The question of form in the forming of questions: The meaning and use of clefted *wh*-interrogatives in Swedish[1]

JOHAN BRANDTLER

Stockholm University

This paper addresses the meaning and use of clefted *wh*-interrogatives (I-clefts) in Swedish. It is shown that I-clefts always relate immediately to the topic under discussion and serve to clarify a matter in relation to this topic. They are never used in out-of-the-blue contexts. I argue that I-clefts have the same information structure as typically assumed for declarative clefts: the clefted clause expresses an existential presupposition and the cleft phrase is the identificational focus of the utterance. I further argue that the implication of existence commonly associated with *canonical* argument questions is weaker (a conversational implicature) than the existential presupposition associated with clefts. The results from an extensive corpus survey show that argument I-clefts (*who*, *what*) constitute no less than 98% of the total number of I-clefts in my material. This frequency is linked to the presuppositional status of the cleft construction: in contexts where the denoted event is presupposed as part of the common ground, the clefted variety is the more effective choice, due to its clear partitioning of focus and ground. The 'cost' of using a more complex syntactic structure (the cleft) is thus counterbalanced by the benefit of being able to pose a question adjusted to the contextual requirements. As non-argument questions are typically presuppositional irrespective of syntactic form, the gain of using a cleft is less obvious – hence their infrequency in the material.

1. INTRODUCTION

This paper is concerned with the meaning and use of clefted *wh*-interrogatives in Swedish. The relevant contrast with canonical *wh*-interrogatives is illustrated in (1) below.

(1) (a) Vem åt du lunch med igår?   Canonical
        who ate you lunch with yesterday
        'Who did you lunch with yesterday?'

(b) Vem var det (som) du åt lunch med igår?   Cleft
    who was it that you ate lunch with yesterday
    'Who was it that you lunched with yesterday?'

Though often overlooked in the literature, clefted interrogatives are natural and 'very common' in Swedish (Teleman, Hellberg & Andersson 1999: 4: 747), especially in spoken and informal written language. Similar claims have been made for Norwegian (Faarlund, Lie & Vannebo 1997: 1091 and Svenonius 1998: 164). In this respect, Swedish and Norwegian (and most probably the remaining Scandinavian languages) differ from other Germanic languages: in Dutch, clefted interrogatives are marked and unnatural; in German, clefted interrogatives are infrequent and contextually marked (Molnár 2016). Though clefted interrogatives do occur naturally in English, they are less frequent than in Norwegian, at least (Gundel 2002). For a more in-depth comparison of the structure and use of (declarative and interrogative) clefts within the Scandinavian languages, see Søfteland (2014: Chapter 3).

The purpose of this paper is twofold: (i) to present a unified semantic analysis of clefted interrogatives in Swedish that encompasses the different *wh*-phrases, and (ii) to account for the discourse potential of clefted *wh*-interrogatives by reference to the syntax and semantics of different *wh*-question types. The study is based on an empirical investigation of the distributional pattern of *wh*-interrogatives in the web-based discussion forum *Flashback Forum*. The results of the quantitative study are complemented by a qualitative analysis of the discourse function of a sample set of more than 5.000 clefted *wh*-interrogatives from the material.

The study begins in Section 3, after a brief presentation of the structure of Swedish clefts in Section 2. The quantitative investigation shows that clefted *wh*-interrogatives account for approx. 4% of the almost 5 million main clause *wh*-interrogatives in the Flashback Forum. The survey further shows that the likelihood of a particular *wh*-interrogative taking the form of a cleft is linked to the *wh*-phrase rather than to the absolute or relative frequency of its occurrence. While interrogatives introduced by *vem* 'who' and *vad* 'what' constitute 58% of the total number of *wh*-interrogatives in the material, they constitute no less than 98% of the total number of clefted *wh*-interrogatives. Also in relative terms, *vem*- and *vad*-interrogatives are the most likely to be formulated as clefts, whereas interrogatives introduced by *hur* 'how' and *var* 'where' strongly resist clefting.

In Section 4, I qualitatively examine how *wh*-interrogatives are used in discourse, by analyzing a random selection of 5,080 clefted *wh*-interrogatives from the same corpus. The interrogatives are roughly divided into two categories: Questions (i.e., interrogatives with the primary function of updating the speaker's knowledge state), and Non-questions (interrogatives with some other communicative purpose). The analysis reveals that clefted *wh*-questions in Swedish primarily function as questions of clarification in contexts where the state-of-affairs expressed in the embedded proposition is presupposed as being part of

756

the situational and/or linguistic context. Clefted *wh*-questions never occur out-of-the-blue, judging from the material. Thus, the clefted question in (1b) above requests referential specification of the addressee's lunch date in a context where the matter of the addressee's lunch is an active topic in the discourse universe. In contrast, the canonical question (1a) is contextually less restricted, and may be used also in situations where the speaker has no preconceived knowledge of whether the addressee even had lunch. I conclude that clefted *wh*-questions in Swedish do *not* have the same general interpretation as canonical *wh*-questions, contra what has previously been claimed for Swedish (Teleman et al. 1999: 4: 517), and in contrast to the situation often cited for the French *wh*-interrogative system (see Tailleur 2013).

In Section 5, I argue that clefted interrogatives have the same information structure as clefted declaratives: the embedded proposition expresses an existential presupposition, and the *wh*-phrase is the identificational focus of the sentence, and thus interpreted exhaustively. I refute the commonly held view that canonical argument *wh*-questions (i.e., *who* and *what*) give rise to existential presuppositions, at least in Swedish. Instead, I argue that the existential implication associated with canonical argument *wh*-questions is better characterized as a generalized conversational implicature, emanating from the choice on the part of the speaker to use a *wh*-question over a yes/no question.

In Section 6 I propose a direct link between the presuppositional status of clefted argument *wh*-interrogatives and their frequency in the material. In contexts where the event denoted by the question is known to both discourse participants, the clefted variety is the optimal choice, due to its clear partitioning of focus and ground. The 'cost' of using a more complex syntactic structure (the cleft) is thus counterbalanced by the benefit of being able to pose an informationally more specific question in the context. As non-argument questions (e.g., *when*, *where*, *how*) are typically presuppositional irrespective of syntactic form, the gain of using a cleft is less obvious – hence their infrequency in the material.

2. CLEFT CONSTRUCTIONS IN SWEDISH

Starting with Prince (1978), clefts have received a fair share of attention in the literature. Needless to say, this short section cannot do justice to this body of literature, and will therefore be limited to the basic structural facts concerning clefts in Swedish. For the interested reader, Reeve (2011) gives an excellent critical overview of different syntactic analyses of (English) clefts; a thorough presentation of previous approaches to clefts in the Scandinavian languages is given in Søfteland (2014: Chapter 3). See also Huber (2002) for a contrastive study on the grammar and information structure of clefts in Swedish and German, and Johansson (2002) for a contrastive study on the use of clefts in Swedish and English.

2.1 *Declarative clefts*

For the current purposes, I assume Lambrecht's (2001: 467) often cited – and relatively theory-neutral – biclausal definition of cleft constructions, as given below.

(2)   A CLEFT CONSTRUCTION (CC) is a complex sentence structure consisting of a matrix clause headed by a copula and a relative or relative-like clause whose relativized argument is coindexed with the predicative argument of the copula. Taken together, the matrix and the relative express a logically simple proposition, which can also be expressed in the form of a single clause without a change in truth conditions.

Declarative clefts in Swedish have a structure similar to that of English *it*-clefts: the matrix is typically introduced by the expletive pronoun *det* 'it' followed by the copula *vara* 'be' in the present or past tense.[2] A schematic overview of the structure of the Swedish declarative cleft construction is given in (2) below.

(3)   The form of Swedish declarative cleft constructions:
det + COPULA (*är/var*) + CLEFT PHRASE + CLEFTED CLAUSE

(4)   (a)   Det är avdelningen   som betalar för middagen.
              it   is department.DEF who pays   for dinner.DEF
              'It's the department that pays for the dinner.'

       (b)   Det var nästa vecka (som) vi skulle träffas.
              it   was next week   that we should meet
              'It was next week we were supposed to meet.'

As seen in (3), I will use *cleft phrase* as an umbrella term for the syntactic constituent following the copula, thus avoiding any discussion regarding its syntactic function within the matrix and the subordinate clause. Similarly, I will use the term *clefted clause* for the relative-like clause following the cleft phrase.

Though the commonality between *it*-clefts and relative clauses is often emphasized in traditional biclausal analyses of cleft constructions (see, e.g., Chomsky 1977; Delahunty 1981, 1984; Hedberg 1990; Clech-Darbon, Rebuschi & Rialland 1999; Reeve 2000, 2011; Haegeman, Meinunger & Vercauteren 2014), already Jespersen (1927) pointed out that the relative clause in cleft constructions is not used to restrict the reference of the constituent following the copula. He bases this claim on the observation that the relative clause may follow 'a word which is in itself so definite that it cannot be further restricted: it was the

---

[2] Whereas most analyses of clefts treat the pronoun as an expletive (see, e.g., Chomsky 1977; Halvorsen 1978; Delahunty 1981; É. Kiss 1998, 1999; Meinunger 1998; Svenonius 1998), Reeve (2000, 2011) argues in line with Hedberg (2000) that *det* is referential, or at least non-expletive. An argument in favor of Reeve's analysis is the fact that *det* cannot be omitted whenever the cleft phrase is promoted to first position, even in languages where expletive pronouns only occur clause initially (such as Icelandic).

battle of Waterloo that decided the fate of Europe' (Jespersen 1927: 89). Taking Jespersen's perspective, Clech-Darbon et al. (1999) see the relative clause as a headless relative clause (see also den Dikken 2013 for a similar proposal), whereas Reeve (2013) argues that the clefted clause behaves in all relevant ways as if it were a syntactic modifier of the clefted XP. In the main reference work of Swedish grammar, *Svenska Akademiens grammatik* (Teleman et al. 1999), cleft constructions are analyzed as a 'special kind' of relativization, and the constituent following the copula as the *correlate* of the relative clause. Others still have proposed monoclausal analyses of cleft constructions (see e.g., Meinunger 1998 and Frascarelli & Ramaglia 2013), arguing that the cleft copula itself does not head a fully projected clausal domain. See Haegeman et al. (2014) for a critique of monoclausal analyses of cleft constructions.

Similar to its distribution in restrictive relative clauses, the introducing complementizer *som* 'that' is optional in the clefted clause, unless there is a subject gap, in which case *som* is obligatory. Compare the clefts in (5) below, taken from Svenonius (1998: 170, 183).

(5) (a) Det var Jon *(som) träffade mig i staden.
        it was Jon that met me in the.town
        'It was Jon that met me in town.'

    (b) Det var Jon (som) jag träffade i staden.
        it was Jon that I met in the.town

    (c) Det var i staden (som) jag träffade Jon.
        it was in the.town that I met Jon

    (d) Det var rött (som) han målade huset.
        it was red that he painted the.house
        'It was red that he painted the house.'

As discussed at some length in Svenonius (1998), the Scandinavian languages display a quite intricate micro-variation with regards to the distribution and choice of the introducing complementizer. While (standard) Swedish only, and always, allows *som*, Norwegian and Danish only use *som* whenever the cleft phrase is a DP; with other kinds of cleft phrases (such as PPs and APs), *som* is ruled out. Danish may also employ the element *der* 'there' with DP cleft phrases, and both Norwegian and Danish allow *at* 'that' with prepositional cleft phrases. For a full overview and syntactic analysis of this rather intricate pattern, the reader is referred to Svenonius (1998).

2.2 *Interrogative clefts*

There are two types of interrogative clefts: polar interrogatives and *wh*-interrogatives. As Swedish polar interrogatives are introduced by the finite verb, the cleft phrase always occurs to the immediate right of *det*; see (6).

759

(6) (a) Är det kaffe (som) du dricker?
      is it coffee that you drink
      'Is that coffee you're drinking?'

As *wh*-phrases obligatorily front in Swedish, clefted *wh*-interrogatives are obligatorily introduced by the cleft phrase. Due to the V2-restriction of Swedish main clauses, *det* inverts with the copula; see (7).

