially, the nature of the elicitation method in combination with the experimental design makes it possible for the results not to show any pattern of preference amongst the informants. In the present study, each informant judged each of the three experimental conditions twice by selecting one of the two sentences each time. Therefore, there might not have been any pattern of preference within the individual informants' choices, for they could have selected a *wh*-question with *že* in one trial of a given experimental condition and a *wh*-question without *že* in the other. The present study, then, allowed every informant to show preference or not.

Forced Choice is less demanding on the informants than, for instance, Magnitude Estimation. The former mode of answering requires a comparison to be made,

while the latter demands a ratio comparison to be given. Even though Magnitude Estimation is becoming increasingly popular amongst linguists, it is important to bear in mind that not all informants are familiar with the task. A great deal of effort is usually required to make sure that error variance due to the informants' unfamiliarity with the task is controlled for. To be on the safe side, the herein-described experiment used a method whose workings were more likely to be known to the informants.

The informants (n=45) were a mixture of students and professionals recruited via email. For convenience, the questionnaire was made accessible on the Internet. The experimental tool used was tailored software developed by Honeypot s.r.o. (www.honeypot.cz). The data gathered during the experiment were imported into an Excel file.

6.2.4 Data filtering and cleaning

As opposed to some other experimental tools (e.g., WebExp), the software used in the present study was not capable of automated filtering or cleaning of any kind. Therefore, the collected data had to be filtered and cleaned manually.

As for filtering, five criteria were set. Their aim was to prepare the data for statistical analyses by excluding informants who had violated at least one of them. First, those who did not select the 'native speaker of Czech' option were excluded. This included informants who selected the 'native speaker of other language' option as well as those who did not provide any specification of their mother tongue. Second, those who completed the questionnaire in five minutes or less were filtered out. The limit of five minutes allowed 15 seconds for each of the 20 questions. The limit might have been too liberal; however, it sufficed to filter out those informants who were most likely not to be paying enough attention to the questionnaire. Third, those who wrongly answered any of the fillers containing an error were excluded. Wrongly answering any of the fillers containing an error was taken to indicate a loss of concentration. Fourth, those who did not specify all of the variables (see (16) below) were filtered out. A statistical analysis with missing data would have been less accurate than one without any missing data. Fifth, those who were familiar with my work were excluded. The chief aim of the herein-listed filtering criteria was to minimise any possible noise in the results. The data left after sifting are used in the reminder of this chapter.

However, it was not enough to filter out the data provided by informants who violated at least one of the five criteria mentioned above. It was also necessary to clean the remaining data before any statistics could have been applied. This was achieved in three steps. First, all fillers and practice items were deleted. This

ensured that the data consisted only of test items. Second, all categories that were not selected to be variables were deleted. This ensured that only those categories that were of statistical interest were considered. Third, the values specified for any of the variables were encoded. The statistical program that was used to analyse the data (i.e., SPSS) required the values to be expressed in numerical form. The cleaned data were used to estimate means with confidence intervals.

6.2.5 Statistical method

Perhaps the most commonly used statistics that could have been applied to the data collected in the present experiment is the Analysis of Variance (ANOVA). However, given the nature of the data, ANOVA would not have been the right choice (see, for instance, Dixon (2008), Jaeger (2008), and Quené & van den Bergh (2008)). First, the data are assumed to be normally distributed in ANOVA. But dichotomous responses produced by the Forced Choice technique do not follow normal distribution. This is so because the informants' performance within any one trial is restricted to range from 0 to 1. To overcome this issue, the responses in each condition can be expressed as proportions. In spite of this provision, however, there might occur ceiling/floor effects when the performance is very good/poor. The distortion effects become less prominent as the performance gets closer to 0.5 for each condition. The problem is that one cannot know *a priori* whether the performance within each condition will be close enough to 0.5 for the distortion effects not to arise. In fact, if one knew that the preference were going to be equal in all conditions, it would make little sense to run an experiment. ANOVA is therefore not well suited to be used with dichotomous data. Second, the observations are assumed to be individual. In a repeated measures design, each informant provides more than one response for each condition. Again, the informants' performance within any one trial is restricted to range from 0 to 1. It follows that, even if the performance proportions for the conditions across informants were close to 0.5, the performance proportions for the conditions within informants could still give rise to ceiling/floor effects were the performance very good/poor. ANOVA is therefore not well suited to be used with repeated measures. The overall conclusion, then, is that running ANOVA on the data collected in the present experiment would be inappropriate.

An alternative is to use one version of Generalised Linear Mixed-effects Model (GLMM). This method provides a more powerful means of analysing the data. Using GLMM requires both random and fixed effects to be explicitly specified in the model. The estimates of neither the effects nor the effect interactions specified in the model are independent: they vary depending on what other terms are specified

in the model. An appropriate strategy is to add one effect or one interaction at a time and then do a model comparison. One can use the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC) to determine the optimal model. As Jaeger (2008) notes, '...the model is considered optimal if it is the model for which the actually observed data are most likely to be observed...' (p.437). AIC and BIC are measures of the model's parsimony. On their own they are not of much use; however, they become useful in model comparison. AIC is often considered to be a more accurate estimator of the fit of a model. The lower the AIC value, the better the model fit. However, one needs to be careful when using information criterions, for they are not always informative. For example, a model with a significant interaction between the predictor and the outcome can receive a higher AIC score than a model without a significant interaction between another predictor and the same outcome. To take AIC as gospel in such a case would not be adequate. The version of GLMM used in the present study has logit as the link function. The underlying data are therefore assumed to be binomially distributed. The link function models the performance proportions as a linear function of the factors in the model. Responses are expressed in terms of response strength. GLMM then uses quasi likelihood to estimate the effects of a model. In a repeated measures design, the multiple performances of an informant are correlated within each condition. The correlation structure selected here is compound symmetry, which is a version of sphericity. GLMM is thus immune to the scaling artifacts that ANOVA suffers from.

