The syntax and pragmatics
of embedded *yes/no* questions

Regine Eckardt

Abstract

The paper investigates the distribution of English *if* and *whether* as complementizers for polar questions in extensional contexts, and specifically the restriction of *if*-questions to negative contexts. I will criticize the recent analysis by Adger and Quer (2001) and suggest that markedness of embedded *if-* questions in certain contexts arises through the conspiracy of several factors. I follow Bolinger (1979) in assuming that the question "*if S*" presupposes an (ir-)relevance asymmetry between the proposition denoted by *S* and its negation. I argue that this leads to competition between embedded *if*-questions and *that*-clauses in the relevant constellations, which in turn excludes the *if*-question as the less optimal variant.

1   Embedded questions

In English, like in many other languages, questions can occur as subordinate clauses under a matrix verb. The matrix verb can be one of questioning, doubting or wondering, like in example (1). The embedded clause here denotes a question and reports that the subject poses or could pose the question, mentally or explicitly, to himself or to others.

    (1)   Susan wondered/asked/did not know who was the murderer.
           Susan wondered/asked/did not know whether Jones was the murderer.

This kind of example, also called *extensional* question embedding, contrasts semantically with examples where the subject holds a certain mental attitude towards the correct answers of the embedded question, like in (2):

    (2)   Elsie knew/told me/discovered who took the book.
           Elsie knew/told/disclosed whether Jones took the book.

These predicates have also been called *intensional* question embeddings. Finer classifications of question embedding verbs have been proposed in the literature (Ginzburg 1996, Lahiri 2002 Ginzburg and Sag 2000), starting from seminal work by Karttunen (1978). The present paper is focused on embedded yes/no questions in intensional contexts, and more specifically on the distribution of the complementizers *if* and *whether* in such embeddings. I will argue that these complementizers carry different presuppositions and that intensional question embeddings compete with embedded *that* clauses. It will be shown that the distribution of *if* and *whether* in positive and negated question embedding contexts can be derived from pragmatic factors. This counters a recent proposal by Adger and Quer (2001) who offer a feature-based analyses of the relevant data.

The paper is organized as follows: In section 2, I will briefly review the relevant examples of embedded *yes/no* questions and the puzzling preferences for complementizing *if* and *whether* reported by native speakers. Section 3 is devoted to an exposition and critique of Adger and Quer's (2001) analysis of these data. In section 4, I propose that the presuppositions of *whether* and *if*, as well as competition between embedded *if* and *that*-clauses, should be used as the basic ingredients for a pragmatic account of the data. This claim will be exemplified on basis of the verb *admit* that, in some sense, constitutes a prototypical embedding verb for the phenomena under debate. In section 5, I will show how this basic account can be generalized to other question embedding verbs. The final section is devoted to the discussion of the notoriously shaky nature of speakers' judgments in this particular area. Side remarks to that end pervade not only the literature (see the very balanced survey in Huddleston and Pullum 2002: 973ff.), they are likewise issued at several places within the present paper. I will argue that we witness a case of ongoing language change in this area, where presuppositions that were a stable part of the lexicon at earlier stages are presently weakened and neglected by speakers, depending on their idiolectal lexicon. This observation lends further support to the proposed analysis in contrast to one which is anchored in the core syntax: While Adger and Quer have little to say about idiolectal variation apart from the remark that they only intend to describe their own, concise judgments about the data, a pragmatic account can link variation to general processes in language change.

As a kind of prelude, let me point out a certain pragmatic asymmetry between embedded constituent questions and *yes/no* questions. Consider a sentence like (3).

(3)   Elsie knew / told me who took the book.

Such a sentence will be unmarked and felicitous, no matter whether the speaker knows the answer to the question '*who took the book?*' or not. The example with the verb *told* even strongly suggests that I know the answer. A sentence like (4), in contrast, sounds somewhat irksome in a situation where the speaker already knows whether there will be a talk on Monday. The context in brackets explicates this situation.

(4)   (We cannot meet on Monday because there will be a talk at 2 pm. – Should I also tell
        Elsie? This is unnecessary, ...)
        Elsie knows whether there will be a talk on Monday.

If we judge the well-formedness of sentence in (4) in isolation, it is clearly grammatical and matches the approved pattern in (2). It is likewise not contradictory or semantically marked in the context in (4). However, informants who are confronted with discourse like in (4) usually show a strong dispreference for the pattern even to the claim that (4) is grammatically ill-formed.[1] Upon closer reflection, they react by saying that an embedded *that*-clause would be the preferred option in this context. Interestingly, embedded constituent questions do not seem to evoke this kind of reaction easily. The discourse in (5) does not suffer from the same markedness as (4).

(5)   (We now got all responses to our invitation to the Christmas party. Tom will come, Sue and Linda, and Peter, Paul and Mary will come. – Should I also tell Elsie? No, that's unnecessary.…)
Elsie already knows who will come to the party.
I already told Elsie who will come to the party.

The contrast in (4) suggests that an embedded *yes/no* question potentially has to compete with the corresponding embedded *that*-clause. Speakers perceive the *that*-clause as a vital alternative to an embedded question in intensional contexts, and the embedded question appears to be the marked, or more complex alternative. If this more complex way of putting things is to be used, the speaker must have good reason to do so. This ties in with comparable analyses for the use of lexical pairs like *kill/cause to die* that have received revived interest in the development of bidirectional optimality theory (Blutner 2000, Jäger 2002).

As we can see from example (5), no similar competition seems to arise between embedded constituent questions and their answers. This observation does not necessarily contradict the proposed explanation of (4) but seems to highlight a structural difference between constituent questions and yes/no questions. The answer to a *yes/no* question consists necessarily of just the confirmation or negation of the proposition at issue. Hence, a *yes/no* question has necessarily always an equally brief answer. Constituent questions, however, are structurally tied to a potential list of answers, as it is structurally possible that more than one instantiation of the *wh*-phrase lead to true propositions. (Note that we are concerned with *structural* possibilities here. Of course, the meaning of the question can *logically* imply that there can be at most one positive answer.) We can therefore assume that, in the search for the optimal realization of a given embedded proposition, our grammar requires us to weigh *yes/no*- questions against *that*-clauses, but *not* constituent questions against their answers—even if that answer be known. Cases like these can help us to come to a better understanding of the range of competing structures in OT accounts, a notoriously debated issue in the development of this kind of approach.