(7) (a) Vem var det som betalade middagen?
      who was it that paid   dinner.DEF
      'Who paid for the dinner?'

    (b) När var det (som) vi skulle träffas?
      when was it that we should meet
      'When did we decide to meet?'

Throughout this paper, I will refer to clefted *wh*-interrogatives like (7) as I-clefts. To avoid confusion, I will not use the term *wh*-cleft, as this term is sometimes used interchangeably with the term *pseudo-cleft* for declarative clefts introduced by a *wh*-phrase. Compare the different structures in (8) below.

(8) (a) Vad är det du helst dricker till maten?   I-cleft
      what is it you rather drink to food
      'What kind of drink do you most prefer with food?'

    (b) Vad jag verkligen föredrar är champagne.   pseudo-/*wh*-cleft
      what I really prefer is champagne
      'What I really prefer is champagne.'

Importantly, the correlation between the cleft phrase and the *wh*-phrase is not absolute, as the cleft phrase can be 'split' in two. While the *wh*-phrase obligatorily fronts, an attribute or specification may remain lower in the matrix. The split cleft phrase is boldfaced in (9) below.

(9) (a) Av ren nyfikenhet, **vad** var det **för film** som "räddade
      of sheer curiosity what was it for movie that saved
      livet" på dig?
      the.life on you
      'Out of sheer curiosity, what/which movie saved your life?'

    (b) **Vem** var det **av dina vänner** som hade flyttat till Kongo?
      who was it of your friends that had moved to Congo
      'Who was it of your friends that had moved to the Congo?'

    (c) **När** var det **nästa vecka** (som) du kunde träffas?
      when was it next week that you could meet
      'When were you able to meet next week?'

Following Teleman et al. (1999), I have analyzed such instances as *complex wh-phrases*. Hence, I assume that the cleft phrase in (9a) is the phrase *Vad för film*,

in (9b) *Vem av dina vänner*, and in (9c) *När nästa vecka*. An argument in favor of this analysis is that the whole phrase may be fronted, even though the split structures in (9) are probably the preferred alternatives, especially for *vad*-clefts.

(10)   (a)   **Vad för** (en) **film**   var det som "räddade livet" på dig?
             what for a   movie was it that   saved   the.life on you
             'What/which movie was it that saved your life?'

        (b)   **Vem av dina vänner** var det som hade flyttat till Kongo?
             who of your friends was it   that had moved to Congo
             'Who was it of your friends that had moved to the Congo?'

        (c)   **När nästa vecka** var det som du   kunde träffas?
             when next week was it   that you could meet
             'When were you able to meet next week?'

In clefted *varför*-interrogatives, an argument to the embedded main verb may constitute the cleft phrase rather than the *wh*-phrase; see (11) below.

(11)   (a)   Varför var det **bara mig** (som) de   ville   prata med?
             why   was it only me   that they wanted talk to
             'Why was it only me they wanted to talk to?'

        (b)   Varför är det **jag** som får   skulden?
             why   is it   me that gets blame
             'Why do I get all the blame?'

As evident from (11), the *wh*-phrase and the cleft phrase do not constitute one split phrase in the sense illustrated in (9) above. As the primary characteristic of an I-cleft is precisely that the *wh*-phrase equals the cleft phrase, clefted *varför*-interrogatives are not proper I-clefts. Consequently, I have omitted clefted *varför*-interrogatives from this survey.

3. THE FREQUENCY OF I-CLEFTS IN SWEDISH

This section presents the frequency of I-clefts in relation to canonical *wh*-interrogatives in Swedish.

3.1 *Method and material*

In order to chart the frequency of I-clefts in Swedish, almost 5 million interrogative clauses were excerpted from the *Flashback Forum* corpus at the Språkbanken corpora resource (Borin, Forsberg & Roxendal 2012). The corpus consists of tagged posts from the Flashback Forum, a popular, Swedish web-based discussion forum.[3] The most recent updates to the corpus were made in March, 2015. For more details concerning the corpus, see the Appendix.

---

[3] URL: http://www.flashback.orgwww.flashback.org.

There are several reasons the Flashback Forum was chosen. The language in discussion forums consists of 'a mixture of personal and group interactions', which tends to reflect a natural and informal style of speech (Biber & Conrad 2009: 91). Thus, we may assume that the language is representative of modern-day, informal Swedish. Furthermore, discussion forums potentially contain a higher share of questions than more 'monolog'-like text genres, like blogs and newspaper texts. In addition, many posts function as direct comments to something the speaker has just experienced. Thus, in a typical post a minimum of text planning has occurred, meaning that the particular linguistic choices (words, word order and constructions) are less likely to result from deliberate attempts by the writer to achieve literary or rhetorical effects compared to more self-conscious genres, such as fiction and blogs.

In order to extract all *wh*-interrogatives in the corpus (i.e., both canonical and clefted), a number of searches were conducted using the following string, where *pos* is part of speech and *VB* is verb:

(12)   [(word = ''Vem'' | word = ''vem'')] [pos = ''VB''].

In sum, a total of 6 different *wh*-phrases were selected for this survey: *vem* 'who', *vad* 'what', *när* 'when', *var* 'locational where', *vart* 'directional where', and *hur* 'how'. The first part of the string in (12) was consequently substituted for each relevant *wh*-phrase. Interrogatives introduced by *varför* 'why' were excluded from the survey, for reasons stated above: in clefted *varför*-interrogatives, the *wh*-phrase neither equals nor is part of the cleft phrase.

As the syntactic structure of a cleft is given (see Section 2.2 above), lexical strings could be used to extract I-clefts from the corpus. To illustrate, the string in (13) extracts all structures of the form *Vem/vem är det* 'who is it' from the corpus, excluding structures where the expletive *det* 'it' is followed by a major delimiter (MAD), i.e., a question mark or a full stop.

(13)   [(word = ''Vem'' | word = ''vem'')] [word = ''är''] [word = ''det'']
        [msd != ''MAD''].

For each *wh*-phrase, I-clefts in both the present and past tense were extracted. By changing the value [word = ''är''] to [word = ''var''], the string extracts all *vem*-clefts in the past tense.

Lexical strings were necessary for the survey, as the machine generated tagging of *Flashback Forum* is not reliable enough to successfully generate searches based on syntactic dependencies. However, the search strings used extracted incorrect (non-cleft) structures of the type illustrated in (14).

(14)   (a)   Vem är det där?
              who is that there
              'Who is that?'

        (b)   Vad är det för något?
              what is that for something?
              'What is that thing?'

762

(c) När   var det då?
   when was it   then
   'So, when was it?'

The entire data sets for *vem*, *när*, *vart* and *var* were manually analyzed for correct cleft structures. Since the data sets for *vad* and *how* were substantially bigger, a manual analysis was carried out on a sample set. An estimated total was then calculated based on the mean number of true cleft structures in the sample; see the appendix for more details concerning this procedure.

While practical, lexical strings obviously restrict the search. First, D-linked *wh*-phrases were deliberately excluded from the survey, since structures like *vilken/vilka X* 'which X' or *hur många X* 'how many X' require a less specified search string. Furthermore, D-linked *wh*-phrases are presuppositional in a sense non-D-linked *wh*-phrases are not, which in turn motivates a separate treatment (see further discussion in Section 7). Second, the lexical search strings used do not allow any intervening material between the *wh*-phrase and the finite verb. Consequently, clause initial complex *wh*-phrases (as in (10) above) are not extracted, though split structures, as in (9), are. For the same reason, structures containing a swearword or some other emphasizing element to the immediate right of the *wh*-phrase, as in (15) below, were not found either.

(15)   (a)   Vem i   helvete är det?
          who in hell   is that
          'Who the hell is that?'

       (b)   Vem i   helvete är det som har sagt det?
          who in hell   is it   that has said that
          'Who the hell has said that?'

As the restriction affects canonical (15a) and clefted (15b) structures alike, this methodological shortcoming probably has little bearing on the overall comparison.[4]

---

[4] As pointed out by an anonymous reviewer, the search string also excludes structures in which the copula is preceded by a modal or temporal auxiliary, as in the declarative and interrogative clefts below.

 (i)   Det borde vara värden   som öppnar vinet.
      it   should be   host.DEF who opens wine.DEF
      'It ought to be the host who opens the wine.'

(ii)   Vem borde det vara som öppnar vinet?
      who should it   be   that opens wine.DEF
      'Who is that ought to open the wine?'

However, interrogative clefts with modal and temporal auxiliaries seem to be very rare in Swedish. A specified search in the same corpus of the structure in (ii) with the auxiliaries *borde* 'should', *kan* 'can' and *bör* 'ought to' generated no hits.

3.2 *The relative frequency of I-clefts*

Let us begin by giving an overview of the material as a whole. In total, the data set contains 4,979,856 main clause *wh*-interrogatives introduced by one of the six relevant *wh*-phrases. Figure 1 illustrates the frequency of each *wh*-interrogative. Please note that both canonical and clefted *wh*-interrogatives are included in the total.
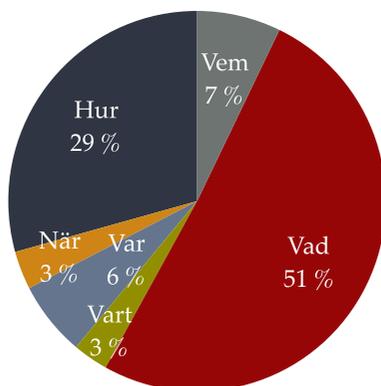

Figure 1
Share of different kinds of *wh*-interrogatives in the material. Total of 4,979,856 interrogative clauses.

As is clearly evident from Figure 1, *wh*-interrogatives introduced by *vad* 'what' are the most frequent in the material (51%), followed by *hur* 'how' (29%). Only 7% of the *wh*-interrogatives are introduced by *vem* 'who', which might be a bit surprising given the prominence of *who*-questions in the linguistic literature. It should also be noted that the frequency hierarchy in Figure 1 corresponds exactly to that of English *wh*-phrases in Siemund (2012: 14).[5]

Singling out I-clefts in the material, these total approx. 4% (202,151 sentences) of the total number of *wh*-interrogatives. As clefted questions in Swedish are assumed to be 'very common' according to Teleman et al. (1999: 4: 747) – and even 'extremely common' according to Svenonius (1998: 164) – 4% is perhaps lower than what one would expect. One should keep in mind, however, that both Teleman et al. and Svenonius refer to spoken language. Although the language

---

[5] As discussed by Siemund (2012: 30), the relative frequency of *wh*-phrases does not correlate in any obvious way with hierarchies based on formal and/or functional complexity: 'nominal interrogative words encoding entities and individuals are found at the top (*what*) and in the middle (*who*) of this hierarchy. Adverbial interrogatives are located to the left and right of *who*.' This leads Siemund (2012: 30) to conclude that the frequency hierarchy 'must be the result of different parameters'.

in the *Flashback Forum* tends to be informal, a corpus of spoken language may reveal a higher frequency ratio.[6]

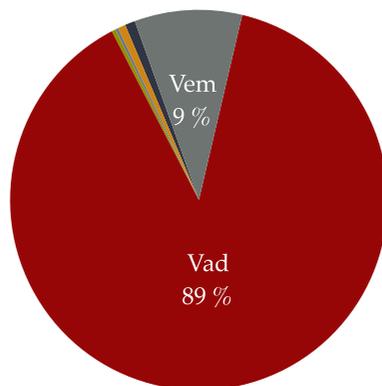As illustrated in Figure 2, the number of I-clefts is not equally distributed across the different *wh*-phrases.

Figure 2
Share of different kinds of I-clefts in the material. Total of 202,151 I-clefts.