6.2.6   Variables

The statistics reported below consider not only the significance of the relation of the independent variable (i.e., condition) and the overall results, but also the significance of the relation of each of the following variables and the overall results.

(16)   An overview of the covariates selected for Experiment I.

      a.   Dialect (categories: Prague; other; do not know)

      b.   Education (categories: below university; university)

      c.   Gender (categories: male; female)

   d.   Language (categories: monolingual; bi/multilingual)

6.2.7 Hypotheses

The following were the hypotheses to be tested by the present experimental study.

(17)   Experimental Hypothesis (H$_1$)[1]
        Participants would have the following preference with respect to the use
        of *wh*-questions with *že*: Given>Implied and Given>New.

(18)   Null Hypothesis (H$_0$)
        Participants would have no preference for either *wh*-questions with *že* or
        wh-questions without *že* in either Given or Implied or New.

6.2.8 Results

There was a highly significant difference ($p<.001$) between Given and Implied/New
(Given>Implied/New). Interestingly, there was no significant difference ($p>.05$)
between Implied and New (Implied=New). In fact, there was no difference at all.

(19)   Experiment I - estimated means with confidence intervals.
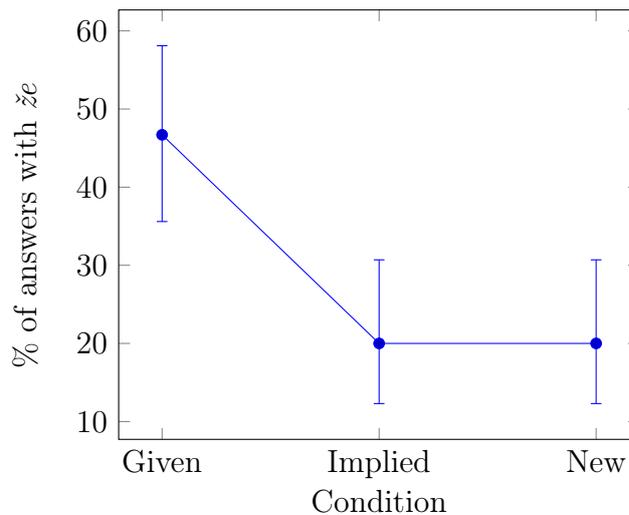
None of the covariates specified in (16) significantly influenced the preference choice
of participants. Given the results, H$_1$ was supported and, by implication, H$_0$ could
be discarded.

    The best fitting model was one that included condition as the only fixed variable.
Some other models scored better in terms of AIC; however, this was expected since
the inclusion of non-significant fixed effects brought the coefficient value closer to
the one associated with the intercept-only model.

---

[1]Note that there was no hypothesis about the preference relation between Implied and New.
In other words, it was not predicted that Implied>New or Implied=New or Implied<New.

6.2.9 Discussion

The results of the present study pattern with Veselovská's (1993) and Gruet-Skrabalova's (2012*b*) observation that matrix *wh*-questions with *že* are used in truly echoic contexts. In other words, native speakers prefer to use *wh*-questions with *že* in both matrix and embedded clauses when the context is truly echoic. Given this, one could contemplate the idea of describing the two as direct and reported echo questions, respectively.

What is interesting about the results is the fact that Implied and New conditions grouped together. This could be taken to indicate that the particle differentiates echoic from non-echoic contexts. However, it cannot be taken to indicate that it differentiates between given and new expressions. The level of givenness decreases from Given to New, with Implied being between the two. One possibility is that no such distinction should be made. Perhaps, expressions are either given or new. On the other hand, the fact that a distinction between the two conditions (i.e., Implied and New) did not show does not prove that it does not exist. It is possible that a larger sample would yield different results.

6.3 Experiment II

The conclusion reached in the previous section was that native speakers prefer to use matrix and embedded *wh*-questions with *že* in truly echoic contexts. It was also noted that the particle might not be sensitive to the level of givenness in non-echoic contexts. Since it is generally more difficult to provide judgments on embedded structures than it is to provide judgements on matrix structures, the present experimental study seeks to test the possibility of *že* in matrix *wh*-questions being sensitive to the degree of givenness. More precisely, it considers the use of matrix *wh*-questions with *že* in non-echoic environments. The results show that native speakers have bias for the use of matrix *wh*-questions with *že* in environments that involve a higher degree of givenness of the utterance containing the expression corresponding to the *wh*-element of the question. Consequently, at least in matrix *wh*-questions, *že* is sensitive to the degree of givenness in contexts that do not invite the use of echo questions.