I will, however, restrict my attention to a limited application of these observations and assumptions here. It will be proposed that markedness effects like in (4)—which arise in a completely context-driven and productive manner in that example—have turned into part of the grammar of certain question embedding verbs (of German and English), and of a certain type of embedded question (*if* in contrast to *whether*) in English. If the proposed analysis is found to be tenable, it will open a new range of mechanisms in which semantics and pragmatics can tie in with the syntax of a language.

2   Embedded yes/no questions: *whether* or *if*

In a recent article, Adger and Quer (2001) discuss a peculiar contrast between English *if*- and *whether*-questions in certain embedded contexts. In particular, they point out that embedded *if*-questions are marked in certain contexts and appear to behave like negative polarity items. The effect is illustrated by the examples in (6) and (7) and particularly by (6.d) and (7.b). In view of the fact that speaker judgments vary widely in this area, I offer their original examples in (6), and (9) to (11).

(6)   a.   The bartender told me that I was drunk.
          b.   The bartender told me who was drunk.
          c.   The bartender told me whether I was drunk.
          d.   *The bartender told me if I was drunk.

(7)   a.   Anton admitted whether he was the heir of Churchill.
          b.   *Anton admitted if he was the heir of Churchill.

According to Adger and Quer, this contrast appears under intensional question embedding verbs like *admit, tell, hear,* or *say.* The authors hypothesize that the relevant class of verbs comprises all and exactly those embedding verbs that can also alternatively embed a *that*-clause.

Matters would be easy if these data could simply be explained as a selection restriction for complementizers, shown by certain verbs. However, further data preclude this kind of explanation. Remarkably, the examples improve in nonveridical contexts. The authors discuss the following examples, resting their argument on the wide notion of nonveridicality proposed by Giannakidou (1998).

(8)   a.   *Anton admitted if he was drunk.
          b.   Anton did not admit if he was drunk.
          c.   Nobody admitted if they were drunk.

(9)   Only Julie admitted/heard/told me if the bartender was happy.
(10) We refused to admit/tell if they had the keys.
(11) Without them admitting if they have the keys, there is nothing we can do for them.

Moreover, the authors sense an improvement if the constructions are used in non-episodic contexts like under modal operators.

(12) Anton would always admit if he was drunk.[2]

The data in (6) to (12) are in fact a challenge. One cannot assume that the embedding predicates simply to not allow complement *if* clauses. According to the authors' diagnosis, the class of contexts in which embedded *if-clauses* are licensed coincides with the class of non-veridical contexts. Given that non-veridical contexts are also those that license polarity sensitive items, it seems natural to exploit this licensing relation in an account of the data in (6) to (12). In essence, the authors propose the following analysis.

1. Unacceptable *if*-clauses occur always as arguments of verbs that also (really?) embed *that*-sentences. Hence, semantically speaking, the argument of the verb must be a proposition rather than a question.
2. Adger and Quer assume that all *wh*-questions in English (including *whether* questions) can be reinterpreted as propositions, namely as the, or one among the true answer(s) to the question. In this reinterpreted sense, questions are always in principle possible arguments for verbs that take propositions as arguments.
3. English *if*-questions can not be reinterpreted as propositions.

4. They propose to use a tacit type shifting operator Δ to make question denotations combinable with a proposition-taking verb. Adger and Quer propose that this should be seen as a kind of generalized quantifier, and discuss possible morphological reflexes of this operator in other languages (though not in English or German). This operator performs on an *if*-question what standard type shifting can achieve for *wh*-questions, namely its coercion into a proposition.[3]
5. In addition, this operator Δ happens to be polarity sensitive.
6. Hence, if Δ occurs in a sentence—in order to semantically adjust an embedded question to the embedding predicate—and is not licensed by a nonveridical context, then the sentence is judged ungrammatical.

This proposal can indeed predict the intriguing interplay between embedded question, embedding predicate, and non-veridical contexts. However, if we evaluate it against a broader background certain questions arise. These will be presented and discussed in the next section.

Before going on, a clarifying statement with respect to the data seems in place. Interestingly, informants report varying reactions to the crucial examples. Usually, hearers find Adger and Quer's starred sentences "odd" or "marked" but not downright "ungrammatical". The authors faithfully report such reactions but ascribe them to the fact that speakers of English might show idiolectal variation here. While it is hard to discover an informant who fully shares Adger and Quer's idiolect, all native speakers that I consulted could agree on markedness effects. These include an estimated total of approximately 20 natives at two workshop presentations, plus two native BE informants who were willing to endure prolonged evaluations of examples. It might turn out as a plus of the pragmatic account that it operates in terms of markedness and competition rather than strict ungrammaticality in a syntactic sense.

3   Open issues in the syntactic account

The first objection to Adger and Quer's proposal is motivated by the general pragmatic nature of polarity sensitive items. Even though work in recent years has made it clear that different languages possess different sets of polarity sensitive items (Giannakidou 1998 and subsequent, Zwarts 1986, 1998 and others), certain characteristics of NPIs remain constant. In particular, it seems clear that the lexical property of being *polarity sensitive* in practically all cases arises through a semantically/pragmatically motivated phase where a certain type of expression is used in order to make an emphatic contribution to discourse. The contrast between (13) and (14) can serve to illustrate the difference. Note that I deliberately refrain from using the more common *Joe doesn't own any money* because the complementary English determiners *some* and *any* are a grammaticalized version of a formerly pragmatic contrast.

(13) Joe has no money.
(14) Joe doesn't own a red cent.

Pragmatic accounts of polarity sensitivity in the tradition of Fauconnier (Fauconnier 1975, Krifka 1995, Lahiri 1998, see also Eckardt 2004a:chap.4, 2005) rest on the primary intuition that polarity sensitive items are used to make emphatic statements. For example, *'not owning a red*

cent' is the most dramatic or extreme property that the subject *'Joe'* can have with respect to a scale of different degrees of being poor. In this and many other cases, the emphatic value of the NPI is part of our current understanding of the word or phrase. In some cases, NPIs have lost the strong rhetorical quality but retain the distributional restrictions (like English *any* in competition to *some*, or German *je* = 'ever' in contrast to *einmal* = 'some time'.) However, so far no NPIs in English have been discovered where this rhetorical quality would be completely absent.