We see from Figure 2 that the overwhelming majority (89%) of I-clefts are introduced by *vad* 'what'. Given that *vad*-interrogatives are the most frequent in the material, this finding would not be surprising had the frequency ratio been more in line with the overall frequency of *vad*-interrogatives (51%). The fact that almost 9 in 10 clefts in the material are introduced by *vad* cannot be explained by absolute frequency alone. Note also that I-clefts introduced by *vem* 'who' are the second most frequent (9%), despite the fact that *vem*-interrogatives on the whole are less frequent in the material than *hur*-interrogatives (see Figure 1). In other words, I-clefts introduced by any other *wh*-phrase than *vad* or *vem* are extremely rare: the remaining four *wh*-phrases (*hur*, *när*, *var*, *vart*) taken together make up a mere 2% of the total number of I-clefts.

Given the varying frequency of the different *wh*-interrogatives in the material, perhaps a more relevant aspect is the relative frequency of I-clefts for each respective *wh*-phrase. That is, how likely is a particular *wh*-phrase to introduce an I-cleft in relation to a canonical interrogative? Table 1 illustrates this correlation.[7] We see from Table 1 that the number of I-clefts occurring per 100 *wh*-interrogatives differs considerably depending on the *wh*-phrase. Argument interrogatives, i.e., *vem* 'who' and *vad* 'what', are the most likely to be formulated

---

[6] Unfortunately, the spoken language corpora at the Språkbanken corpora resource are relatively small. The spoken language corpus *Talbanken* yielded no hits for the string *Vem var det*.

[7] Recall that the total number of I-clefts introduced by *vad* and *hur* as given in Table 1 are estimates, based on the mean share of I-clefts occurring in a sample set of the material. See Table A3 in the appendix for more detailed information regarding this estimate.

765

Wh-phrase	Total	I-cleft	I-clefts/ 100 interrog.
Vad 'what'	2,533,080	179,267	7,1
Vem 'who'	357,572	18,595	5,2
När 'when'	162,061	1,356	0,84
Vart 'where'	150,383	713	0,47
Var 'where'	317,100	452	0,14
Hur 'how'	1,459,660	1,768	0,12
Total:	4,979,856	202,151	4,05

Table 1
The relative frequency of I-clefts for each respective *wh*-phrase.

as clefts, whereas *hur-* and *var*-interrogatives very rarely take the form of a cleft (0,12% and 0,14%, respectively). Of the non-argument interrogatives, *när*-interrogatives occur most frequently as clefts, although almost seven times less frequently than *vem*-interrogatives.

Combining the results of Table 1 with Figures 1 and 2, we are able to draw a few conclusions regarding the distribution of I-clefts in Swedish. First, there is no apparent correlation between the absolute frequency of a *wh*-interrogative and the absolute frequency of I-clefts in the material. For example, *vem*-clefts (9%) are more than ten times as frequent as *hur*-clefts (1%), despite the fact that *hur*-interrogatives (29%) are four times more frequent in the corpus than *vem*-interrogatives (7%). Second, there is no obvious correlation between the absolute and relative frequency of I-clefts in the material. Though the absolute frequency of *när*-clefts is very low in the material (1%), the ratio of *när*-clefts in relation to *när*-interrogatives is actually higher than for any other non-argument *wh*-phrase. These two observations strongly suggest that the likelihood of a particular *wh*-interrogative being formulated as a cleft is somehow linked to the *wh*-phrase rather than to frequency. In order to understand this frequency variation, we now turn to the issue of how I-clefts are used in discourse.

4. THE DISCOURSE POTENTIAL OF I-CLEFTS

In order to chart how I-clefts are used in Swedish, a total of 5,080 I-clefts were randomly chosen from the material and subsequently analyzed. For each *wh*-phrase included in the survey, a maximum of 500 randomly selected I-clefts were manually extracted from the corpus in the present and the past tense, respectively, giving a total maximum of 1,000 clefts per *wh*-phrase. In cases where the total number of I-clefts was less than 500, all hits were extracted. Table 2 presents the total number of extracted I-clefts for each *wh*-phrase.

The extracted I-clefts have been analyzed with regard to their function in the linguistic and situational context in which they occur. As the corpus supports

Tense	Vem	Vad	Var	Vart	När	Hur	Total
Present	500	500	389	500	500	500	**2,889**
Past	500	500	63	128	500	500	**2,191**
Total	1,000	1,000	452	628	1,000	1,000	**5,080**

Table 2
Extracted number of I-clefts per *wh*-phrase.

extended contexts, every linguistic item can be accessed in its original environment, i.e., the original forum post. That means that the I-clefts have not been categorized solely on the basis of the context immediately preceding or following the interrogative.

Two broad functions have been distinguished: I-clefts used for updating the speaker's knowledge state (Questions) and I-clefts used for some other communicative purpose (Non-Questions). Questions are uttered with the *expectation* of an answer (though an answer is not always provided), whereas the speaker does not expect an answer when uttering a Non-question I-cleft (though an answer *is* sometimes provided). The authentic interrogatives in (16) below illustrate each category. Unless stated otherwise, every linguistic sample in this section and onwards is taken from the data set.

(16)   (a)   *Question I-cleft:*

Ok vem var det som ringde från dolt   nummer 00.04?
ok who was it   that called from hidden number
'Who called from a hidden number at 12.04 a.m.?'

(b)   *Non-question I-cleft:*

Vem var det som bestämde att   yoghurt skulle inkluderas i
who was it   that decided   that yoghurt should include   in
duschkräm?
shower-gel
'Who decided that shower-gel should include yoghurt?'

The interrogative in (16a) is a direct request to reveal the identity of an anonymous caller at 12.04 a.m. In contrast, the speaker of (16b) does not wish to identify the inventor of yoghurt-based shower-gel: rather, she is using the interrogative to make a mocking comment about the perceived absurdity of using yoghurt in soap.

The distinction between a question and non-question is not always clear-cut, as illustrated by the following examples:

(17)   (a)   Vad är det som är så roligt?
what is it   that is so funny
'What's so funny?'

(b)   Hur är det egentligen man ska   göra?
how is it   actually   one shall do
'How are you actually supposed to do this?'

767

The *vad*-cleft in (17a) can certainly be used as a straightforward question, but is perhaps more often used to emphasize that the speaker does *not* find the situation particularly funny. Similarly, (17b) tends to be interpreted as a frustrated comment on the speaker's own lack of understanding ('I really don't understand how to do this'), but might equally well function as a true information-seeking question. In most cases such as these, the extended context has been sufficient to disambiguate the interpretation. However, some I-clefts could not be disambiguated by the context, and some I-clefts seem to be deliberately ambiguous within the context, being both Questions and Non-Questions at the same time. These unclassifiable I-clefts have been put in a third category, simply named Ambiguous.

4.1 *The distribution of question and non-question I-clefts*

Sections 4.2 and 4.3 discuss the distinction between Question and Non-Question I-clefts in more detail. But let us first look at the distribution of these two types in the material. Figure 3 shows the ratio of Question and Non-Question I-clefts in the corpora.
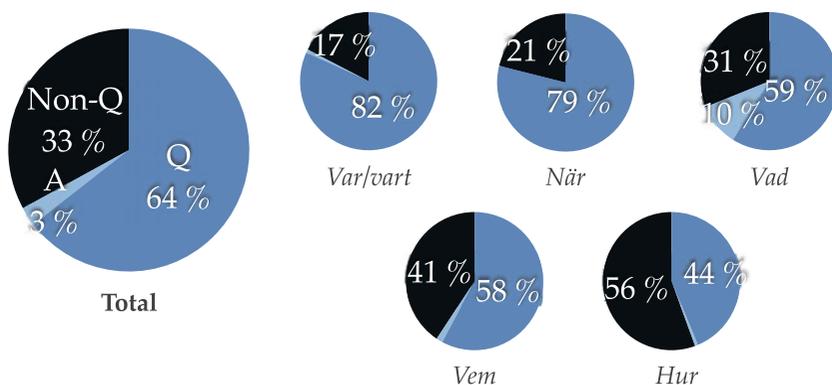

Figure 3
Share of Question (Q), Non-Question (Non-Q) and Ambiguous (A) I-clefts in the data set.
Total of 5,080 I-clefts.

As seen from Figure 3, Question I-clefts constitute 64% of the total number of I-clefts in the extracted material, but the frequency ratio is not equally distributed among the different *wh*-phrases. The vast majority of *var/vart*- and *när*-interrogatives are used to request information (79%–82%), whereas only 44% of the *hur*-interrogatives are used for this purpose. Note also that less than two thirds of the *vad*- and *vem*-clefts are used as questions (58%).

Combining the results of Figure 3 and Table 1, we may draw the following conclusions. Though the relative frequencies of *var*- and *hur*-clefts are equally low in the material, they are used differently: 8 in 10 *var*-clefts are questions, whereas only 4 in 10 *hur*-clefts are questions. In other words, there is no

immediate correlation between frequency and use, so that a rarely (or commonly) occurring *wh*-cleft is more likely to be used as a question (or non-question). Again, both relative frequency and use seem to be linked to the *wh*-phrase.

4.2 *Question I-clefts*

Let us now look at Question I-clefts in more detail. These I-clefts are used sincerely with the purpose of updating the speaker's knowledge state with new information:

(18) (a) Vad var det som gick snett?
          what was it   that went astray
          'What went wrong?'

       (b) När   var det Belgien skapades föresten[*sic*]?
          when was it   Belgium created   by-the-way
          'When was Belgium created, by the way?'

The most important observation concerning the Question I-clefts in the material is that they never occur out-of-the-blue. The information status of the proposition expressed in the clefted clause is either *old* or *inferrable* in the sense of Delin (1989): it either matches an explicit, previous utterance or is created by inference from the preceding discourse context. Compare the (italicized) clefted clauses in (19) below.

(19) (a) Du skrev även att dessa kvinnor blir   ännu mer
          you wrote also that these women becomes even more
          förtryckta, **vem** är det *som förtrycker dem* ?
          oppressed who is it   that oppresses them
          'You wrote that the oppression against these women is getting worse,
          but who is oppressing them?

       (b) Är lite nyfiken, **vad** var det *du inte gillade med*
          am little curious what was it   you not liked with
          konceptet?
          concept.DEF
          'I'm a bit curious; what didn't you like about the concept?'

       (c) Det där lät   ju inte så kul, **när** var det *du jobbade*
          that there sounded PRT not so funny when was it   you worked
          där?
          there
          'That didn't sound too great, when did you work there?'

       (d) **Var** är det *jag ska   skriva på*?
          where is it   I   should sign on
          'Where should I sign?'

769

In (19a), the proposition in the clefted clause is clearly inferrable from the previous utterance, as the claim that women are being oppressed presupposes an oppressor. The propositions in the clefted clauses in (19b), (19c) and (19d) are informationally old, as they match utterances made earlier in the discourse: the addressee has previously complained about the concept of platform video games and a negative working experience at the company *Stenbolaget* (lit. 'The Stone Company'), respectively. In the conversation leading up to (19d), the speaker has been urged to (electronically) sign a petition, but cannot find the link to do so.

The immediate connection to the topic under discussion is common to all Question I-clefts in the material. The primary function of a clefted *wh*-question in Swedish thus seems to be one of clarification: the cleft requests referential specification of a particular person/thing/place/time that has been actualized in the immediately preceding discourse context. In other words, clefted questions aim to clarify aspects of the current topic. No Question I-cleft in my data set is used out-of-the-blue, nor to introduce a new topic to the discourse universe.

In Brandtler (2012: 200–201) I make a similar observation, arguing that clefting seems to be preferred in Swedish whenever the speaker poses a question that relates to the immediate, surrounding context. Thus, the canonical interrogatives are marked in the following contexts:

(20) [Context: A sudden noise is heard.]

    (a) Vad var det som lät?
         what was it that sounded
         'What was that sound?'

    (b) #Vad lät?
          what sounded

(21) [Context: The addressee just ends a telephone call.]

    (a) Vem var det som ringde?
         who was it that called
         'Who called?'

    (b) #Vem ringde?
          who called

The results of the current survey empirically confirm my previous claim. It also seemingly conforms to Gundel's (2002: 125) observation that 'non-clefted *wh*-questions where the material in the cleft clause is strongly presupposed (. . .) are judged to be highly unnatural, if not totally unacceptable, in Norwegian.'