6.3.1 Materials

The questionnaire used in the present study comprised test items, fillers and practice items. Each of these consisted of a short story followed by a question and two sentences. The individual stories were taken directly from Experiment I. They were designed in such a way so as to help the informants contextualise the

situation in which the two sentences could have been used. In total, 26 different stories were part of the experiment. Of these, 12 stories were associated with the test items, 12 with the fillers and 2 with the practice items.

As has been mentioned above, each story was followed by a pair of sentences. Similarly to Experiment I, each pair of sentences presented some kind of minimal contrast. The variation was restricted to one of the following: the presence or absence of a particular word (see (20) and (21)), the choice of a particular word (see (22) and (23)) and the linear position of a particular word (see (24) and (25)). In sum, one sentence could differ from the other sentence within a pair in that it had one more word (see (21)) or one different word (see (23)) or one word in a different position (see (25)).

(20) Kdo   se   vyspal   s   Martinem?
      who.NOM REFL.CL sleep.PST with Martin.INS
      Who slept with Martin?

(21) Kdo   že se   vyspal   s   Martinem?
      who.NOM ŽE REFL.CL sleep.PST with Martin.INS
      Who slept with Martin?

(22) Jaké   je   v Thajsku   ubytování?
      what-kind-of be.PRS in Thailand.LOC accommodation.NOM
      What kind of accommodation there is in Thailand?

(23) Jaké   je   v Thajsku   jídlo?
      what-kind-of be.PRS in Thailand.LOC food.NOM
      What kind of food there is in Thailand?

(24) Kam   odletěl míč?
      where.LOC fly.PST ball.NOM
      Where did the ball fly?

(25) Kam   míč   odletěl?
      Where.LOC ball.NOM fly.PST
      Where did the ball fly?

Each test item featured a pair of sentences that presented a contrast between matrix *wh*-questions with and without *že* (see (20) and (21) above). The moved *wh*-expression was either a subject or an object or an adjunct. Similarly, each filler featured a pair of sentences. However, none of the fillers comprised a matrix *wh*-question with *že* (see (24) and (25) above). The moved *wh*-expression was either an object or an adjunct. The minimal contrast was achieved either by inserting a word or by moving a word. Four of the sentences, each in a different filler, were ungrammatical. The ungrammaticality always stemmed from a severe word order violation. The purpose of these sentences was to check whether the informants

were paying attention in the course of the test. Finally, the practice items, which were not part of the test *per se*, allowed the informants to get to grips with the answering mechanism. Again, each practice item featured a pair of sentences. One of the sentences included an error that was based on the incompatibility of the answer with the given story, rather than on ungrammaticality. This was meant to make the informants aware of the importance of understanding the stories, as opposed to solely focusing on the syntax of the sentences.

The stories took the form of text. In the test, the stories remained visible for the duration of answering the related questions. In contrast to stories, the sentences took the form of audio recordings. The recordings were by either a male or a female speaker. They were made in the UCL's phonetic laboratory in Chandler House, 2 Wakefield Street, London, WC1N 1PF. The audio recordings of *wh*-questions with *že* were adjusted so that the pitch contours were as identical to those of the corresponding *wh*-questions without *že* as possible. The adjustments were made in an informed way, using quantitative target approximation (see Prom-on et al. (2009)). A sample of the adjusted and the unadjusted audio recordings was presented to a group of native speakers to check whether they could distinguish the former from the latter. Each native speaker was presented with the following question: 'Which of the following two recordings do you consider to be more natural?' Native speakers did not show a bias for either the adjusted or the unadjusted audio recordings. Consequently, the adjusted audio recordings were used in the experiment to minimise the possible influence of the pitch on the informants' preference judgments. The audio recordings were identical for both versions of a given story associated with test items.

6.3.2   Design

The experiment assumed a one-way design with two levels of independent variable. The phenomenon studied was the sensitivity of *že* in matrix *wh*-questions to the level of givenness. The closest one can get to matrix *wh*-questions with *že* are matrix *wh*-questions without *že*. Therefore, each test item comprised two sentences that formed a minimal pair consisting of a particular *wh*-question with *že* and a corresponding *wh*-question without *že*. To find out about the particle's sensitivity to the level of givenness, the 12 stories associated with the test items were a product of creating two versions of six distinct stories. The two versions did not force the request-for-repetition reading, and were mere requests for new information. The first version involved the speaker implying that such and such was the case, without explicitly revealing any more relevant details. The second version did not involve any explicit or implicit mention of anything by the speaker, but only a

thought on part of the listener, who presumed that such and such was the case. The labels given to the two conditions were Implied and New, respectively. Each condition formed one level of the single factor of the one-way design.

(26)   An overview of the one-way design with two levels of the independent variable.

Implied	New
že	*že*
∅	∅

Every informant judged each of the two conditions (i.e., Implied and New) thrice, and saw exactly one version of each of the six stories. In addition, every informant saw all twelve fillers and both practice items along with their associated stories. In every script, the order of test items, fillers and practice items was the following.

(27)   An overview of the distribution of test items, fillers and practice items
within the script.