The analysis of Adger and Quer implies the existence of invisible NPIs. This claim contradicts everything that we know about the genesis of polarity sensitive items at present. In the initial stage, it is vital for the element to be able to carry stress in order to give rise to an emphatic statement. It is unclear how a phonologically null element could enter this development.

Perhaps we could assume that *if* might have functioned as the exponent of the emphatic quality of the statement somehow. But once again, nothing in the historical development of embedded *if* questions supports the claim that an emphatic quality ever played a role in the use of this construction. This is witnessed, for instance, by the detailed discussion of question patterns in Old and Middle English that are offered in The Cambridge History of the English Language (1992; Vol 1 (Hogg) and 2 (Blake). The respective authors on syntax, Traugott and Fischer, offer detailed accounts of the use of question words in OE and ME. Neither, however, reports a particular rhetoric quality of *if* questions. This coheres with the examples and uses that are listed in the Oxford English Dictionary. Whatever differences may guide the use of *if* contra *whether*, there seem to be no indications whatsoever that *if*-clauses should be an emphatic variant of *whether*-clauses. Let me illustrate what that would amount to: In contrasting (15) and (16), no scholar ever described variants like (16) as synonymous to *Joe did not <u>admit in any sense</u> whether he was the heir of Beethoven* or *Joe did not admit whether he was the heir of Beethoven <u>in any sense whatsoever</u>*.

(15) Joe didn't admit whether he was the heir of Beethoven.
(16) Joe didn't admit if he was the heir of Beethoven.

Of course, this objection can only affect Adger and Quer's analysis to the extent that we believe in the pragmatic account of polarity sensitivity. It is not an objection against their synchronic implementation. However, there has been an increasing effort in recent years to achieve a contentful understanding of the phenomenon of polarity sensitivity, an explanation of the licensing conditions in terms of an item's pragmatic lexical properties rather than as a curious allergy to positive contexts. Adger and Quer's proposal certainly runs counter to this strong trend in semantic research.

The second objection again does not affect the implementation of Adger and Quer's proposal for English. However, it casts doubt on the analysis from the viewpoint of language comparison. The analysis can not be extended to similar constellations that can be found in German, particularly with the embedding verb *sich sicher sein, ob...* ('to be sure'). The data are listed in (17) and (18).

(17) sich sicher sein, ob

a. *?*Peter ist sich sicher, ob er den Herd ausgeschaltet hat.*
   Peter is himself sure 'ob' he the stove off.switched has
   'Peter is sure if he switched off the stove'

b. **Alle sind sich sicher, ob Pisa in Italien liegt*
   all are themselves sure 'ob' Pisa in Italy lies
   'All are sure if Pisa lies in Italy'

c. **Ich bin mir sicher, ob Johann kommt*
   I am myself sure 'ob' Johann comes
   'I am sure if Johann will come'

(18) a. *Peter ist sich nicht sicher, ob er den Herd ausgeschaltet hat.*
        Peter is himself not sure 'ob' he the stove off-switched has
        'Peter is not sure if he switched off the stove'

     b. *Ich bin mir selten sicher, ob ich den Herd ausgeschaltet habe*
        I am myself rarely sure 'ob' I the stove off-switched have
        'I am rarely sure if I switched off the stove'

     c. *Wenig Jugendliche sind sich mit 18 sicher ob sie eine*
        Few teens are themselves with 18 sure 'ob' they a
        eigene Familie gründen wollen.
        own family found want
        'Few teens are sure at the age of 18 if they want to found a family'

General facts about German *ob* preclude an analysis of the contrasts in (17) and (18) along the lines of Adger and Quer's theory. German *ob* usually behaves like English *whether*. In particular, it can occur under intensional question embedding verbs, as shown in (19). Note that the German example patterns with the English *whether* sentence, and not with the English *if* example. Hence, embedded *ob*-questions should be reinterpretable as propositions without further operators.[4]

(19) *Otto weiß jetzt, ob er Krebs hat.*
     Otto knows now 'ob' he cancer has
(20) Now, Otto knows whether he has cancer.
(21) *Now, Otto knows if he has cancer.

In view of these facts, the markedness of the examples in (17) can not be captured by recurring to a tacit generalized quantifier which performs the type shift from question to proposition, and which is in addition polarity sensitive.

A final objection concerns Adger and Quer's proposal as part of a coherent overall theory of embedded questions. In earlier accounts of question embedding verbs, it was usually assumed that the reinterpretation of a question as a proposition occurs as part of the verb's lexical meaning. Propositional attitude verbs like *know* or *determine* have a basic use as propositional attitude verbs and a use as an operator on question where the question is reinterpreted as a proposition. Propositional attitude verbs that do not allow for this reinterpretation can only embed declarative sentences (This is essentially the line followed by Lahiri, 1998).

Adger and Quer, in contrast, assume that *wh*-questions can always be reinterpreted as propositions. However, we clearly find verbs that do not accept embedded questions; *believe* is a

prominent example. In Adger and Quer's account, this restriction can not be semantic in nature, because all questions except *if*-questions allow for a type shift. Therefore, it must be regulated in syntactic terms, for instance by structurally allowing or refuting an argument with *wh* features.

The feature selection properties of a verb normally do not change when the verb enters into complex constructions. Embeddability of questions, however, does. Some examples are listed in (22).

(22) a. *Ich vermute, wer morgen kommt.*
I   assume who tomorrow come
'I assume who will come tomorrow'
Hans hat vermutet, ob Pisa in Italien liegt.
Hans has assumed whether Pisa in Italy   lies
'Hans has assumed whether Pisa is in Italy'

(22) b. *Ich kann nur vermuten, wer morgen kommt.*
I   can only assume/guess who tomorrow come
'I can only guess who will come tomorrow'

(22) c. *Ich kann lediglich / höchstens / allein vermuten, wer morgen kommt.*
I   can only, at best, just   guess   who tomorrow comes
'I can only guess who will come tomorrow'

The embedded questions in (22.a) are ungrammatical and seem to indicate that '*vermuten*' does not license a *wh*-complement. The sentences in (22.b) and (22.c), however, are fully acceptable. A *syntactic* licensing mechanism that derives the potential to license *wh*-features from complex predicates in a compositional manner remains as yet to be developed.