Furthermore, the Question I-clefts in the material are regularly used to remind the speaker of a state-of-affairs. This use is acknowledged also by Teleman et al. (1999: 4: 747), who mention that a clefted *wh*-question may signal that 'the speaker has known the answer previously or is acquainted with the question somehow.' Such questions are often supplemented with comments such as *nu* 'now' or *nu igen* 'now again'.

770

(22) (a) var   var det nu   du jobbade rent geografiskt?
         where was it   now you worked PRT geographically
         'Where did you work geographically, again?'

     (b) Vad var det craving hette nu   igen på svenska?
         what was it   craving was   now again in Swedish
         'What's the Swedish word for craving, again?'

To summarize this section, Question I-clefts in Swedish are strongly presup-
positional: the proposition expressed in the clefted clause is uncontroversially
part of the discourse universe, by virtue of being either informationally old
(a repetition of a previous utterance) or inferrable from the preceding context.
Question I-clefts are primarily used for the purpose of clarification: they request
referential specification of a person/thing/time etc. currently under discussion.
These observations lead us to conclude that there is a subtle, yet consistent,
contrast in Swedish between clefted and canonical *wh*-questions. In other words,
canonical and clefted *wh*-questions cannot be used interchangeably in discourse.

4.3 *Non-question I-clefts*

Non-question I-clefts are *wh*-interrogatives used for some other purpose than
updating the speaker's knowledge state. While any interrogative not used as
a proper question is traditionally labeled rhetorical (e.g., *Are you out of your
mind?* or *Nice weather, isn't it?*), such a broad definition may blur some subtle
but poignant differences with regard to the use of various I-clefts. At least
three distinct uses of Non-Question I-clefts have been discerned in the material,
although space restricts me from discussing each type at any length. See Brandtler
(in prep.) for a more in-depth discussion of Non-Question I-clefts in Swedish.
    Some Non-Question I-clefts express the speaker's attitude toward a state-of-
affairs. In terms of traditional rhetoric, these interrogatives would be labeled
erotema. Often (but not always) the utterance marks an emotional and/or negative
stance toward what is expressed in the clefted clause.

(23) (a) Vem är det som läser sånt här egentligen?
         who is it   that reads that here really
         'Who actually reads stuff like this?'

     (b) Vad är det du gnäller om?
         what is it you whine about
         'What are you whining about?'

In the discussion where (23a) occurs, the *vem*-cleft serves to belittle the quality
(and existence) of a particular blog. In another thread, the function of (23b) is to
mock the addressee's complaints. The interrogatives in (23) can quite naturally
embed under matrices such as *I can't (for the life of me) understand* . . . .

Other Non-Question I-clefts are used for highlighting or emphasizing a state-of-affairs made salient from the preceding context. Often, the speaker immediately provides the 'answer' to the (insincerely) posed question. Some of these interrogatives are used stylistically to sum up a longer discussion in a few words; see (24a). Others are used for calling the addressee's attention to something, often by referring to an earlier comment or remark made by the addressee; (24b).

(24) (a) Vad var det Mona sa   nu: "det är häftigt att betala skatt"
         what was it Mona said now it is cool   to pay   taxes
         'As Mona put it: it's cool to pay one's taxes.'

     (b) Hur var det nu   du sa, vi är alla människor.
         how was it   now you said we are all humans
         'As you put it yourself, we're all humans.'

Note that there is typically no question mark to indicate the end of the interrogative, nor does the entire sentence end with a question mark.

Lastly, some Non-Question I-clefts are rhetorical questions in the ordinary sense: they express questions to which the answer is obvious given the situational context. I assume, in accordance with Han (2002), that the rhetorical flavor is best captured in terms of (un)informativity: a rhetorical question presupposes the least informative answer, i.e., the answer most likely to be true, given the contextual situation and the conversational common ground. Some examples of Swedish interrogative clefts used rhetorically are found in (25) below.

(25) (a) Jaha, vem var det som blev   arg   nu då
         well who was it   that became angry now then
         'Well, who's angry now?'

     (b) När   var det senast det gick något   bra på Kanal   3?
         when was it   last   it   went anything good on channel 3
         'When was the last time Channel 3 aired anything good?'

As evident from the context in which these interrogatives occur, the answers to these 'questions' are self-evident and therefore uninformative. In (25a), the obvious answer is *you*, i.e., the addressee, and in (25b) *never*. Note that the rhetorical question in (25b) can be paraphrased as a negative assertion: *There's never anything good on Channel 3*.

4.4 *Brief summary*

In an attempt to analyze the use of I-clefts in Swedish, I have broadly distinguished between Question and Non-Question I-clefts.

Question I-clefts are primarily used for clarifying an aspect of the topic under discussion. They may also be used for reminding the speaker of a state-of-affairs. Non-Question I-clefts are used for expressing speaker attitude, for highlighting a state-of-affairs or for rhetorical purposes. I conclude that Swedish clefted and canonical interrogatives cannot be used interchangeably in discourse.

5. THE SEMANTICS AND INFORMATION STRUCTURE OF CLEFTS

In this section and the next, the semantics and information structure of both declarative and interrogative clefts are discussed, in order to understand (i) when and why a clefted interrogative is chosen over a canonical question; (ii) why some *wh*-interrogatives occur more frequently as clefts than others.

5.1 *Declarative clefts*

In the comprehensive literature on cleft constructions, there is a broad consensus that clefts have an information-structural function. I understand information structure to mean the structuring of the utterance into a *focal* (new, informative, rhematic) part and a *ground* (known, contextually bound, thematic) part. The proposition expressed in the clefted clause is typically presented as presupposed and thus represents the ground part, whereas the cleft phrase tends to express the focal part of the sentence.[8]

It has been widely acknowledged that cleft constructions express an existential presupposition, i.e., 'a presumption on the part of the speaker that the property denoted by the cleft clause is true of some individual' (Reeve 2011: 149); see, e.g., Jackendoff (1972); Halvorsen (1978); Atlas & Levinson (1981); Percus (1997); Hedberg (2000). This presumption is constant under negation, as seen from the fact that both sentences in (26) below presuppose that Mary hit someone.

(26) (a) It was JOHN that Mary hit.

   (b) It wasn't JOHN that Mary hit.

I assume with Vercauteren (2016: 47–51) that the existential presupposition associated with the clefted clause is semantic in nature, on par with the existential presupposition of definite descriptions. In order for a cleft to be felicitously used in discourse, however, the existential presupposition must be (pragmatically) *satisfied* (see Heim 1983): the information contained in the clefted clause must already be part of the conversational common ground when the sentence with the presupposition is uttered. This 'dual' representation of the cleft construction serves to capture the fact that a semantic presupposition typically is pragmatically presupposed in the context as well.[9] Note that an unsatisfied presupposition does not automatically lead to infelicitousness, as the semantic presupposition may also

---

[8] However, Hedberg & Fadden (2007) show that the clefted clause in *it*-clefts functions in equal measure as the topic *or* the comment (i.e., focus/new information) in spoken English; cf. also Delin (1995). This finding shows that the information structure of English *it*-clefts is not as static as is usually assumed.

[9] Beaver & Geurts (2014) give the following characterization of a satisfied presupposition.

   (i) S is satisfied in a context C iff $C + S = C$ (i.e., updating C with S has no effect).

   (ii) S presupposes S' iff S' is satisfied in all contexts where update with S is defined.

be accommodated in the sense of Lewis (1979); see also Beaver (1992). Hence, it is important to point out, as Johansson (2001: 557) does, 'that although the cleft clause in all types of cleft is presuppositional (. . .) this does not entail that it must contain information that is given. Instead, it seems clear that the cleft clause in all types of cleft may contain information that is either given or new.'

Hedberg (2000) proposes to treat cleft constructions and definite descriptions in a parallel fashion, based on the obligatory existential condition. The parallelism between clefts and definite descriptions becomes even greater when considering the exhaustivity condition associated with the cleft phrase, i.e., the 'presumption that the individual denoted by the clefted XP is the only (or maximal) contextually relevant individual of which the property denoted by the cleft clause holds' (Reeve 2011: 149). The information structure of the cleft in (26a) can be schematized as follows:

(27)   Sentence: It was JOHN that Mary hit.
       Presupposition: Mary hit someone
       Assertion: Mary hit John
       Exhaustive focus: John

Naturally, the asserted part is not inherent to the cleft *per se*: it evaporates in a conditional setting, whereas the presupposition and exhaustive focus remain (e.g., *If it was John that Mary hit, I'd be surprised*).

É. Kiss (1998, 1999) puts forward two strong arguments in favor of analyzing the cleft phrase as the exhaustive focus of the sentence. First, the cleft phrase cannot combine with *also/even*, adverbs known to be inherently incompatible with exhaustive focus; see (28a). Second, the cleft phrase cannot be a universal quantifier. A universal quantifier is per definition incompatible with exhaustive focus, as exhaustive focus involves 'identification by exclusion' (Kenesei 1986), and a bare universal does not exclude any members of the set; see (28b).

(28)   (a)   It was ??*also* / *even **sherry** that John drank.
       (b)   *It was **everything** that John drank.

É. Kiss's definition of identificational focus is given in (29).

(29)   *The function of identificational focus*: An identificational focus represents a subset of the set of contextually or situationally given elements for which the predicate phrase can potentially hold; it is identified as the exhaustive subset of this set for which the predicate phrase actually holds.

<div align="right">(É. Kiss 1998: 245)</div>

Lambrecht (2001: 477) makes a similar assumption, arguing that the relation between the focus and the proposition expressed in the relative clause is 'not one of predication but of *identification*' (italics in original).

Importantly, identificational focus contrasts with 'ordinary' focus, which is informational and non-exhaustive. É. Kiss (1998: 250) uses the following examples to illustrate the difference between the two types of foci.

<div align="center">774</div>

(30) (a) Mary picked **a hat and a coat** for herself.   info. foc
      (b) Mary picked **a hat** for herself.

(31) (a) It was **a hat and a coat** that Mary picked for herself.   id. foc
      (b) It was **a hat** that Mary picked for herself.

She states that while (30b) is a logical consequence of (30a), (31b) is *not* a logical consequence of (31a); instead it contradicts (31a).

5.2 *Clefted interrogatives*

What has been said so far relates to declarative *it*-clefts. However, the analysis quite naturally extends to clefted interrogatives. There is no reason to assume that the clefted clause should be interpreted any differently in interrogative clefts: the existential presupposition is as evident in clefted polar questions (32a) as in clefted *wh*-questions (32b).

(32) (a) Är det din   plånbok som ligger på bordet?
           is it   your wallet   that lies   on the.table
           'Is that your wallet on the table?'
      (b) Vad är det som ligger på bordet?
           what is it   that lies   on the.table.DEF
           'What's that on the table?'

Both interrogatives in (32) share the presupposition that something is on the table.

An important difference between clefted polar and *wh*-interrogatives, however, is the effect of a negative answer on the existential presupposition. The presupposition of the polar question is clearly constant under negation: a negative answer in response to (32a) simply denies that the thing on the table is the addressee's wallet. In contrast, a negative answer in response to (32b) is contradictory: it contradicts the presupposition that there actually is something on the table. Hence, (affirmative) clefted interrogatives cannot be felicitously answered in the negative. While it is probably the case that sincere questions, clefted or not, are most often asked with the *expectation* of a non-negative answer, a negative answer need not be infelicitous. Some canonical questions are resolved equally well by a negative answer, such as *What did you buy for your mother's birthday?*. However, the existential presupposition of the clefted clause is a condition under which the question I-cleft is answerable: in situations where the presupposition fails to be true, the question is not 'valid', as it were (cf. also Kennan & Hull 1973 and Comorovski 1996).

A strong argument in support of this claim comes from the interpretation of rhetorical *vem*-clefts. In the material, 291 *vem*-interrogatives (of 1,000) function as rhetorical questions. Though affirmative rhetorical questions typically imply negative answers (making them similar to negative assertions), every single rhetorical *vem*-question in the material implies a non-negative answer. Consider (33) below.