SCRIPT	
Item Number	Item Type
1	Practice
2	Practice
3	Filler
4	Test
5	Filler (error)
6	Filler
7	Test
8	Filler
9	Filler (error)
10	Test
11	Filler
12	Filler
13	Test
14	Filler (error)
15	Filler
16	Test
17	Filler
18	Filler (error)
19	Test
20	Filler

Using the above template, two scripts were created. To ensure that both scripts
comprised exactly one version of each story and three exemplars of each experi-
mental condition, the selection method depicted in (28) below was used.

(28)   Experiment II - selection method.

	Version 1 (Implied)	Version 2 (New)
Story 1	Script 1	Script 2
Story 2	Script 2	Script 1
Story 3	Script 1	Script 2
Story 4	Script 2	Script 1
Story 5	Script 1	Script 2
Story 6	Script 2	Script 1

To introduce further variation, two versions of each script were created: one half of the audio recordings of test items, fillers and practice items was by a male speaker and the other half by a female speaker. The order of test items, fillers and practice items was fixed. However, all experimental items were randomly inserted into the fixed, prefabricated positions.

Instructions for informants appeared at the beginning of the questionnaire. The end was reserved for demographic and other information. More precisely, informants were asked to specify their native language, dialect, age, gender, highest achieved education, occupation and name. The last of these was required only to enable the informants' identification in potential follow-up experimental studies.

6.3.3   Methodology

The elicitation method used in the experiment was Forced Choice. The informants were asked to select one of the two sentences that were offered as answers to the following question: 'Which of the two sentences, according to you, sounds better in a given context?' The informants were not asked to select the sentence that they thought more grammatical. In the present study, each informant judged both experimental conditions thrice by selecting one of the two sentences each time.

The informants (n=70) were a mixture of students and professionals recruited via email. For convenience, the questionnaire was made accessible on the Internet. The experimental tool used was tailored software developed by Honeypot s.r.o. (www.honeypot.cz). The data gathered during the experiment were imported into an Excel file.

6.3.4   Data filtering and cleaning

As opposed to some other experimental tools (e.g., WebExp), the software used in the present study was not capable of automated filtering or cleaning of any kind.

Therefore, the collected data had to be filtered and cleaned manually.

As for filtering, six criteria were set. Their aim was to prepare the data for statistical analyses by excluding informants who had violated at least one of them. First, those who did not select the 'native speaker of Czech' option were excluded. This included informants who selected the 'native speaker of other language' option as well as those who did not provide any specification of their mother tongue. Second, those who completed the questionnaire in four minutes or less were filtered out. The limit of four minutes allowed 12 seconds for each of the 20 questions. The limit might have been too liberal; however, it sufficed to filter out those informants who were most likely not to be paying enough attention to the questionnaire. Third, those who wrongly answered any of the fillers containing an error were excluded. Wrongly answering any of the fillers containing an error was taken to indicate a loss of concentration. Fourth, those who did not specify all of the variables (see (29) below) were filtered out. A statistical analysis with missing data would have been less accurate than one without any missing data. Fifth, those who were familiar with my work were excluded. Sixth, those who did not play each audio recording at least once were excluded. The chief aim of the herein-listed filtering criteria was to minimise any possible noise in the results. The data left after sifting are used in the reminder of this chapter.

However, it was not enough to filter out the data provided by informants who violated at least one of the six criteria mentioned above. It was also necessary to clean the remaining data before any statistics could have been applied. This was achieved in three steps. First, all fillers and practice items were deleted. This ensured that the data consisted only of test items. Second, all categories that were not selected to be variables were deleted. This ensured that only those categories that were of statistical interest were considered. Third, the values specified for any of the variables were encoded. The statistical program that was used to analyse the data (i.e., SPSS) required the values to be expressed in numerical form. The cleaned data were used to estimate means with confidence intervals.

6.3.5 Statistical method

Given that the data collected during the present experiment were of the same nature as those of Experiment I, all the remarks from Section 6.2.5 also apply here.

6.3.6 Variables

The statistics reported below consider not only the significance of the relation of the independent variable (i.e., condition) and the overall results, but also the significance of the relation of each of the following variables and the overall results.

(29)   An overview of the covariates selected for Experiment II.

    a.   Dialect (categories: Prague; other; do not know)

    b.   Education (categories: below university; university)

    c.   Gender (categories: male; female)

    d.   Language (categories: monolingual; bi/multilingual)

In addition, to check whether the gender of the speaker of the audio recordings influenced the outcome, the covariate Voice Gender was included.

6.3.7   Hypotheses

The following were the hypotheses to be tested by the present experimental study.

(30)   Experimental Hypothesis (H_1)
    Participants would have the following preference with respect to the use of *wh*-questions with *že*: Implied>New.

(31)   Null Hypothesis (H_0)
    Participants would have no preference for either *wh*-questions with *že* or *wh*-questions without *že* in either Implied or New.

6.3.8   Results

There was a highly significant difference ($p<.001$) between Implied and New (Implied>New).