In view of these objections to Adger and Quer's proposal, it seems reasonable to consider possible alternatives. In the next sections, I would like to lay out the elements of a pragmatically based account for the data under consideration.

4. *If* or *whether*, the simple case

The pragmatic account for the different uses of embedded *if-* and *whether*-questions will rest on the competition between embedded assertions and questions that was illustrated at the beginning. I propose that embedded *if*-questions interact with the meaning of embedding verbs in such a way as to be equivalent to an embedded assertion. The embedded assertion is then preferred as the simpler, or less marked way to make the same statement and consequently, the embedded *if*-sentence is judged as marked. In the present section, I will spell out this general idea for a basic case.

The first issue to resolve is the nature of the difference between an *if*-question and a *whether*-question. For expository reasons, I will assume that the meaning of a question is represented by the set of possible answers, i.e. use Hamblin's analysis (Hamblin 1973). The argument is compatible with more refined treatments of question semantics (e.g. Groenendijk and Stokhof, 1984, Krifka 2001 or Ginzburg and Sag, 2002). For *yes/no*-questions this amounts to the following representation:

(23) $[\![if/whether\ S]\!] = \{\ [\![S]\!],\ \neg[\![S]\!]\ \}$

One might wonder why the English language affords two words for a syntactic function as simple as a question complementizer. In fact, there seems to be no officially acknowledged difference between the two, and speakers' strategies to chose between them seem to vary to a certain degree. Nevertheless, the following trend is supported by the data and was described as early as Bolinger (1978).

(24) The Bias of *if*: The use of an embedded *if* S presupposes that the two possible answers are not of equal status for future actions, opinions, plans, etc. Commonly, the positive case *S* is the relevant case for some decision, opinion, etc., the case that really changes something. The negative case is perceived as the 'otherwise' case with no, or no interesting consequences.

(25) The Non-Bias of *whether*: The use of an embedded *whether* S does not presuppose any asymmetric interests between the positive case 'S' and the negative case 'non S'. The speaker assumes a balanced interest in either one of the two possible answers.

It might be important to point out here that the bias of *if* is not one in terms of the expected answer. All interlocutors can be truly ignorant about the answer of an embedded *if* question. They do not hold a bias about the likely answers, but the two possible answers are not equally relevant. Before illustrating this effect with some examples, let me state the presuppositions of an embedded *if*-question in these terms:

(26) $[\![if\ S]\!] = \{\ [\![S]\!],\ \neg[\![S]\!]\ \}$
Presupposition: RELEVANT($[\![S]\!]$).

The notion of being 'relevant' is an unspecific cover term that will have to be instantiated in each context of use. It is understood that not both *p* and non-*p* can be 'relevant' in this sense. Some first examples are listed in (27) and (28).

(27) a. The captain needs to know your name and whether you tend to get seasick.
b. The captain needs to know your name and if you tend to get seasick.

Question (27.a) at the beginning of a cruise seems to suggest that the captain has an interest in both types of passengers, those who get seasick as well as those who do not. This might be because staff collects statistical data about the percentage of passengers who need paper bags. Question (27.b) appears to be more likely in a context where the staff already has prepared means to defend seasickness and wants to contact affected passengers. The question suggests that passengers who are "good sailors" are not the core interest of those who maintain the question.

Note that both embedded questions are conjoined with a nominal argument of the matrix verb. This ensures that the *if*-clause is actually interpreted as a complement of the verb and not as a postposed conditional.[5]

Similar asymmetric expectations arise for embedded positive and negative questions.

(28) a. We must find out who's the first speaker and also if he needs a beamer.

b. We must find out who's the first speaker, and also if he doesn't need a beamer.

c. We must find out whether the first speaker needs a beamer (or not).

The sentence in (28.a) is more likely in a context where we only start racing for a beamer as soon as we learn that the first speaker will need one. The sentence in (28.b) suggests a different situation: We seem to have a beamer and need to find out whether the first speaker can spare it. The last variant reports neutral interest in the speaker's need for a beamer without preconceptions.

In fact, it is highly interesting to take a closer look into the actual use that speakers of English make of *whether* questions. Native speakers report that they are very frequently used with the postponed disjunct '*or not*' like indicated in brackets in example (28.c). The same tendency can also easily be confirmed in the medium which is at once the richest and least reliable current corpus-like structure for English, the internet. Although one should not rest a study on such findings alone, superficial evaluation strongly suggests that in written sources in the internet, speakers of English use *whether*-questions with the '*or not*' tag much more frequently than they do with *if*-questions (where it can only be found rarely). Let me offer some impressionistic numbers here: Doing web search for the terms *if* plus *"or not"* and picking an arbitrary choice of 50 hits, I got a ratio of 18 *whether-or not* constrictions to 6 *if-or not* constructions (in addition to 24 other uses, and 2 double sites). I will come back to the issue of language *use* in contrast to current grammar in section six.

A final example to illustrate the bias of *if*-questions can be found in Bolinger (Bolinger, 1978).

(29) a. Please tell me if you want to marry me.
b. Please tell me whether you want to marry me.

Bolinger notes that the variant in (29.a) seems to be the more successful strategy while (29.b), inappropriately, has a strangely disinterested ring to it. The proposal in (24) can explain this difference. Under normal circumstances, the second party in (29) will expect that a positive answer is the most relevant alternative for the speaker; joy, happiness, wedding preparations and a major change in life being some of the expectable consequences. A negative answer, in contrast, may have some emotional impact but is not normally the starting point for further activities. The question as put in (29.b) does not presuppose this asymmetry. The speaker seems to envisage two equally well-planned ways for future events for both the negative and the positive answer. This is a sober strategy, but disclosing it is not very encouraging for the addressee in (29.b).

Minimal pairs at the demarcation line between the use of *if* and *whether* are not easy to construct. Nevertheless, where differences do appear, they seem to point in the direction suggested in (24) and (25). The distinction is likewise in accord with the origin of '*if*' (thought the reader will realize that the same source for German *ob* has not led to any bias in current use). Hence, I will use (24)(25) as baseline for the next steps.