(33) (a) för   när de   blir   sjuka, vem är det som blir
         because when they become sick who is it   that becomes
         tvungen att betala?
         forced to pay
         'Because when they get sick, who will have to pay?'

     (b) Förlåt mig, men vem är det egentligen som är naiv?
         excuse me but who is it   really   that is naive
         'I'm sorry, but who's being naive here?'

In (33a), the speaker implies that *we* (i.e., the tax payers) have to pay for the healthcare of retired people living abroad. In (33b), the speaker implies that the addressee is being naive, in response to a previous exchange. The fact that the presupposition of existence associated with the clefted clause is present even in rhetorical contexts unequivocally indicates that it is inherent to the cleft construction itself.

At this point, it needs to be emphasized that the implication of existence discussed above is by no means unique to clefted *wh*-interrogatives: it is considered a standard property of *wh*-questions in general (see Katz & Postal 1964, Postal 1971, Karttunen 1977; for a more recent analysis also Haida 2008). As observed by Postal (1971: 73, fn. 6), (35) is just as much of a contradiction as (34), which suggests that the presupposition of the *wh*-phrase is just as strong as the presupposition associated with the existential quantifier in (34).

(34)   #Something is on the table and nothing is on the table.

(35)   #Although nothing is on the table, what is on the table?

According to most advocates of Postal's view, the presupposition of existence is inherent to the *wh*-phrase, in the same way as it is assumed to be inherent to definite descriptions. As a consequence, the negative answer is, strictly speaking, not an answer at all – it is a denial of the presupposition of existence, just as was argued above. The response in (36) is thus similar to the explicit denial in (37).

(36) A: Who came?>>
          Someone came.
      B: No-one.

(37) A: The King of France is bald.>>
          There is a king of France.
      B: There is no king of France.

One obvious problem with this alleged parallelism is the trivial but crucial fact that argument *wh*-questions *can* be felicitously answered in the negative. That is, the negative answer does not necessarily lead to a logical contradiction in the sense of Postal. Consider the two exchanges in (38) below.

(38) A: Vem åt Maria lunch med igår?   Canonical
          who ate Maria lunch with yesterday
          'Who did Maria lunch with yesterday?'

      B: Ingen.
          no-one

(39) A: Vem var det (som) Maria åt lunch med igår?   Cleft
          who was it   that Maria ate lunch with yesterday
          'Who was it that Maria lunched with yesterday?'

      B: #Ingen.
           no-one

The negative answer to the canonical question (38) resolves the issue at stake equally well as a non-negative answer. In contrast, the negative answer to the clefted *wh*-question in (39) is infelicitous, as discussed above, as it contradicts the existential presupposition that the proposition expressed in the clefted clause is true of some individual. In other words, the alleged existential presupposition associated with the canonical *wh*-question in (38) can be easily challenged, whereas the existential presupposition of the clefted *wh*-question in (39) cannot. This is a rather confounding conclusion under the assumption that both question types involve the same kind of semantic presupposition.[10]

I would argue that the implication of existence associated with canonical argument *wh*-questions is an implicature rather than a presupposition. This proposal is in line with Lie (1978), who argues that canonical *wh*-questions in Norwegian *supposes* rather than **pre**supposes existence. In order to arrive at this conclusion, first consider the question pair in (40).

(40) (a) Who has called this morning?

      (b) Has anyone called this morning?

Lambrecht (1994: 285) argues that an open question like (40a) is best analyzed as a 'conventionalized shortcut' for the more elaborated and cumbersome sequence *Did anybody call this morning, and if so, who?*. Quite naturally, posing a *wh*-question is a two-stage process. The primary function is to request referential specification of an underspecified member of a set of people/things/times etc. But ontologically prior to such a request is the issue of whether or not there exists any such person/thing/time etc. Simply put, by asking (40a) one simultaneously asks (i) whether or not there exists any x such that x called this morning and (ii) in case there was, what the identity of x is.

---

[10] While both Faarlund et al. (1997: 943) and Teleman et al. (1999: 4: 517) maintain that canonical and clefted *wh*-questions share the same presuppositions in Norwegian and Swedish, Hansen & Heltoft (2011: 1796) – the main reference grammar of Danish – argue that there has to be a presuppositional difference between the two constructions, as the cleft would be pleonastic otherwise.

From a communicative perspective, the *wh*-question in (40a) and the yes/no question in (40b) are, of course, not entirely parallel. By choosing to ask the *wh*-question, the speaker signals epistemic bias, i.e., that she holds it likely that *somebody* has called during the course of the morning. This view is explicated in Tomioka (2009: 264–65).

> To the question *What did Sue buy?*, one can truthfully answer by saying *She bought nothing* without creating ill effects of presupposition failure. In general, the proposition created by existentially binding a Wh-variable is not a presupposition but should be considered an epistemic bias on the part of the speaker (cf. Han & Romero 2004). When the speaker asks *What did Sue buy?*, she is inclined to believe that Sue bought something. It is nonetheless uncertain that the addressee shares the same belief. Thus, the speaker should not be too shocked to find that her bias turns out to be false.

Importantly, the yes/no question in (40b) would be the more natural choice should the speaker's expectations be epistemically unbiased.

In his article on the social economics of questions, Levinson (2012: 23) hypothesizes that speakers will never ask a *wh*-question where a yes/no question would do. According to Levinson, *wh*-questions invoke higher social costs than yes/no questions, in terms of an increasing risk of face loss, and/or clues to the speaker's interests and concerns etc. The more specific the question, the higher the social cost. Hence, Levinson hypothesizes, speakers strive to ask for the smallest informational increment they think they need: the use of 'queclaratives' (declaratives disguised as questions) is the least costly option, followed by polar questions.

I would argue that the existential implication associated with the question in (40a) arises as a consequence of the social and informational hierarchies hypothesized by Tomioka and Levinson. By choosing the costliest option (the *wh*-question), the speaker signals that a less costly question would not do in that particular context. In other words, posing a *wh*-question signals that the issue at stake is *not* primarily the truth of the proposition. Had the speaker been less certain as to the truth of the proposition, she would have chosen a less costly option, such as a yes/no question, in order not to risk 'that her bias turns out to be false'. Hence, via the Gricean Maxim of Quantity ('make your contribution as informative as is required'), a *wh*-question gives rise to a generalized conversational implicature to the effect that the proposition is true of someone/something. As with any conversational implicature, however, the implication of existence can be canceled. In the case of argument questions, the implicature is canceled by a negative answer, as we saw in (38a) above.[11]

Assuming with Tomioka (2009) that a canonical *wh*-question expresses 'epistemic bias on the part of the speaker', we may capture the difference in implicational strength between canonical and clefted argument *wh*-questions in

---

[11] Dryer (1996: fn. 14) also entertains the idea that the existential implication associated with *wh*-questions is a conversational implicature, although he does not pursue this idea further.

Swedish. A clefted *wh*-question expresses an existential presupposition that is contextually satisfied, meaning that it is already part of the common ground of the discourse participants. A canonical argument *wh*-question, in contrast, expresses a generalized conversational implicature to the effect that the speaker holds the proposition in the clefted clause as true; this implicature arises from the active choice on the part of the speaker of posing a socially costlier *wh*-question over a yes/no question. If this characterization is correct, (affirmative) yes/no questions, canonical *wh*-questions and clefted *wh*-questions in Swedish form a scale ranging from 'no implicature', via 'generalized implicature' to '(semantic) presupposition'. For our purposes, the most important observation is that the existential presupposition associated with clefted *wh*-interrogatives is inherent to the cleft construction as such, and different from the implication of existence associated with canonical argument *wh*-interrogatives.

5.3 *The cleft phrase*

Let us finally turn to the cleft phrase in I-clefts, i.e., the *wh*-phrase. In Swedish *it*-clefts, the cleft phrase is primarily used for focusing anaphoric elements (pronouns and adverbials), as shown by Johansson (2001); the same holds for Norwegian (Søfteland 2014: 175). As pointed out by Søfteland (2014: 214), I-clefts are different in this respect, as the cleft phrase (i.e., the *wh*-phrase) never denotes referentially known information. Again, this distinguishes clefted *wh*-interrogatives from clefted polar interrogatives, in which the cleft phrase tends to be anaphoric. Consider the following Norwegian examples of clefted polar questions, taken from Søfteland (2014: 215).

(41) (a) Ove Røssbakk$_i$, var det **han**$_i$ som skulle vere med?
        Ove Røssbakk was it he that should be with
        'Ove Røssbakk, was it him that wanted to participate?'

    (b) Northåv$_j$, var det **det**$_j$ han hette?
        Northåv was it that he named
        'Northåv, was that his name?'

Given that the *wh*-phrase is the informational focus of a *wh*-question, the main issue to address in this section is whether the cleft phrase has an exhaustive interpretation in the sense of É. Kiss (1999), as argued to be the case in declarative clefts. In order to address this issue, we must look more closely at the semantics of *wh*-phrases.

Perhaps the most influential semantic analysis of questions originates with Hamblin (1973), who argued that the meaning of a *wh*-question is a set of propositions constituting its possible answers. Under such an analysis, (i) a question like *Who left?* denotes a set of propositions expressed by sentences such as 'Bill left', 'Mary left' etc., and (ii) *wh*-phrases have the same denotation as ordinary noun phrases and equal existential quantifiers or indefinites. Thus, basic *wh*-phrases like *who* and *what* translate into *somebody* and *something*, see (42a).

779

For the sake of illustration, the semantic value of the question *Who left?* is given below.

(42) Who left?

    (a) $[\![\text{who}]\!](w) = [\![\text{somebody}]\!](w) = \lambda Q \,\exists x\, [\mathbf{person}(x)\,(w) \wedge \boldsymbol{Q}(x)\,(w)]$

    (b) $[\![\text{who left}]\!](w) = \lambda p \,\exists x\, [p(w) \wedge \mathbf{person}(x)\,(w) \wedge p = \lambda w\, (\mathbf{left}(w)\,(x))]$

    (c) $Q = \{\text{Bill left, John left, Mary left}\}$

A persistent problem with the Hamblin-based approach is that it does not specify in more detail what 'possible answers' are supposed to be. Karttunen (1977) proposed that the set of possible answers should be delimited to include only the 'true and complete' answers. For example, if Bill, John, and Mary left in the actual world, the denotation of the *wh*-question in (42) is the set of propositions shown in (42c) – and no other.[12] Each proposition in the set constitutes a partial answer to the question, and jointly they constitute its true and complete answer (Karttunen 1977: 20). The analyses of Higginbotham & May (1980) and Groenendijk & Stokhof (1982, 1984) are similar in assuming that a *wh*-question is only completely answered by an exhaustive list of individuals who satisfy the relevant predicate.

But not all *wh*-questions require an exhaustive answer. So called 'mention-some' questions can be resolved by mentioning one of several possible persons, places, reasons etc.

(43) (a) Var   kan jag köpa svenska dagstidningar?
         where can I   buy Swedish newspapers
         'Where can I buy Swedish newspapers?'

    (b) Hur går jag till stationen?
         how go I   to station.DEF
         'How do I get to the station?'

The questions in (43) do not require an exhaustive list of places that sell Swedish newspapers or ways of getting to the station in order to be resolved. Note, however, that mention-some questions resist clefting in Swedish, as illustrated in (44) below (see also Lie 1978: 396 for a similar claim regarding Norwegian):

(44) (a) ??Var   är det jag kan köpa svenska dagstidningar?
          where is it I   can buy Swedish newspapers
         'Where is it I can buy Swedish newspapers?'

    (b) ??Hur är det jag går   till stationen?
          how is it I   walk to station.DEF
         'How is it I get to the station?'

---

[12] In the extensional language of the formula, explicit variables stand for possible worlds. When *w* is free, its value is the actual world. The condition $p(w)$ thus ensures that only the true answers are considered in the set Q.