(32)   Experiment II - estimated means with confidence intervals.
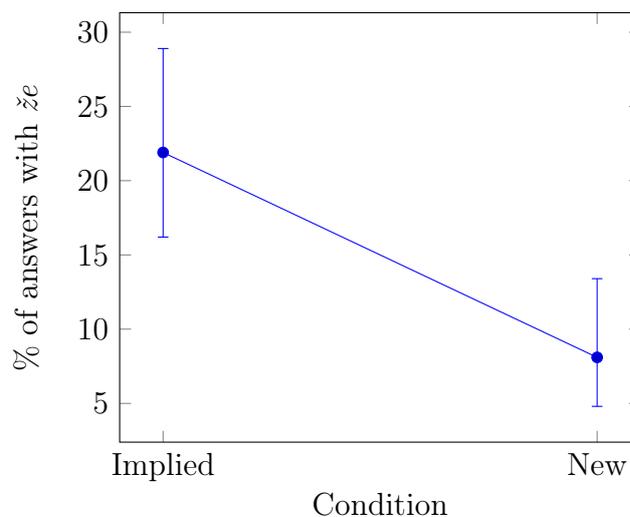
None of the covariates specified in (29) significantly influenced the preference choice of participants, and neither did the covariate Voice Gender. Given the results, H_1

was supported and, by implication, H_0 could be discarded.

Again, the best fitting model was one that included condition as the only fixed variable. Some other models scored better in terms of AIC; however, this was expected since the inclusion of non-significant fixed effects brought the coefficient value closer to the one associated with the intercept-only model.

6.3.9   Discussion

The results of the present study support the assumption that *že* in matrix *wh*-questions is sensitive to the degree of givenness. It remains an open question whether *že* in embedded *wh*-questions has the same property. The results of Experiment I did not show any significant difference between Implicated and Presumed. The differences in estimates for the two conditions in Experiment I and Experiment II might be due to sample size: Experiment II (n=70) had more than 1.5 times as many participants as Experiment I (n=45). In addition to this, each informant judged each condition thrice in Experiment II, but only twice in Experiment I. Expressing this in numbers, each experimental condition was judged 210 times in Experiment II, but only 90 times in Experiment I. Hence, there is a possibility that, were the sample size increased for Experiment I, the difference between the two conditions would show even for embedded *wh*-questions with *že*.

6.4   Conclusion

The above experiments show that the particle *že* in Czech *wh*-questions is sensitive to the level of givenness of the proposition associated with its containing *wh*-question. Whether this means that there are both direct and reported echo questions is debatable. It is debatable because echo questions can be defined in various ways. If echo questions were defined strictly structurally, as *in-situ wh*-questions, then none of the questions that were included in the two experiments could be considered echo questions. This follows from the fact that they all involve *wh*-movement. When it comes to *in-situ wh*-questions, Czech patterns with English: the *in-situ* counterparts of the *wh*-questions used in test items are possible to construct for Experiment II, but not for Experiment I. One obvious problem is that not all *in-situ wh*-questions are echo questions. So, a strictly structural definition seems to be inadequate. If, however, echo questions were defined semantically/pragmatically, as requests for repetition, then the results of Experiemnt I and II could be taken to prove the existence of reported echo questions. Perhaps, the semantics of the particle *že*, which is in accord with the licensing conditions on the use of echo questions ('echo' under the strict structural definititon), makes it possible to report

an echo question.

In addition, the results of the two experimental studies reported above inform our understanding of how fine-grained information structure is, or should be. While the results of Experiment I are compatible with the 'standard' bi-categorial classification system (i.e., [given - new]), the results of Experiment II are compatible with a classification system that makes either sub-categorial distinctions (i.e., [given - new [implied - brand new]]) or further categorial distinctions (i.e., [given - implied - new]). The results thus suggest that the representation of discourse with respect to the speaker's knowledge needs to be more granular than usually assumed in the realm of information structure.

Chapter 7

Conclusion

The main aim of the thesis was to compare the uses of the particle *že* in various syntactic constructions. When used as a subordinating conjunction or a question tag, the particle functions as a semantically empty linker that can lexicalise either C or I, depending on the structural make-up of its containing clause. This type of *že* can link a head of an extended projection to a dependent extended projection selected by that head. When used as a discourse particle, *že* functions as a semantically contenful linker that can lexicalise C or appear in SpecCP. This type of *že* can link an utterance to an expression in some preceding discourse.

Based on these observations, one might wonder whether all three uses could be unified. Philip (2013) argues that linkers mark a relation that exists independently between two elements. In some languages, linkers may be omitted, in others they cannot. The (im)possibility of omitting an element is therefore not indicative of its status as a linker. Philip indicates that one of the crucial diagnostics for linkers that function as subordinating conjunctions is their inability to appear in matrix clauses. Of course, matrix clauses are not selected, so they themselves cannot contain a linker indicating their subordination to a head of some higher extended projection. Following this line of reasoning, one could argue, as I did, that *že* used as either a subordinating conjunction or a question tag is a linker. However, when used as a discourse particle, *že* does not mark this sort of relation. Instead, it marks a relation between its containing utterance and an expression in some preceding discourse. This explains why it can be used in both matrix and embedded clauses. On this analysis, *že* would mark an independently existing relation even when used as a discourse particle. This would allow for a unified analysis of all three uses of the particle outlined above.