We can now investigate the interaction of embedding predicates with the bias of *if*. Let me start by discussing one of the most stable and uncontroversial among Adger and Quer's embedding

verbs, *admit*. It gives rise to several presuppositions, most prominently it is veridical: The embedded proposition is presupposed to be true. However, it is another presupposition that seems to be of particular interest for our case: Sentences like '*A admitted that p*' presuppose that the proposition *p* that was admitted was worth to be kept secret. This *secrecy* presupposition excludes many contents as possible objects of '*admitting*' and, in particular, will in most cases lead to the constellation that among a proposition *p* and its negation ¬*p*, only one can be '*admitted*' while the other needs to be '*told*' or '*announced*'.

(27) a. Joe admitted that he had made mistakes in the exam.
       b. Joe *?admitted/ announced that he hadn't made any mistakes.

Of course, this effect depends on the given context. What can be *secret* for one person in one context can be *not secret* for other people under different circumstances. There are also circumstances where a writer wants to withhold any information about a given topic, as in typical front page headlines *"Chancellor Schröder finally admits whether hair was dyed"*. The default case, however consists in *p* being worth to be hidden iff its opposite ¬*p* is not, as illustrated in (27). Our assumptions about *admit* are summarized in (28).

(28) ADMIT(*a, p*)
       Presupposition: *p* is true, EMBARRASSING(*p*)
       Default assumption: EMBARRASSING(*p*) → ¬EMBARRASSING(¬*p*)

What happens if the bias of *admit* meets the bias of an embedded *if*-question? Remember that the bias of *if* contains an open parameter that needs to be specified in context. There is an asymmetry in RELEVANCE, but what counts as RELEVANT needs to be decided in context. I assume that under these circumstances, RELEVANCE is instantiated by the expectation that some EMBARRASSING fact is at stake. Something is relevant-qua-being embarrassing. In sum, the following lexical information is combined in the derivation:

(29) [[*if S*]] = { [[S]], [[¬S]] }
       Presupposition: RELEVANT([[*S*]])
       Default assumption: RELEVANT(*p*) → ¬RELEVANT(¬*p*)

(30) [[admit, if S]]

   1. The lexical meaning of the embedding predicate allows for a reinterpretation of the semantics of the embedded question as a proposition.
       [[A *admit, if S*]] iff ADMIT(*a, p*) for the true answer *p* to the question ,*if S*'.
   2. Only embarrassing things can be admitted: The presupposition of ,*admit*' requires that this *p* must be EMBARRASSING(*p*)
   3. Instantiation of the presupposition of '*if S*': *S* is relevant qua being embarrassing.
   4. The only possible way to verify these different side requirements consists in the following instantiation: The true answer must be *p* = [[*S*]] because the admitted fact must be embarrassing (by 'admit') and relevant (by 'if').
   5. In summary: [[*a admits if S*]] holds true if and only if [[*a admits that S*]]

6. Due to competition between synonymous phrases, *a admits **if** S* is regularly suppressed by the more optimal (less complex, and more specific) alternative *a admits **that** S*.

The last step follows the recent trend to explain certain grammaticality judgments in terms of competition between different candidates. One prominent example in the domain of semantics/pragmatics is Bidirectional Optimality Theory (Blutner 2000, Jäger 2002) which was successfully applied to explain pairs like *kill, cause to die*. In the present case, we face a similar form of competition. I suggest that the lexical meaning of *if,* together with the lexical meaning of embedding verbs like *admit*, yield the net effect that an embedded *if*-question in veridical contexts can only convey the same meaning as an embedded *that*-clause. (The negative case will be discussed presently.) Under these circumstances, the use of an embedded question is more costly than an embedded assertion, because the proposition *that S* needs to be derived as the positive answer to the question *if S*. The case mirrors the markedness effects that were diagnosed at the beginning of the paper for embedded constituent questions.

How can this proposal explain the effect, that *admit* can embed an *if*-sentence in a nonveridical context? Consider once again a sentence like in (31).

(31) The queen did not admit if her hair was dyed.

The crucial difference arises already at the very beginning of the derivation when the embedded question is reinterpreted as proposition. The combination of the question meaning with the negated embedding predicate conveys something like in (32):

(32) For all true answers p to the question *'if the queen's hair was dyed'* it was not the case that ADMIT(queen,p).

Importantly note that the negated sentence in (31) does not just negate that the queen admitted something about her hair. As we know, '*admit*' is veridical, and hence the mere negation of a positive sentence will retain the presupposition that the embedded clause is true. But the truth of the embedded clause is exactly what remains open in a sentence like (31): (31) does not make any claims about the hair color of the queen.

The presuppositions that are carried by various parts of the sentence can still be unified in the manner described above, because *true* answers and *relevant* answers to a question are not the same. We may still assume that the proposition that would be worth admitting, or relevant-qua-embarrassing, is denoted by the sentence under *if*. Yet, (31) states that no admissions took place at all. Hence, an embarrassing proposition might have been withheld, or it might not have been 'admitted' simply because it was not true. Sentence (31) is open in this respect and consequently not synonymous with the positive statement (33).

(33) The queen did not admit that her hair was dyed.

Therefore, no logical equivalence between (31) and (34) arises. But if the two sentences do not convey the same message, then they will not compete.

The slight contrast between (31) and (34) predicted by this account should consist in that (34) is more neutral as to whether it is more embarrassing to dye, or not to dye one's hair. This made-up example is very nicely backed up by real uses of the collocation 'admit whether' like in the sentences in (35) and (36).

(34) The queen did not admit whether (or not) her hair was dyed.
(35) Most of them were afraid to **admit whether** or not it was a successful rape or whether they had ever been victims.
(36) And how would we do that? Read a porno magazine and **admit whether** we were turned on or not, (...)

These were again harvested from the web. Search for the collocation *'admit whether'* indeed mostly yields hits where *whether* does not introduce a subordinate clause. Approximately one out of 20 cases show in fact subordination, and these mostly report some symmetric constellation like the two above. I tried to complement these wild data by search in the online version of the BNC, but the collocations *"admit whether"* and *"admitted whether"* yield 0 hits in this corpus.
In contrast, embedded *if*-questions under *admit* do not compete with embedded *that*-sentences when it is left open (at least for some persons, some time, some possible world) whether there was anything to admit or not. Like other uses of *if*-questions, the speaker conveys an asymmetry in relevance between the positive and the negative answer. Embedded *if*-questions rest on the notion of relevant-qua-embarrassing also in nonveridical contexts. The following sentences show the typical and most frequent type of use:[6]

(37) "you are both courageous enough to admit <u>if mistakes were made</u>"
      "I'm the first to admit <u>if I'm wrong</u>"
      "I would admit <u>if I'd contemplated suicide</u> (which I have), but..."