The awkwardness of the questions in (44) points to an important observation: clefted *wh*-questions in Swedish require exhaustive answers. This observation, in turn, seems natural if we assume with É. Kiss (1998: 245) that the cleft phrase (i.e., the *wh*-phrase) is the *identificational* focus of the sentence, requesting 'the exhaustive subset of [the] set for which the predicate phrase actually holds'.[13]

The identificational focus of the cleft phrase in combination with the existential presupposition of the clefted clause ensures that the *wh*-phrase of a clefted argument *wh*-interrogative denotes a non-empty set of people/things. This analysis predicts that negative polarity items (NPIs) should be ruled out in argument *wh*-clefts, as they are within the scope of an existential quantifier (i.e., ⟦who⟧ = ⟦somebody⟧). This prediction is borne out, as observed by Brandtler (2012: 194). Consider the canonical and clefted *vem*-questions in (45) below.

(45)　(a)　Vem har *någonsin* varit i　Seoul?
　　　　　　who has　NPI　　　been to Seoul
　　　　　　'Who has ever been to Seoul?'

　　　(b)　*Vem är det som *någonsin* varit i　Seoul?
　　　　　　who is it　that NPI　　　been to Seoul?

The canonical *vem*-question in (45a) is close in meaning to the yes/no question *Has anyone ever been to Seoul?*, leaving open the possibility of the *wh*-phrase denoting an empty set of people. The existential presupposition of the clefted *vem*-question (45b), however, forces the interpretation of the *wh*-phrase as an existential quantifier, which effectively blocks the occurrence of NPIs within the scope of the *wh*-phrase.[14]

5.4　*Section summary*

In this section, I have argued that I-clefts can be given the same information-structural analysis as is typically assumed for declarative *it*-clefts. The clefted clause constitutes the ground, expressing a contextually satisfied existential presupposition. Support for this analysis comes from the fact that the existential

---

[13] As pointed out by an anonymous reviewer, the clefts in (44) tend to be interpreted as information-*confirming* questions, i.e., as requests to remind the speaker of something that has been previously mentioned, such as a number of places where you can buy Swedish papers (in the context of (44a)). However, the availability of this interpretation only serves to emphasize the point: it seems very difficult to interpret the questions in (44) as information-seeking, mention-some questions. Perhaps a more correct characterization would be that I-clefts cannot receive a mention-some reading in Swedish, rather than saying that mention-some questions cannot be clefted.

[14] The NPI-licensing property of *wh*-questions has received a fair amount of attention in the literature; see e.g., Progovac (1994), Krifka (1995), Han (2002), Guerzoni (2003), van Rooy (2003), Mayr (2013), Nicolae (2015). Though there is no consensus as to the licensing mechanism of canonical *wh*-questions like (45a), the *anti*-licensing mechanism of Swedish *vem*-clefts as in (45b) is straightforward under the proposed analysis: the very definition of an NPI is that it cannot occur within the scope of existential determiners.

presupposition is present even in rhetorical *vem*-questions. The cleft phrase expresses the identificational focus of the utterance, and is as such exhaustive. An argument in favor of this analysis is that I-clefts seemingly cannot receive a mention-some reading in Swedish.

Canonical argument *wh*-questions in Swedish are not presuppositional. Instead, they express epistemic bias on the part of the speaker, signaled by the speaker's active choice of using a *wh*-question over a yes/no question. This choice gives rise to a generalized conversational implicature to the effect that the proposition holds true of someone/something.

6. MATCHING THE MEANING AND FREQUENCY OF I-CLEFTS

In this final section, I attempt to relate the findings of my empirical survey to my account of the meaning and information structure of I-clefts as presented in the previous section. In order to facilitate the discussion, I distinguish between three different groups of *wh*-words: argument (*vem* 'who', *vad* 'what'), framing (*när* 'when', *var/vart* 'where'), and, constituting a group of its own, *hur* 'how'.

6.1 *Argument I-clefts*

We saw in Section 4.1 that argument *wh*-interrogatives are the most likely to appear as clefts in the material. Though the majority of these I-clefts are questions (58% of the *vem*-clefts, and 59% of the *vad*-clefts), a considerable number are non-questions.

I argued in Section 5.2 that clefted *wh*-questions have an existential presupposition to the effect that the property denoted by the clefted clause holds true of some individual. This presupposition is contextually satisfied, meaning that the information expressed is already part of the common ground. Consider the (authentic) questions in (46) below.

(46) (a) Vem var det som de   grep   i skogspartiet bredvid
           who was it   that they arrested in forest.part next.to
           skolan?
           school.DEF
           'Who did they arrest in the small forest by the school?'

      (b) Hallå ja, vem är det jag pratar med?
           hello yes who is it   I   talk   with
           'Yes, hello, who am I talking to?'

The question in (46a) is posted in a thread discussing the school shooting at Sandy Hook Elementary School in Newtown, Connecticut in 2012. The immediately preceding posts (all posted within a few minutes of each other) concern the identity of the shooter, as two names were circulating in the news at the time. Thus, the proposition in the clefted clause is clearly part of the common ground

782

of the discourse participants. In (46b), the proposition *I am talking to someone* is, of course, trivially known to the participants.

From a communicative perspective, a clefted question is a highly efficient means of updating the speaker's knowledge state in contexts such as those illustrated in (46) above. With the clear partition of (exhaustive) focus and ground, the information packaging of the I-cleft is optimal for singling out a specific aspect that needs to be clarified in relation to the topic under discussion. Naturally, a canonical argument *wh*-question is not entirely ruled out in contexts where the event denoted by the question is known to both discourse participants, but it is, in information-structural terms, a less efficient way of updating the speaker's knowledge state. By using the cleft construction, the speaker clearly signals how the question relates to the common ground and to the ongoing topic of discussion, instead of (in the case of a canonical *wh*-question) only implicating it.

I propose that the frequency of argument I-clefts in the material is immediately linked to the existential presupposition associated with the cleft construction. That is, the fact that 98% of the I-clefts in the material are *vem-* and *vad-* interrogatives can be attributed to the difference in implicational strength between canonical and clefted argument interrogatives. Whereas the I-cleft semantically and pragmatically presupposes the existence of an individual/thing that satisfies the (embedded) predicate, the canonical variety merely – and indirectly – implicates such an existence. The 'cost' of using a more complex syntactic structure (the cleft) is thus counterbalanced by the benefit of being able to pose an informationally more specific question in the given context.

6.2 *Framing I-clefts*

As shown in Table 1 and Figure 2, framing I-clefts are relatively rare in absolute numbers: less than 2% of the total number of clefted interrogatives are introduced by *när* 'when' or *var/vart* 'locational/directional where'. Clefts introduced by *var/vart* are rare also in relative terms, whereas *när*-clefts are the third most likely *wh*-phrase to introduce a cleft (following *vad* 'what' and *vem* 'who'). There is, however, a striking difference in frequency compared to argument I-clefts, and the reason may be related to the meaning of canonical framing *wh*-questions.

The function of framing *wh*-questions, irrespective of syntactic form, is to request specification of a spatiotemporal anchor to the event under discussion. In contrast to argument questions, framing *wh*-questions express full propositions. Prototypically, a question of the form *When p* or *Where p* presupposes *p*, as illustrated below:

(47)   (a)   **When** did you move to New York? >>
             You have moved to New York.

       (b)   **Where** did you buy that sweater? >>
             You have bought that sweater.

783

The presuppositional status of framing *wh*-questions is supported by two independent facts. First, they are not 'conversational shortcuts' in the sense of Lambrecht (1994) (cf. (40) above); i.e., they do not encompass an unsettled yes/no question:

(48) (a) When did you move to New York? ≠

     (b) Have you moved to New York?

(49) (a) Where did you buy that sweater ≠

     (b) Did you buy that sweater?

Second, they cannot be answered in the negative. The only way of answering framing questions in the negative is to explicitly challenge the presupposition, as in (50) below.

(50) A: When did you move to New York?

     B: Well, I haven't *moved* exactly. I'm just renting a room for now.

It seems reasonable to assume that the relative infrequency of framing I-clefts in the material is connected to the presuppositional status of the canonical variety: the information-structural gain of choosing the syntactically more complex construction is less obvious with framing interrogatives than with argument interrogatives. The puzzling issue to address, then, is under what circumstances a clefted framing *wh*-interrogative is preferred over a canonical one.

Looking first at *var*-questions, the clefts in the material seem to be used similarly to argument I-clefts: they request clarification of an aspect related to the topic under discussion; see examples (51) below.

(51) (a) jag har windows 8.1 Var   är det jag ska   högerklicka
         I   have Windows 8.1 Where is it   I   should right-click
         menar du?
         mean   you
         'I have Windows 8.1. Where did you say I should right-click?'

     (b) Var   är det du har läst att ''de   är totalt likgiltiga''?
         where is it   you have read that they are totally indifferent
         'Where did you read that "they are totally indifferent"?'

In (51a), the speaker is referring to a previous suggestion by the addressee to right-click to show display options in the traybar. In (51b), the speaker is literally quoting the addressee, asking him to specify where he found that text passage. In neither question is there an issue of whether the requested location exists: rather, the speaker requests a more *exact* specification. In some *var*-questions, this specification is made explicit by the addition of *exakt* 'exactly' to the immediate left of the *wh*-phrase.

(52) (a) Exakt   var   är det du kollar minnesanvändningen?
         exactly where is it   you check memory.use
         'Exactly where do you check memory use?'

(b) Exakt var   är det det har brunnit?
exactly where is it   that has burned?
'Exactly where has it been burning?'

Framing I-clefts introduced by *när* 'when' are used in much the same way, as illustrated in (53).

(53) (a) När   är det du ska   rycka in?
when is it   you shall join   up
'So when are you joining up?'

(b) När   är det spelet   släpps?
when is it   game.DEF releases
'When is the game released?'

The cleft in (53a) connects to a previous question raised by the addressee on how to prepare for his upcoming military service. The clefted question in (53b) is posted in a discussion thread about the upcoming release of the video game *Age of Conan*. Thus, the clefted *wh*-questions in (51)–(53) seem to be motivated by the need to specify an exact place or point in time. Note that an approximate place/time has been mentioned previously in the discourse. I thus conclude that clefted framing *wh*-questions are primarily used for specifying already mentioned times or places.

Approximately 22% of the clefted *när*-interrogatives in the material are Non-questions (see Figure 3). Of these, more than two thirds are rhetorical questions, or 146 in total. Interestingly, all but three examples involve the adverbs *sedan* (*sen*) 'since' or *senast* or *sist* 'last time'. Whereas *sedan* always precedes the *wh*-phrase, *senast* and *sist* may occur either in the matrix clause (see (25b) above) or the clefted clause, (54b) below.

(54) (a) **Sedan** när   var det sysslolösa invandrargrupper som stod   för
since   when was it   idle   immigrant.groups who stood for
huliganvåldet?
hooligan.violence.DEF
'Since when are idle immigrants responsible for hooligan violence?'

(b) När   var det vi såg röken   av din   skapare **senast**?
when was it   we saw smoke.DEF of your maker   last
'When was the last time we saw a trace of your maker?'

The rhetorical questions in (54) can be paraphrased as negative assertions: *Idle immigrants have never been responsible for hooligan violence* and *We have never seen the traces of your maker*. Example (54b) is especially noteworthy, as the expression *se röken av* (literally 'see the smoke of') is a negative polarity item (NPI), and hence cannot occur in affirmative assertions:

(55) (a) Jag har aldrig sett röken   av honom.
I   have never seen smoke.DEF of him
'I have never seen a trace of him.'

785

(b) *Jag har sett röken av honom.
    I   have seen smoke.DEF of him

Intriguingly, it seems that the presence of *senast* 'latest' in clefted *när*-interrogatives may trigger a non-factual reading of the embedded proposition, which in turn enables a rhetorical interpretation. Clefted rhetorical *när*-clefts are thus different from clefted rhetorical *vem*-questions (discussed in Section 5.2), in that the existential presupposition may evaporate. It should be pointed out, however, that both *senast* and *sist* (though apparently not *sedan*) are compatible with sincere questions as well, as illustrated in (56) below. Hence, it is not the presence of these adverbs alone that triggers the rhetorical reading.

(56) (a) När var det **senast** du åt där?
        when was it   last   you ate there
        'When was the last time you ate there?'