Admittedly, what would be more problematic would be to explain the syntactic distribution of *že*. Even when functioning as a subordinating conjunction, the particle does not have to always appear adjacent to the phrase headed by the embedding predicate. This was shown to hold for non-initial conjuncts of embedded

coordinated clauses. Perhaps, the adjacency requirement is relaxed in clauses that are not selected. Furthermore, *že* in *wh*-questions cannot be sentence initial. It is not immediately clear what it should be adjacent to. It is possible that the particle has to be adjacent to the projection that contains the expression that the rest of its containing utterance is linked to. This projection would correspond to the fronted phrase. The cases where *že* appears inside the fronted phrase could be captured in the same way. However, it would be difficult to explain why it can appear only inside fronted phrases comprising a *wh*-element, and not inside fronted phrases comprising a full NP.

I conclude that *že* (almost) always functions as a semantically empty linker. The lack of complexity means that it is compatible with the existing hypotheses about syncretism in distributed morphology and nano syntax. The principal distinction between distributed morphology and nano syntax lies in their respective assumptions about lexical entries. While distributed morphology assumes that they are minimally specified, nano syntax assumes that they are maximally specified. The descriptive equivalence achieved by these theories often arises from the very distinct ways in which lexical items compete for lexical entries. In distributed morphology, the (morpho)syntactic and semantic properties that the different uses of *že* share in common would be encoded in the lexical entry. Everything else would be encoded in syntax. In nano syntax, the (morpho)syntactic and semantic properties that the different uses of *že* share in common would be a subset of what is encoded in the lexical entry. Everything else would be also encoded in the lexical entry. As far as syncretism is concerned, the difference between the two frameworks boils down to their respective assumptions about the relation between the properties of the various uses of *že*. Nano syntax, as opposed to distributed morphology, would require the properties of each use to stand in either a subset or a superset relation to the properties of the other uses. However, if the various uses of *že* have the same (or minimally different) properties, it is impossible to show which of the two theories makes better predictions regarding syncretism.

Bibliography

Bašić, M. (2004), Nominal subextractions and the structure of nps in serbian and english, Master's thesis, Universitetet i Tromsø.

Büring, D. (1995), The 59th street bridge accent: on the meaning of topic and focus, PhD thesis, Universität Tübingen.

Büring, D. (2003), 'On d-trees, beans, and b-accents', *Linguistics and Philosophy* **26**(5), 511–545.

Büring, D. (2007), Semantics, intonation, and information structure, *in* G. Ramchand & C. Reiss, eds, 'The Oxford Handbook of Linguistic Interfaces', Oxford University Press, Oxford, chapter 14, pp. 445–473.

Büring, D. (to appear), (contrastive) topic, *in* C. Féry & S. Ishihara, eds, 'Handbook of Information Structure', Oxford University Press.

Cable, S. (2010*a*), 'Against the existence of pied-piping: evidence from tlingit', *Linguistic Inquiry* **41**(4), 563–594.

Cable, S. (2010*b*), *The grammar of Q: Q-particles, wh-movement, and pied-piping*, Oxford Studies in Comparative Syntax, Oxford University Press, Inc., Oxford.

Chomsky, N. & Halle, M. (1968), *The Sound Pattern of English*, Harper and Row, New York.

Citko, B. & Gračanin-Yuksek, M. (2013), 'Towards a new typology of coordinated wh-questions', *Journal of Linguistics* **49**(1), 1–32.

Constant, N. (2012), Topic abstraction as the source for nested alternatives: a conservative semantics for contrastive topic, *in* N. Arnett & R. Bennett, eds, 'Proceedings of the 30th West Coast Conference on Formal Linguistics', Vol. 30, pp. 120–130.

Constant, N. (2014), Contrastive topic: meanings and realizations, PhD thesis, University of Massachusetts Amherst.

Daneš, F. (1957), *Intonace a věta ve spisovné čeština [Sentence-intonation in present-day Standard Czech]*, Československá akademie věd, Praha.

de Haan, G. J. (2001), 'More is going on upstairs than downstairs: Embedded root phenomena in west frisian', *Journal of Comparative Germanic Linguistics* **4**(1), 3–38.

deHaan, G. J. & Weerman, F. (1985), Finiteness and verb fronting in frisian, *in* H. Haider & M. Prinzhorn, eds, 'Verb Second Phenomena in Germanic Languages', Vol. 21 of *Publications in Language Sciences*, Foris Publications, Dordrecht, chapter 4, pp. 77–110.

Dixon, P. (2008), 'Models of accuracy in repeated-measures designs', *Journal of Memory and Language* **59**(4), 447–456.

Dotlačil, J. (2004), The syntax of infinitives in czech, Master's thesis, Universitetet i Tromsø.

Dotlačil, J. (2007), Why clitics cannot climb out of cp: A discourse approach. A draft of a proceedings paper for FASL 15.

Featherston, S. (2004), 'Bridge verbs and v2 verbs - the same thing in spades?', *Zeitschrift für Sprachwissenschaft* **23**(2), 181–209.

Féry, C. & Samek-Lodovici, V. (2006), 'Focus projection and prosodic prominence in nested foci', *Language* **82**(1), 131–150.

Firbas, J. (1966), Non-thematic subjects in contemporary english, *in* J. Vachek & M. Komárek, eds, 'Les problèmes du centre et de la périphérie du système de la langue', Vol. 2, Academia Éditions de l'Académie tchécoslavaque des Sciences, Prague, pp. 239–256.