Observe that, counter to the expectations of Adger and Quer, real examples like these do not exhibit typical NPI contexts. If we follow the competition model proposed here, we expect that (37) are perfect uses of embedded *if*, because they all leave it open that *no mistakes might have been made*, that *I might not be wrong* or that *I might not contemplate suicide.* The affinity of embedded *if* for non-veridical contexts hence shows a superficial similarity to NPI licensing but is an essentially different phenomenon.

The proposal as it stands so far works well for matrix verbs with specific presuppositions about their complement. It can not so far explain why verbs like *know* or *tell* should give rise to similar markedness effects. I will turn to these in the next section.

5   The general case

In the present section, I will address the behavior of *if*-questions under verbs like *know* or *tell*. For the purpose of illustration, let me use *tell* as the standard example. I will not, in the present section, be concerned with data judgments or with the exact boundaries of the verb class to be investigated. However, I will return to these issues in the last section. As for now, I want to

model data that follow the Adger and Quer *scheme*, i.e. examples like the following and specifically the markedness of sentence (38.c):

(38) a. The queen told Edward that her hair was dyed.
     b. The queen told Edward whether her hair was dyed.
     c. *The queen told Edward if her hair was dyed.
     d. The queen did not tell Edward if her hair was dyed.

In the last section, we assumed that an *if S* question carries a bias. The positive proposition ⟦ S ⟧ is perceived as the RELEVANT case in context, whereas the negative proposition ⟦¬S ⟧ is of minor or no interest. We moreover assumed that the notion of RELEVANT is to be seen as a cover term that needs to be specified in a given utterance.

I propose that the specification of what is RELEVANT is in part driven by the embedding verb, in sentences like (38.c) and (38.d). There is a conventionalized unification of the meaning of the matrix verb with the notion of relevance as a presupposition of the *if*-question. It is exemplified in (39), and (40) offers the general scheme.

(39) A tells (if S)
      RELEVANT in the given context amounts to: BE-WORTH-TELLING. Per default, it is expected that only one of a proposition and its negation will be WORTH-TELLING.
(40) A <verbs> (if S)
      RELEVANT in the given context amounts to: BE-WORTH-<VERBING>

Note the clear parallel between this assumption and the analysis of *if*-questions as complements of *admit* in the last section. While the previous case could be motivated by the lexical meaning of *admit*, the move in (40) might strike the reader as circular. In the final section, this conventionalized instantiation of RELEVANT will be motivated by the origin of the construction from conditionals. For the moment, however, let us recapitulate how the assumption in (39) can explain the data in (38). The steps in (41) closely follow the derivation in (30).

(41) ⟦A*tell, if S* ⟧
      1. The embedded question is reinterpreted as a proposition:
          ⟦A*tell, if S* ⟧ is true iff TELL(*a, p*) for the true answer *p* in ⟦ *if S* ⟧.
      2. The presupposition of ,*if S*' consists of RELEVANT(⟦*S*⟧) and by default, ¬RELEVANT(⟦¬*S*⟧)
      3. Instantiation of the presupposition of '*if S*': Being RELEVANT amounts to BE-WORTH-TELLING.
      4. It follows that TELL(*a, p*) for the true answer *p* in ⟦ *if S* ⟧ only occurs if *p* is worth telling and hence, the true answer *p* in ⟦ *if S* ⟧ must have been ⟦*S*⟧.
      5. In summary: ⟦*a tell if S* ⟧ holds true if and only if ⟦*a tell that S* ⟧
      6. Due to competition between synonymous phrases, *a tell **if** S* is regularly suppressed by the more optimal (less complex, and more specific) alternative *a tell **that** S*.

In analogy to the previous example, we can show that the embedded *if*-question does not compete with the *that*-clause exactly if there is at least one instance in the domain of quantification of the sentence where the *if*-question does not amount to the same proposition as the *that*-clause. This is so if no telling takes place and if, at the same time, it is open whether it does not occur due to the fact that *S* is not true, or because the given subject chose not to tell about the truth of *S*.

What is the general difference between an embedded *whether* question in contrast to an embedded *if* question? The present analysis predicts that embedded *whether*-questions are more balanced and that the act of *telling* (*guessing, deciding, revealing*, ...) will take place also if the negative proposition ¬S is true. Consider the sentences in (42) and (43).

> (42) Anna may decide whether there will be a party.
> (43) Anna may decide if there will be a party.

The predicted slight difference should be this. Sentence (43) should be appropriate in a situation where the normal course of events will be one without a party. So, the real decision of Anna is one *in favor* of a party whereas there is nothing to decide in order to have none. Sentence (42) should report that Anna has to make a volitional decision *pro* or *contra* a party. There might be some people who favor the idea that a party should take place, and hence it is not clear what the course of events will be if Anna fails to make a decision.

If native speakers of English are interviewed about their use of minimal pairs like in (42)/(43) and also in (38.b/c), they usually report unclear intuitions. On the one hand, it can usually not be excluded that Anna, instead of remaining stubbornly tacit, will say 'no party' at some point even though the situation was described with the sentence in (43). On the other hand, a non-event can be epistemically as explicit as an explicit denial. As a consequence, sentences like (42) seem to sound natural even if Anna in fact just *does* remain silent (e.g. *fails* to send out invitations, *fails* to post an announcement, etc.) in the 'no party' case. The only testable intuition (at least for some speakers) seems to be that episodic past or present tense positive reports about acts of *telling, announcing, deciding, knowing* etc. need to use a *whether* question as complement.

Before closing the section, I will take a brief look at alternative questions. One surprising kind of pattern emerges when we look at alternative questions embedded with *if* and *whether* respectively. The disjunction 'or not' in spontaneous written language use occurs in the overwhelming majority of cases in embedded *whether* questions. Typical examples are like in (44).

> (44) "The focus of the Asia Pacific Hospice Conference (...) centered on **whether or not** physicians should inform cancer patients of their diagnosis."
> "You will be expected to decide after this **whether or not** you wish to join the group."
> "When the Prison Service is deciding **whether or not** to grant temporary release to a prisoner, it will (...)"
> "The applicant will be told if a charge will apply, so they can decide **whether** to proceed with the request **or not**."