    (b) När var det **sist** du hade pojkvän   / vän
        when was it   last you hade boyfriend   friend
        'When was the last time you had a boyfriend / friend?'

Judging from the material, it seems as though *senast* or *sist* are more likely to trigger rhetorical readings whenever they occur in the clefted clause; in every (sincere) question with *senast* and *sist*, the adverb occurs in the matrix clause. However, the issue of whether the syntactic distribution of *senast* and *sist* can affect the interpretation of the interrogative will be left open to future research.

    Summarizing this section, I have argued that the relative infrequency of framing I-clefts can be attributed to the fact that canonical framing *wh*-interrogatives are presuppositional: there is in most contexts no immediate communicative gain in choosing the more complex syntactic structure. However, the framing I-clefts we do find are used for specifying a more exact place or point in time than what has previously been mentioned in the conversation.

6.3 Hur-*clefts*

Let us finally turn to I-clefts introduced by *hur* 'how'. As illustrated in Figure 2, *hur*-clefts constitute less than 1% of the total number of I-clefts in the material. In relative terms, *hur* is also the least likely *wh*-phrase to introduce an I-cleft (0,12%); see Table 1. The infrequency of *hur*-clefts is especially striking in relation to the overall frequency of interrogatives introduced by *hur* in the material: *hur*-interrogatives constitute almost a fourth of the total number of interrogatives (see Figure 1). In addition, the majority of the *hur*-clefts are Non-questions, i.e., they are not primarily used to request information (see Figure 3).

    Compared to I-clefts introduced by other *wh*-phrases, *hur*-clefts are quite formulaic: 61% of the Non-questions involve the main predicate *säga* 'say/tell',

786

and 21% of the Question *hur*-clefts involve the main predicate *göra* 'do'.[15] The generic pronoun *man* 'one/you' (literally 'man') is the subject in the majority of these constructions, as illustrated in (57) below.

(57) (a) Hur är det man gör?
           how is it   you do
           'How do you do this?'

      (b) Hur är det man säger på svenska "sådan husse sådan hund"
           how is it   you say in Swedish like   master like   dog
           'What do you call it in Swedish, "like master like dog".'

Apart from the generic predicate *göra* 'do', other frequent main predicates in question *hur*-clefts are *fungera* 'work/function', *räkna ut* 'calculate' and *ligga till*, literally 'lying to', but meaning how things stand. In question *hur*-clefts, then, the *wh*-phrase almost exclusively functions as a manner adverbial, requesting a specification of how to execute a particular action.

In Non-question *hur*-clefts, in contrast, the most frequently occurring predicates apart from *säga* 'say/tell' are *skriva* 'write', *uttrycka* 'express', as well as *skrika* 'scream' and *sjunga* 'sing'. In combination with these predicates, the meaning of *hur* is close to *vad* 'what': it requests an argument to the verb (what is said, written, expressed etc.). Consider also the idiomatic translation of (57b) above, involving *what* rather than *how*.

This observation suggests a rather intricate relation between the meaning and use of *hur* 'how' in I-clefts. The manner adverbial *hur* is used in questions, whereas argumental *hur* is used in non-questions.

Similar to framing *wh*-questions, *how*-questions are presuppositional irrespective of form: a question of the form *How p* presupposes *p*. Again, it seems reasonable to assume that the infrequency of *hur*-clefts is related to the presuppositional status of the canonical variety, just as argued for framing I-clefts. The observation that the main bulk of the *hur*-clefts in the material is made up of only a few constructions further strengthens the argument that *hur*-interrogatives generally resist clefting in Swedish.

7. CONCLUSION

When posing a question, the speaker has a repertoire of structures at her disposal. It seems reasonable to assume that the final choice (canonical interrogative, cleft, declarative) is not entirely random, but reflects the speaker's view of the hearer's information state at the time of the utterance, as well as her subconscious calculation of the social costs involved. Engdahl (2006: 93) argues that the

---

[15] In the present tense, no less than 81% of the Non-question *hur*-clefts involve the predicate *säga* 'say/tell'.

way questions are realized in discourse rather systematically correlates with 'the speaker's view of what the hearer might know and what has happened so far in the conversation.'

The current study of the meaning and use of clefted interrogatives in Swedish supports Engdahl's view. Clefted interrogatives do not have the same general interpretation as canonical interrogatives in Swedish, and the choice of using a clefted interrogative is discourse-motivated.

More specifically, clefted *wh*-interrogatives in Swedish always relate immediately to the topic under discussion and, when used as questions, serve to clarify a matter in relation to this topic. Clefted *wh*-interrogatives are not used in out-of-the-blue contexts.

I have argued that interrogative cleft constructions can be given the same information-structural analysis as declarative cleft constructions: the clefted clause expresses an existential presupposition, and the cleft phrase (the *wh*-phrase) constitutes the identificational focus of the utterance. The implication of existence associated with canonical argument questions is a conversational implicature, triggered by the choice on the part of the speaker of using a *wh*-question rather than a yes/no question.

The results of a corpus survey of informal Swedish showed that argument clefts are the most common in both absolute numbers and relative frequency. The presuppositional status of clefted argument *wh*-interrogatives was argued to be immediately linked to their frequency in the material. In contexts where the event denoted by the question is known to both discourse participants, the clefted variety is the optimal choice, due to its clear partitioning of focus and ground. Since interrogatives introduced by *when*, *where* and *how* are presuppositional regardless of syntactic form, the gain of using a cleft is less obvious – hence the infrequency of *when*-, *where*- and *how*-clefts in the material.

There are a number of issues that, unfortunately, have only been touched upon in this paper. First, the current study did not include D-linked *wh*-phrases. As D-linked *wh*-phrases imply the existence of a contextually determined set of entities, they are presuppositional in a sense that non-D-linked *wh*-phrases are not. Since, judging from my results, the main function of clefting a *wh*-interrogative is precisely to signal existence, one may suspect that D-linked *wh*-phrases resist clefting to a higher degree than non-D-linked *wh*-phrases. Future studies will show whether this prediction is borne out. Second, the issue of how clefted interrogatives are used rhetorically deserves to be more properly addressed, especially considering that rhetorical *vem*-questions maintain an existential presupposition, whereas rhetorical *när*-questions tend to be interpreted as non-factual. The licensing of NPIs in clefted interrogatives is another topic that requires further looking into. Lastly, the observed fact that (declarative and interrogative) clefts are much more common in the Scandinavian languages than in the other Germanic languages, including English, is certainly both interesting and confounding, and has not received the attention it deserves. These open issues, together with the claims made in this paper, point to the importance of

acknowledging the differences between canonical and clefted interrogatives, both with regard to their meaning and possible uses in discourse.

APPENDIX

The study investigated the frequency and distribution of clefted *wh*-interrogatives (I-clefts) in the Flashback Forum via the Språkbanken corpora resource (www.spraakbanken.gu.se). At the time of the survey, the Flashback corpus consisted of 15 subcorpora, each representing an overarching topic of discussion. For unknown technical reasons, the subcorpus *Hem, bostad & familj* 'Home, housing & family' was not available during parts of the survey, and was therefore excluded. Table A1 summarizes the size and tokens of the 14 remaining subcorpora at the time of the survey.

Corpus	No. of tokens	No. of sentences	Last update
Dator & IT ('Computer & IT')	272,251,243	18,379,450	Jan. 25, 2015
Droger ('Drugs')	198,213,690	12,923,116	Jan. 29, 2015
Fordon & trafik ('Vehicles & Traffic)	46,274,293	3,053,377	Jan. 30, 2015
Kultur & Media ('Culture & Media')	280,105,484	20,396,623	Feb. 04, 2015
Livsstil ('Lifestyle')	91,949,837	6,575,267	Feb. 05, 2015
Mat, dryck & tobak ('Food, Drink & Tobacco')	52,953,030	3,805,314	Feb. 06. 2015
Politik ('Politics')	399,937,411	24,551,461	Feb. 11, 2015
Resor ('Travels')	28,632,782	1,855,627	Feb. 11, 2015
Samhälle ('Society')	453,217,217	28,497,194	Feb. 13, 2015
Sex ('Sex')	66,619,716	4,547,429	Feb. 15, 2015
Sport & träning ('Sport & Fitness')	176,120,282	2,681,979	Feb. 26, 2015
Vetenskap & humaniora ('Science & Art')	289,259,367	17,987,642	Mar. 02, 2015
Övrigt ('Miscellaneous')	91,608,784	6,619,644	Feb. 09, 2015
Om Flashback ('About Flashback')	21,120,960	1,403,771	Feb. 16. 2015
Total	**2,468,264,096**	**153,277,894**	

Table A1
The Flashback Forum: size and tokens of the surveyed subcorpora.

Table A2 summarizes the number of *wh*-interrogatives in each respective subcorpus as generated by the search strings. Note that these figures have not been checked manually, meaning that the actual number of main clause *wh*-interrogatives may deviate slightly from the results in Table A2.

789

Corpus	**Wh**-phrase						**Total**
	Vem	*Vad*	*Var*	*Vart*	*När*	*Hur*	
Dator	18,638	305,717	31,362	21,947	16,114	216,107	**609,883**
Droger	17,480	178,627	21,666	12,726	13,091	116,114	**359,704**
Fordon	3,589	50,568	5,718	2,986	2,132	23,489	**88,482**
Kultur	55,900	268,106	38,592	16,452	20,069	119,404	**518,523**
Livsstil	7,661	93,528	12,289	8,624	4,629	43,259	**169,990**
Mat	3,793	45,306	7,245	3,957	2,228	23,701	**86,230**
Politik	75,782	418,321	47,836	17,896	30,636	224,989	**815,460**
Resor	1,640	28,965	4,985	4,198	1,928	22,697	**64,413**
Samhälle	76,764	492,836	70,541	27,974	31,155	291,226	**990,496**
Sex	12,503	60,806	8,347	3,674	2,932	34,142	**122,404**
Sport	26,238	156,342	17,785	9,059	13,675	82,610	**305,709**
Vetenskap	36,537	303,931	34,141	12,924	15,279	194,607	**597,419**
Övrigt	17,542	108,231	13,824	6,785	6,710	55,241	**208,333**
Om Fl.	3,507	21,796	2,769	1,181	1,483	12,074	**42,810**
Total:	**357,572**	**2,533,080**	**317,100**	**150,383**	**162,061**	**1,459,660**	**4,979,856**

Table A2
Number of *wh*-interrogatives per *wh*-phrase in the subcorpora.

Table A3 presents the total number of I-clefts in the material in the past and present tense.

Tense	Vem	Vad	Var	Vart	När	Hur	Total
Present	14,684	161,810	389	585	504	775	178,747
Past	3,911	17,457	63	128	852	993	23,404
Total	18,595	179,267	452	713	1,356	1,768	202,151

Table A3
The distribution of I-clefts in the present and past tense per *wh*-term.

The entire data sets for *vem-*, *var-*, *vart-* and *när*-clefts were checked manually for incorrect (non-cleft) structures. The figures in Table A3 thus reflect the actual number of I-clefts in the material rather than the total number of hits generated by the search string. As the data sets for *vad-* and *hur*-interrogatives were too large to check in their entirety, the numbers in Table A3 are estimates based on the ratio of I-clefts in a random sample set. The size of each sample set is specified below.

- **I-clefts introduced by *Vad/vad***: the estimated number of I-clefts is based on the average share of I-clefts in 2% (or 5,551 sentences) of the total number of hits in the present tense (277,547) and 10% (or 2,548 sentences) of the total number of hits in the past tense (25,484). In the present tense, 58,3% (3,237 sentences) were I-clefts, giving an estimated total of 161,810 *vad*-clefts in the present tense (58,3% of 277,547). In the past tense, 68,5% (1,747 sentences) were I-clefts, giving an estimated total of 17,457 *vad*-clefts in the past tense (68,5% of 25,484).
- **I-clefts introduced by *Hur/hur***: the estimated number of I-clefts is based on the average share of I-clefts in 5% (or 3,590 sentences) of the total number hits in the present tense (71,799) and 20% (or 3,009 sentences) of the total number of hits in the past tense (15,046). In the present tense, only 1,08% (39 sentences) were I-clefts, giving an estimated total of 775 *hur*-clefts in the present tense (1,08% of 71,799). In the past tense, 6,6% (199 sentences) were I-clefts, giving an estimated total of 993 *hur*-clefts in the past tense (6,6% of 15,046).