Götze, M., Weskott, T., Endriss, C., Fiedler, I., Hinterwimmer, S., Petrova, S., Schwarz, A., Skopeteas, S. & Stoel, R. (2007), Information structure in cross-linguistic corpora: Annotation guidelines for phonology, morphology, syntax, semantics, and information structure, *in* S. Dipper, M. Götze & S. Skopeteas, eds, 'Interdisciplinary Studies on Information Structure', Vol. 7, pp. 147–187.

Gruet-Skrabalova, H. (2011), Czech questions with two wh-words, *in* P. Kosta & L. Schürcks, eds, 'Formalization of Grammar in Slavic Languages: Contributions to the Eighth International Conference on Formal Description of Slavic Languages - FDSL VIII 2009', Peter Lang GmbH, Frankfurt am Main, pp. 179–192.

Gruet-Skrabalova, H. (2012*a*), What kind of element is že in czech?, *in* M. Ziková & M. Dočekal, eds, 'Slavic Languages in Formal Grammar: Proceedings of FDSL 8.5, Brno 2010', Peter Lang GmbH, Frankfurt am Main, pp. 33–47.

Gruet-Skrabalova, H. (2012*b*), What kind of element is že in czech?

Gruet-Skrabalova, H. (2015), Verbs and particles in minimal answers to *yes-no* questions, *in* G. Zybatow, P. Biskup, M. Guhl, C. Hurtig, O. Mueller-Reichau & M. Yastrebova, eds, 'Slavic Grammar from a Formal Perspective: The 10th Anniversary FDSL Conference, Leipzig 2013', Vol. 35 of *Linguistik International*, Peter Lang GmbH, Frankfurt am Main, pp. 197–215.

Hagstrom, P. & McCoy, S. (2003), Presuppositions, wh-questions, and discourse particles: Russian že, *in* W. Browne, J.-Y. Kim, B. H. Partee & R. A. Rothstein,

eds, 'Formal Approaches to Slavic Linguistics: The Amherst Meeting', Vol. 11, Michigan Slavic Publications, Ann Arbor, Michigan, pp. 201–218.

Hamblin, C. L. (1958), 'Questions', *Australasian Journal of Philosophy* **36**(3), 159–168.

Heck, F. (2008), *On pied-piping: wh-movement and beyond*, Vol. 98 of *Studies in Generative Grammar*, Mouton de Gruyter, Berlin.

Heck, F. (2009), 'On certain properties of pied-piping', *Linguistic Inquiry* **40**(1), 75–111.

Iatridou, S. & Kroch, A. (1992), The licensing of cp-recursion and its relevance to the germanic verb-second phenomenon, *in* 'Working Papers in Scandinavian Syntax', Vol. 50, pp. 1–24.

Jackendoff, R. S. (1972), *Semantic Interpretation in Generative Grammar*, Studies in Linguistics, The MIT Press, Cambridge, Massachusetts.

Jaeger, T. F. (2008), 'Categorical data analysis: Away from anovas (transformation or not) and towards logit mixed models', *Journal of Memory and Language* **59**(4), 434–446.

Junghanns, U. (2001), On rightward backgrounding, *in* G. Zybatow, U. Junghanns, G. Mehlhorn & L. Szucsich, eds, 'Current Issues in Formal Slavic Linguistics', Linguistik International, Peter Lang GmbH, Frankfurt am Main, pp. 329–343.

Karttunen, L. (1977), 'Syntax and semantics of questions', *Linguistics and Philosophy* **1**(1), 3–44.

Kortlandt, F. (1983), 'From proto-indo-european to slavic', online.
URL: *http://www.kortlandt.nl/publications/art066e.pdf*

Krifka, M. (2001), 'Quantifying into question acts', *Natural Language Semantics* **9**(1), 1–40.

Kučerová, I. (2007), The Syntax of Givenness, PhD thesis, Massachusetts Institute of Technology.

Kuno, S. (1982), The focus of the question and the focus of the answer, *in* R. Schneider, K. Tuite & R. Chametzky, eds, 'CLS: Papers from the parasession on nondeclaratives', Chicago Linguistics Society, Chicago, IL, pp. 134–157.

Ladd, R. D. (1981), A first look at the semantics and pragmatics of negative questions and tag questions, *in* 'Proceedings of Chicago Linguistic Society', Vol. 17, pp. 164–171.

Lenertová, D. (2001), On clitic placement, topicalization and cp-structure in czech, *in* G. Zybatow, U. Junghanns, G. Mehlhorn & L. Szucsich, eds, 'Current Issues in Formal Slavic Linguistics', Linguistik International, Peter Lang GmbH, Frankfurt am Main, pp. 294–305.

Lenertová, D. & Junghanns, U. (2007), Wide focus interpretation with fronted

focus exponents in czech, *in* K. Schwabe & S. Winkler, eds, 'On Information Structure, Meaning and Form: Generalizations across Languages', Vol. 100 of *Linguistik Aktuell/Linguistics Today*, John Benjamins Publishing Company, Amsterdam/Philadelphia, pp. 347–363.