We also find alternative embedded *if*-questions. This, at first sight, seems to pose a challenge to the analysis that I present here. If we postulate that only the positive case *S* of *if S* can ever be relevant, how could anyone possibly use an alternative *if* question? Curiously, speakers prefer different patterns to build alternative questions embedded with *if*. Typical examples are listed in (45).

(45)   "You're then told if the number you've keyed in is **greater or smaller** than the mysterious number."
"Once you have done that you will then be told if the number you picked is **wrong or right.**"
"You will be told if the item is **available or in use**."

In all these cases, the disjuncts are both presented with a positive description. The speaker views them as two different, but equally relevant possibilities. Only against the background of world and lexical knowledge, the negative answer to the (first) positive *if S$_1$* question is logically equivalent to the positive complementary case (' ... *or S$_2$*'). Such circumstances, then, seem to be appropriate to overwrite the default assumption in (23) that only one among *p* and not-*p* can be relevant.
Without aiming at a full analysis of sentences with the *'or not'* coda, I submit that the present analysis offers a balanced interaction between hard constraints and soft parameters that can capture borderline phenomena like the ones quoted. Therefore, examples like in (45) support rather than falsify the present analyses. Adger and Quer mention the role of the *'or not'* coda in a footnote and state that their account offers no explanation.

6. The wider perspective
In the present article, I took up Adger and Quer's observation about the restriction of embedded *if* questions to nonveridical contexts. After criticizing their earlier proposal, I developed an account which rests on the following basic ideas.

- Embedded *yes-no* questions compete with embedded *that*-clauses. Only those embedded questions are licensed that lead to a sentence that is in some sense better than the competing embedded assertion.
- An embedded assertion is preferred to an embedded polar question if it is clear in context that the speaker already knows the coorect answer to the question and could hence be more informative, and if no further communicative factors sanction withheld information (e.g. newspaper headline context).
- *if*-questions presuppose the bias that the positive case *S* is relevant whereas the negative case ¬*S* is not. Relevance is viewed as an open parameter that will be instantiated in specific utterance situations.
- Under verbs of questioning, the choice of what counts as relevant is left to the non-linguistic context. In intensional embedding contexts, however, there is a conventionalized interaction with the meaning of the embedding verb.
- Specifically, the bias of *if* is instantiated by an expectation about what is worth to be <verb>-ed (e.g. *told, remembered, admitted, revealed*). The restriction turns the sentence

equivalent to a sentence with an embedded *that*-clause in veridical episodic contexts. This competition leads to markedness of the embedded *if*-clause.
- No similar bias arises for *whether*-questions which explains their wider distribution.

I would like to end with some considerations with respect to language development. First note that the analysis that I proposed in section five comes very close to the meaning of an *if*-sentence as a postponed conditional clause. This can be illustrated with the example in (46).

(46) *Sue will tell us if she is hungry.*
(47) a. *If she is hungry, Sue will tell us* (so).
     b. *Sue will tell us if she is hungry (and what she wants to eat).*

The two variants in (47) disambiguate (46), and it is generally assumed that *if* as question complementizer arose by reanalysis of sentences like (46). If we follow propositional logic and interpret *if-then* as material implication, the structure in (47.a) conveys a proposition that can be paraphrased as follows:

(48) In case *'Sue is hungry'* is true at some time *t*, then she will tell us that she is hungry. In the case that *'Sue is hungry'* is false at some time *t*, no predictions are made. (It is strongly suggested, though, that Sue will not falsely claim that she is hungry *then*).

In view of the fact that the conditional is the source structure for complement *if* clauses, it is not surprising that the latter still carry the expectation that the positive case is the 'relevant' case. The task of tracing this development in the data would be rewarding but is clearly beyond the scope of this paper. However, such considerations can give us a hint as to where the current use of *if* questions may be located on a developmental scale, and why speakers' judgments might be variant.

At the onset of this paper, I pointed out a general necessity for embedded *yes/no* questions to compete with embedded *that*-sentences. It is interesting to see that the actual use of embedded *whether* is severely shaped by this necessity as well. Of course, there are uses where the author/speaker wants to keep some proposition secret for good reason, like in the headline use in (49).

(49) Charles finally reveals whether he will marry Camilla. (Read more on page...)

Such uses, however, are rarer than one might expect, and an overwhelming majority of '*<verb> whether'* constructions occur in non-veridical contexts. Consider, for example, the following episodic use of *whether*. In fact, it amounts to quantification over cases.

(50) ... He looked at the logs and the structures. He told me whether it was worth working on....

The situation reported is one where an expert and a potential buyer inspect an old log house and work shed to see whether reconstructions are feasible. The diagnosis that is summarized here as *"whether it was worth working on"* (with a positive answer, by the way) must have consisted of a series of local evaluations of the building's parts and possible improvements. Therefore, and in

order to describe the evaluation situation from a viewpoint where the result was still open, the writer can use embedded *whether*.

This seems to lead to the situation that language learners don't get enough evidence any more to form a tenable hypothesis about the current lexical content of *if* in contrast to *whether*. Moreover, *if* is the shorter one of two competing, highly frequent function words. One would therefore expect that the differences between the two items are blurred as the shorter item is on the rise. One indication for an ongoing development are examples of *if*-complements with an 'or not' coda. Although they are rare in spontaneous data, they can be produced (see also Adger and Quer 2001, ftn. 2). Hence, we find examples like the following:

(52) "You will never be **told if** your answers are correct **or not**."
"You can fill in your card details and within a minute you will be **told if** the payment was successful **or not**."

To give some impression of the frequency of such uses, I can offer the following figures: In a search for "*<verb> if*" + "*or not*", roughly one hit out of three will in fact contain '*or not*' as a coda of some '*whether'* clause. Hence, the '*or not'* coda in spontaneous writing is still strongly tied to *whether*. Only approximately one hit out of fifty provides an alternative *if* question. (The other hits contain "*<verb> if*" and *or not* in different sentences, or in a different kind of construction.) I interpret these findings as follows: Speakers still sense the bias requirement of *if*, but it appears to be demoted from a strict pragmatic requirement to a mere pattern of usage. Speakers also see that the bias requirement can be neutralized if an alternative *if* question is used, originally with a positive description of the complementary alternative.[7] Having lost sense for the fact that the *positive* description of the negated case was once crucial, speakers use standard negation as *the* pattern in order neutralize the bias of an *if* question.