REFERENCES

Atlas, Jay & Stephen C. Levinson. 1981. *It*-clefts, informativeness, and logical form. In P. Cole (ed.), *Radical pragmatics*, 1–62. New York: Academic Press.

Beaver, David. 1992. The kinematics of presupposition. In P. Dekker & M. Stockhof (eds.), *Proceedings of the Eighth Amsterdam Colloquium*, 17–36. Amsterdam: University of Amsterdam, ILLC.

Beaver, David & Bart Geurts. 2014. Presupposition. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. Stanford, CA: Stanford University; winter 2016 edn. https://plato.stanford.edu/archives/win2016/entries/presupposition/.

Biber, Douglas & Susan Conrad. 2009. *Register, genre, and style*. Cambridge: Cambridge University Press.

Borin, Lars, Markus Forsberg & Johan Roxendal. 2012. Korp – the corpus infrastructure of Språkbanken. *Proceedings of LREC 2012*, 474–478. Istanbul: ELRA.

Brandtler, Johan. 2012. *The evaluability hypothesis. The syntax, semantics and pragmatics of polarity item licensing*. Amsterdam: John Benjamins.

Chomsky, Noam. 1977. On *wh*-movement. In Peter W. Culicover, Thomas Wasow & Adrian Akmajian (eds.), *Formal syntax*, 71–132. New York: Academic Press.

Clech-Darbon, Anne, Georges Rebuschi & Annie Rialland. 1999. Are there cleft sentences in French? In Georges Rebuschi & L. Tuller (eds.), *The grammar of focus*, 83–118. Amsterdam: John Benjamins.

Comorovski, Ileana. 1996. *Interrogative phrases and the syntax-semantics interface*. Dordrecht: Kluwer Academic Publishers.

Delahunty, Gerald. 1981. *Topics in the syntax and semantics of English cleft sentences*. Ph.D. dissertation, University of California, Irvine.

Delahunty, Gerald. 1984. The analysis of English cleft sentences. *Linguistic Analysis* 13, 33–113.

Delin, Judy. 1995. Presupposition and shared knowledge in *it*-clefts. *Language & Cognitive Processes* 10.2, 97–120.

Delin, Judy L. 1989. *Cleft constructions in discourse*. Ph.D. dissertation, University of Edinburgh, Edinburgh.

den Dikken, Marcel. 2013. Predication and specification in the syntax of cleft sentences. In Hartmann & Veenstra (eds.), 35–70.

Dryer, Matthew S. 1996. Focus, pragmatic presupposition, and activated propositions. *Journal of Pragmatics* 26, 475–523.

Engdahl, Elisabet. 2006. Information packaging in questions. In Olivier Bonami & Patricia Cabredo Hofherr (eds.), *Empirical issues in syntax and semantics*, vol. 6, 93–111. Paris: Colloque de Syntaxe et Sémantique à Paris.

Faarlund, Jan Terje, Svein Lie & Kjell Ivar Vannebo. 1997. *Norsk referansegrammatikk*. Oslo: Universitetsförlaget.

Frascarelli, Mara & Francesca Ramaglia. 2013. (pseudo)clefts at the syntax-prosody-discourse interface. In Hartmann & Veenstra (eds.), 97–138.

Groenendijk, J. & M. Stokhof. 1982. Semantic analysis of *wh*-complements. *Linguistics and Philosophy* 5, 175–233.

Groenendijk, J. & M. Stokhof. 1984. *Studies on the semantics of questions and the pragmatics of answers*. Ph.D. dissertation, University of Amsterdam, Amsterdam.

Guerzoni, Elena. 2003. *Why even ask? On the pragmatics of questions and the semantics of answers*. Ph.D. dissertation, MIT, Cambridge, MA.

Gundel, Jeanette K. 2002. *It*-clefts in English and Norwegian. In Hilde Hasselgård, Bergljot Behrens, Cathrine Fabricius-Hansen & Stig Johansson (eds.), *Information structure in a cross-linguistic perspective*, 175–196. Amsterdam: Rodopi.

Haegeman, Liliane, André Meinunger & Aleksandra Vercauteren. 2014. Syntax and semantics of *It*-clefts. A tree-adjoining grammar analysis. *Journal of Linguistics* 50.2, 269–296.

Haida, Andreas. 2008. The indefiniteness and focusing of question words. In Andreas Haida, T. Friedman & S. Ito (eds.), *SALT XVIII*, 376–392. Ithaca, NY: Cornell University.

Halvorsen, P-K. 1978. *The syntax and semantics of cleft sentences*. Ph.D. dissertation, University of Texas at Austin, Austin.

Hamblin, C. L. 1973. Questions in Montague grammar. *Foundations of Langauge* 10, 42–53.

Han, Chung-hye. 2002. Interpreting interrogatives as rhetorical questions. *Lingua* 112, 201–229.

Han, Chung-hye & Maribel Romero. 2004. On negative yes/no questions. *Linguistics & Philosophy* 27, 609–658.

Hansen, Erik & Lars Heltoft. 2011. *Grammatik over det danske sprog*. Copenhagen: Det danske sprog- och litteraturselskab.

Hartmann, Katharina & Tonjes Veenstra (eds.). 2013. *Cleft structures*. Amsterdam: John Benjamins.

Hedberg, Nancy. 1990. *Discourse pragmatics and cleft sentences in English*. Ph.D. dissertation, University of Minnesota.

Hedberg, Nancy. 2000. The referential status of clefts. *Language* 76, 891–920.

Hedberg, Nancy & Lorna Fadden. 2007. The information structure of *It*-clefts, *Wh*-clefts and reverse *Wh*-clefts in English. In Nancy Hedberg & Ron Zacharski (eds.), *The grammar-pragmatics interface: Essays in honor of Jeanette K. Gundel*, 49–76. Amsterdam/Philadelphia: John Benjamins.

Heim, Irene. 1983. On the projection problem for presuppositions. In M. Barlow, D. Flickinger & M. Westcoat (eds.), *Second Annual West Coast Conference on Formal Linguistics*, 114–126. Stanford, CA: Stanford University.

Higginbotham, James & Robert May. 1980. Questions, quantifiers, and crossing. *The Linguistic Review*.1, 41–80.

Huber, Stefan. 2002. Es-*clefts und* det-*clefts. Zur Syntax, Semantik und Informationsstruktur von Spaltsätzen im Deutschen und im Schwedischen.* Ph.D. dissertation, Lund University, Lund.

Jackendoff, Ray S. 1972. *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.

Jespersen, Otto. 1927. *A modern English grammar on historical principles*, vol. Part III. Heidelberg: Carl Winters Universitätsbuchhandlung.

Johansson, Mats. 2001. It and wh clefts in English and Swedish. *Linguistics* 39.3, 547–582.

Johansson, Mats. 2002. *Clefts in English and Swedish: A contrastive study of IT-clefts and WH-clefts in original texts and translations*. Ph.D. dissertation, Lund University, Lund.

Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics & Philosophy* 1.1, 3–44.

Katz, J. J. & P. M. Postal. 1964. *An integrated theory of linguistic descriptions*. Cambridge, MA: MIT Press.

Kenesei, Istvan. 1986. On the logic of word order in Hungarian. In A. Abraham & S. de Mey (eds.), *Topic, focus and configurationality*, 143–159. Amsterdam: John Benjamins.

Kennan, Edward L. & Robert D. Hull. 1973. The logical presuppositions of questions and answers. In J. S. Petöfi & D. Franck (eds.), *Präsuppositionen in Philosophie und Linguistik*, 441–466. Frankfurt am Main: GFR: Athenäum.

Krifka, Manfred. 1995. The semantics and pragmatics of polarity items. *Linguistic Analysis* 25.3-4, 209–257.

Lambrecht, Knud. 1994. *Information Structure and Sentence Form. Topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press.

Lambrecht, Knud. 2001. A framework for the analysis of cleft construction. *Linguistics* 39, 463–516.

Levinson, Stephen. 2012. Interrogative intimations: On a possible social economics of interrogatives. In Jan de Ruiter (ed.), *Questions: Formal, functional and interactional perspectives*, 11–32. New York: Cambridge University Press.

Lewis, David. 1979. Scorekeeping in a language game. *Journal of Philosophical Logic* 8, 339–359.

Lie, Svein. 1978. Cleft wh-questions in Norwegian and their presuppositions. In Kirsten Gregersen, Hans Basbøll & Jacob Mey (eds.), *The Fourth Scandinavian Conference of Linguistics*, 62–77. Odense: Odense University Press.

Mayr, Clemens. 2013. Downward monotonicity in questions. In Emmanuel Chemla, Vincent Homer & Gregoire Winterstein (eds.), *Proceedings of Sinn und Bedeutung 17*, 345–362. Paris: ENS.

Meinunger, André. 1998. A monoclausal structure for (pseudo-)cleft sentences. In N. Tamanji & Kiyomi Kusumoto (eds.), *Proceedings of NELS 28*, 283–398. Amherst, MA: GLSA.

Molnár, Valéria. 2016. Questions in focus – focus in questions. Presentation at the 22nd Germanic Linguistics Annual Conference, University of Iceland, May 21, 2016.

Nicolae, Andreea. 2015. Questions with NPIs. *Natural Language Semantics* 23, 21–76.

Percus, O. 1997. Prying open the cleft. In K. Kusumoto (ed.), *Proceedings of NELS 22*, 337–351. Amherst, MA: GLSA.

Postal, Paul M. 1971. *Cross-Over Phenomena*. New York: Holt, Rinehart, and Winston, Inc.

Prince, Ellen F. 1978. A comparison of wh-clefts and it-clefts in discourse. *Language* 54.4, 883–906.

Progovac, Ljiljana. 1994. *Negative and positive polarity*. Cambridge: Cambridge University Press.

Reeve, Matthew. 2000. *Clefts*. Ph.D. dissertation, University College London, London.

Reeve, Matthew. 2011. The syntactic structure of English clefts. *Lingua* 121, 142–171.

Reeve, Matthew. 2013. The cleft pronoun and cleft clause in English. In Hartmann & Veenstra (eds.), 165–186.

van Rooy, Robert. 2003. Negative polarity items in questions: Strength as relevance. *Journal of Semantics* 20.3, 239–273.

Siemund, Peter. 2012. Interrogative clauses in English and the social economics of questions. *Journal of Pragmatics* 119, 15–32.

Svenonius, Peter. 1998. Clefts in Scandinavian. *ZAS Papers in Linguistics* 10, 163–190.

Søfteland, Åshild. 2014. *Utbrytingskonstruksjonen i norsk spontantale*. Ph.D. dissertation, University of Oslo, Oslo.

Tailleur, Sandrine. 2013. *The French wh interrogative system:* est-ce que, *clefting*? Ph.D. dissertation, University of Toronto, Toronto.

Teleman, Ulf, Staffan Hellberg & Erik Andersson. 1999. *Svenska Akademiens grammatik*. Stockholm: Norstedts Ordbok.

Tomioka, S. 2009. *Why* questions, presuppositions, and intervention effects. *Journal of East Asian Linguistics* 18, 253–271.

Vercauteren, Aleksandra. 2016. *A conspiracy theory for clefts: The syntax and interpretation of cleft constructions*. Ph.D. dissertation, Ghent University, Gent.

É. Kiss, Katalin. 1998. Identificational focus versus informational focus. *Language* 74, 245–273.

É. Kiss, Katalin. 1999. The English cleft construction as a focus phrase. In L. Mereu (ed.), *Boundaries of morphology and syntax*, 217–229. Amsterdam: John Benjamins.

Author's address: *Stockholm University,*
Universitetsvägen 10D, 106 91 Stockholm, Sweden
johan.brandtler@su.se