Mathesius, V. (1947), *Čeština a Obecný Jazykozpyt: Soubor Statí*, Melantrich, Praha.

McCloskey, J. (2006), Questions and questioning in a local english, *in* R. Zanuttini, H. Campos, E. Herburger & P. Portner, eds, 'Crosslinguistic Research in Syntax and Semantics: Negation, Tense, and Clausal Architecture', Georgetown University Press, Washington, D.C., chapter 4, pp. 87–126.

Merchant, J. (2001), *The syntax of silence: sluicing, islands, and the theory of ellipsis*, Oxford Studies in Theoretical Linguistics, Oxford University Press, Oxford.

Meyer, R. (2006), O pozici podřadících spojek v čnk, *in* F. Štícha & J. Šímandl, eds, 'Gramatika a korpus', ÚJČ AV ČR.

Meyer, R. (2010), Mood in czech and slovak, *in* B. Rothstein & R. Thieroff, eds, 'Mood in the Languages of Europe', Vol. 120 of *Studies in Language Companion Series*, John Benjamins B.V., Amsterdam/Philadelphia, pp. 358–375.

Munn, A. B. (1993), Topics in the syntax and semantics of coordinate structures, PhD thesis, University of Maryland.

Neeleman, A. & van de Koot, H. (2006), Syntactic haplology, *in* M. Everaert & H. van Riemsdijk, eds, 'The Blackwell Companion to Syntax', Vol. IV of *Blackwell Handbooks in Linguistics*, Blackwell Publishing Ltd, chapter 69, pp. 685–710.

Neeleman, A. & van de Koot, H. (2008), 'Dutch scrambling and the nature of discourse templates', *The Journal of Comparative Germanic Linguistics* **11**(2), 137–189.

Nevins, A. (2010), Haplological dissimilation at distinct stages of exponence.

Nissim, M., Dingare, S., Carletta, J. & Steedman, M. (2004), An annotation scheme for information status in dialogue, *in* M. T. Lino, M. F. Xavier, F. Ferreira, R. Costa & R. Silva, eds, 'Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC 2004)', Lisbon, pp. 1023–1026.

Philip, J. (2013), '(dis)harmony, the head-proximate filter, and linkers', *Journal of Linguistics* **49**(1), 165–213.

Prince, E. F. (1981), Toward a taxonomy of given-new information, *in* P. Cole, ed., 'Radical Pragmatics', Academic Press, Inc., New York, pp. 223–255.

Prom-on, S., Xu, Y. & Thipakorn, B. (2009), 'Modeling tone and intonation in mandarin and english as a process of target approximation', *The Journal of the Acoustical Society of America* **125**(1), 405–424.

Proto-Indo-European Etymological Dictionary (2009), online.
  URL: *http://dnghu.org/en/proto-indo-european-language/*

Quené, H. & van den Bergh, H. (2008), 'Examples of mixed-effects modeling with crossed random effects and with binomial data', *Journal of Memory and Language* **59**(4), 413–425.

Riester, A. (2008), A semantic explication of information status and the underspecification of the recipients' knowledge, *in* A. Grønn, ed., 'Proceedings of SuB12'.

Roberts, C. (2012), 'Information structure in discourse: towards an integrated formal theory of pragmatics', *Semantics and Pragmatics* **5**(6), 1–69.

Rooth, M. (1992), 'A theory of focus interpretation', *Natural Language Semantics* **1**(1), 75–116.

Rudin, C. (1988), 'On multiple questions and multiple wh fronting', *Natural Language and Linguistic Theory* **6**(4), 445–501.

Schwarzschild, R. (1999), 'Givenness, avoidf and other constraints on the placement of accent', *Natural Language Semantics* **7**(2), 141–177.

Šimík, R. & Wierzba, M. (2015), 'The role of givenness, presupposition, and prosody in czech word order: An experimental study', *Semantics and Pragmatics* **8**, 1–103.

Sturgeon, A. (2008), *The left periphery: the interaction of syntax, pragmatics and prosody in Czech*, Vol. 129 of *Linguistik Aktuell/Linguistics Today (LA)*, John Benjamins Publishing Company, Amsterdam.

Truckenbrodt, H. (2013), 'An analysis of prosodic f-effects in interrogatives: Prosody, syntax and semantics', *Lingua* **124**, 131–175.

Urešová, M. (2008), Anglické tázací dovětky. A seminar paper on tag questions in Czech and English.

Veselá, K., Peterek, N. & Hajičová, E. (2003), 'Topic-focus articulation in pdt: Prosodic characteristics of contrastive topic', *The Prague Bulletin of Mathematical Linguistics* **79–80**.

Veselovská, L. (1993), Wh movement and multiple questions in czech, Master's thesis, Durham University, Durham.

Veselovská, L. (1995), Phrasal movement and X-morphology: word order parallels in Czech and English nominal and verbal projections, PhD thesis, Palacký University, Olomouc.

Wagner, M. (2012), 'Contrastive topics decomposed', *Semantics and Pragmatics* **5**(8), 1–54.

Wurmbrand, S. (2015), 'Cross-linguistic investigation of restructuring and finiteness:

implications for clausal domains and locality', Eastern Generative Grammar (EGG) Summer School.