In sum, we can speculate that the current preference for *whether* in veridical, non-quantificational contexts is a remnant pattern from an older but vanishing, more specific difference between *if* and *whether*. The data may be unclear from the point of view of synchronic linguistic theory. From a diachronic perspective, the phenomenon deserves deeper investigation. Speakers of English seem to be in the transient phase between two clearly defined kinds of use of embedded *yes/no* questions. Further stages remain to be awaited.

References

Adger, David and Josep Quer. 2001. "The syntax and semantics of unselected embedded questions". *Language* **77** (1): 107-133.
Blake, Norman (ed.) 1992. *The Cambridge History of the English Language.* Vol. II. Cambridge: Cambridge University Press.
Blutner, Reinhard. 2000. "Some aspects of optimality in natural language interpretation". *Journal of Semantics* 17: 189-216.
Bolinger, Dwight. 1978. "Yes/No questions are not alternative questions". In: *Questions*. Hiz, Henry (ed.), 87-105. Dordrecht: Reidel.
Eckardt, Regine. 2004. *Meaning Change under Reanalysis*. Habilitationsschrift, HU Berlin. To appear at Oxford University Press.

Eckardt, Regine. 2005. "Too poor to mention: Subminimal Events and Negative Polarity." In: *Event Arguments: Foundations and Applications*. Claudia Maienborn and Angelika Wöllstein-Leisten (eds.), 301- 331. Tübingen, Niemeyer.

Fauconnier, Gilles. 1975. "Pragmatic scales and logical structures". *Linguistic Inquiry* 6: 353-375.

Giannakidou, Anastasia. 1998. *Polarity Sensitivity as (Non)Veridical Dependency*. Amsterdam and Philadelphia: John Benjamins.

Ginzburg, Jonathan. 1996. "Interrogatives: Questions, Facts and Dialogue." In: *The Handbook of Contemporary Semantic Theory*. Shalom Lappin (ed.), 385-458. Malden, Blackwell.

Ginzburg, Jonathan and Ivan A. Sag. 2000. *Interrogative Investigations*. Stanford: CSLI Publications.

Groenendijk, Jeroen and Martin Stokhof. 1984. *Studies on the semantics of questions and the Pragmatics of Answers*. PhD dissertation, University of Amsterdam.

Hamblin, C.L. 1973. "Questions in Montague English". *Foundations of Language* 10:41-53.

Hogg, Richard M. (ed.) 1992. *The Cambridge History of the English Language*. Vol. I. Cambridge: Cambridge University Press.

Huddleston, Rodney, Geoffrey L. Pullum (2002): *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Jäger, Gerhard. 2002. "Some notes on the formal properties of bidirectional Optimality Theory". *Journal of Language, Logic and Interpretation* 11(4): 427-451.

Krifka, Manfred. 1995. "The semantics and pragmatics of polarity items". *Linguistic Analysis* 25: 209-257.

Krifka, Manfred. 2001. "For a structured account of questions and answers". In: *Audiatur Vox Sapientia. A Festschrift for Arnim von Stechow*. C. Fery & W. Sternefeld (eds.), 287-319. Berlin: Akademie Verlag (studia grammatica 52).

Lahiri, Utpal. 1998. "Focus and Negative Polarity in Hindi". *Natural Language Semantics* 6: 57-123.

Lahiri, Utpal. 2002. *Questions in Embedded Contexts*. Boston: MIT Press.

Zwarts, Frans. 1986. *Categoriale grammatica en algebraïsche semantiek: een onderzoek naar negatie en polariteit in het Nederlands*. Ph.D. diss., UGroningen.

---

[1] This was a colleagues immediate reaction.

[2] Generally, *if* clauses in such constructions are potentially ambiguous between an embedded sentence reading and a postposed antecedent of a conditional. Many starred examples in the text would be acceptable in the latter analysis. Note that this ambiguity can be avoided by choosing verbs like *admit* which obligatorily require a complement phrase.

[3] Note that from a semantic perspective, this move is in fact not tenable. Type shifting is an operation on semantic objects. As far as the authors tell us in their paper, the semantic objects denoted by *if*-questions and *whether*-questions are the same. It is unclear how type shifting can "sense" whether a semantic object has come about as the denotation of one, or the other, kind of syntactic structure, and can apply, or not apply, respectively.

[4] Note that the star in example (21) reflects Adger and Quer's abstract diagnosis rather than a real speaker judgment. In view of the high degree of individual variation here, it is unclear what status other speakers' intuitions (including, unfortunately, the reader's) should have for Adger and Quer's proposal. A personal evaluation by Adger and Quer would be the only

methodologically sound criterion to asset the grammaticality of (21)—and they can only confirm the *, or disclaim their own theory.

[5] I thank Philippe Schlenker who convinced me that this potential ambiguity can severely disturb the judgments.

[6] I want to briefly specify and defend my data acquisition strategies for the present paper. The collocations "verb + if" and "verb + whether" were searched in Google and got hits that range between ca. 800 and several hundred thousands for the constructions that are discussed here. Where too many interfering hits occurred, data were cut down by specifying pronoun subjects and objects. Hits were listed in batches of 30 per page. A pattern is called "typical" here when it occurred at least three or four times per 30 hits over at least five pages that were scanned, and if the opposing pattern – where there is one at stake – does not or occurs at best once in this range of hits. Web pages outside the UK or USA were not taken into account, nor poetry or writers with evidently nonstandard or non-native competence. This mode of data acquisition does not meet statistical requirements for real corpus search. However, the range of data comes as close to spontaneous spoken language as possible unless time and manpower can be invested in tape recordings and transcripts. The sheer quantity of data available online is notoriously tempting. The trends that I report here were so stable that I am confident they can be reproduced. What seems highly desirable, though, are qualified modes to take a sample corpus from the web in a manner that precludes manipulation or unwelcome statistical side effects.

7 A nice user's quote in this context is this: "... like in 'he told me whether he'd go to Paris or to Rome', which is in a sense two yes/no questions in one."