

# Hans-Christian Schmitz and Bernhard Fisseni

## Repairs for Reasoning\*

**Abstract:** We describe and experimentally investigate phenomena of *modal enrichment*, that is, phenomena in which a recipient *non-literally* interprets an utterance by creating and applying a modal operator. We give competing explanations for these phenomena – namely an explanation according to which modal enrichment is a repair procedure for making the utterance match a script of information processing vs. an explanation according to which modal enrichment is triggered by rhetorical structure.

---

## 1 Introduction

Let us consider the following example, discussed by Schmitz (2011):

(1) A: Wie spät ist es? (What time is it?)

B: Es ist 5 nach 3, meine Uhr geht aber 5 Minuten vor. (It's 5 past 3, but my watch is 5 minutes fast.)

When you ask test subjects what time it is according to B's answer, by far most of them will reply that it is (exactly) 3 o'clock, although this contradicts the literal meaning of the answer sentence. Schmitz (2011) performed such experiments several times. The results of one experiment are given in Table 1<sup>1</sup> – the data proved to be reliable in repetition. In addition, Schmitz asked the subjects how they arrived at the *non-literal* interpretation, and they replied that they interpreted the answer in the sense of "By my watch it's 5 past 3, but my watch is 5 minutes fast" which entails that it is (exactly) 3 o'clock.

|         | 15:00      | 15:05     |
|---------|------------|-----------|
| exp. 1: | 36 (85.7%) | 6 (14.3%) |

**Table 1:** What time is it according to B's answer? (42 test subjects)

The example has not been made up. The answer was originally uttered by Benedikt Löwe (at that time a logician at the University of Bonn who did not intend to become a provider of linguistic examples). When asked what he meant, he

answered that he meant that it was exactly 3 o'clock. Thus, the test subjects grasped the intended meaning, they interpreted Benedikt's answer correctly.

How did they do so? Obviously, they expanded the literal meaning of the answer sentence by deriving a modal operator such as 'by the speaker's watch' and applying it to the first conjunct. Then they inferred the answer to A's question from the *non*-literal interpretation of the first conjunct and the literal interpretation of the second conjunct.

Schmitz (2011) calls meaning expansion by creation and application of a modal operator a form of enrichment, namely 'modal enrichment'. He argues that modal enrichment is a conventionalized operation of expanding the truth-conditional content of sentences and gives several examples to prove that modal enrichment can take place in communication. Moreover, he shows that under certain conditions it is possible that modal enrichment is licensed for the recipient although not intended by the speaker. In such a situation, a misunderstanding can be avoided if the speaker explicitly blocks the enrichment of his utterance. Schmitz (2011) argues that in German this blocking function is fulfilled by the adverb "tatsächlich" (*in fact*).

The present paper can be seen as a sequel to Schmitz (2011): instead of elaborating on the fact *that* modal enrichment occurs and the question of how it can be avoided, we will focus on the issue of *why* recipients perform operations of modal enrichment. We will experimentally investigate some phenomena in closer detail, in particular the initial watch-example (1), in order to expose under which conditions modal enrichment takes place and under which conditions it does not. The derived data will be the basis for explanations of the phenomena. In particular, we will discuss two approaches in closer detail and show how these can be evaluated. According to the first approach (*Reasoning*), the literal interpretation of (1) does not match a presupposed script of deriving and conveying information. It therefore has to be *repaired* by modal enrichment. According to the second approach (*Rhetorical Structure*), modal enrichment is not a repair procedure but rather a primary (close to hard-wired) way of interpreting information. We elaborate upon this approach with respect to current theories of polyphony.

The outline of the paper is as follows: in section 2 we will report on experimental investigations in modal enrichment in order to prove the reliability and strength of the effect and determine under which circumstances it can be expected. In section 3, we will investigate the impact of the conjunction "aber" (*but*) on modal enrichment. In section 4, we will elaborate upon the mentioned approaches for explaining modal enrichment and discuss their evaluation. Finally, in section 5, we will sum up our findings.

## 2 Modal enrichment

In this section, we will discuss experimental data regarding the interpretations of some key examples, in particular variations of the initial watch-example (1).

### 2.1 Experimental settings

All our experiments took place in a classroom setting: the examples were projected onto a wall and read aloud twice. The test subjects were then asked a question – after the presentation of example (1), e.g., they were asked what time it is according to B's answer. They noted down their answers on paper. In most experiments, we did not predefine a list of answers to choose from, that is, the questions were not multiple choice questions (exceptions will be discussed below). The experiments took place between 2005 and 2009 at the Universities of Bonn, Frankfurt am Main and Duisburg-Essen. The test subjects were mainly undergraduate students in computational linguistics, linguistics or German studies. It can be assumed that all of them were naïve regarding the issues investigated, meaning that they were not linguistically biased. All test subjects were native speakers of German, except in the first experiment mentioned in the Introduction. In the experiments, we investigated the interpretation of German sentences in dialogue contexts. In the present paper we also provide English translations of the examples, although these were not provided during the experiments. We assume that the observed phenomena of modal enrichment also occur with the English translations; the experimental data, however, are only valid for the German examples.

### 2.2 Data

We start with the original watch-example (1). Schmitz (2011) tested this example several times with different test groups: in all experiments the vast majority of subjects modally enriched the answer sentence. In the original example, “*aber*” (*but*) appeared as an adverb. The tendency towards a *non*-literal interpretation persisted if “*aber*” was changed into a conjunction between the two answer clauses (example 2a), if it was replaced with “*und*” (*and*, example (2b)) or if it was completely left out (example (2c)). The results of respective experiments are given in Table 2.<sup>2</sup>

|                             | 15:00      | 15:05     | 15:10    | '?' |
|-----------------------------|------------|-----------|----------|-----|
| exp. 2a (48 test subjects): | 41 (85.4%) | 6 (12.5%) | 1 (2.1%) |     |
| exp. 2b (22 test subjects): | 18 (81.8%) | 3 (13.6%) | 1 (4.6%) |     |
| exp. 2c (48 test subjects): | 39 (81.2%) | 8 (16.6%) | 1 (2.1%) |     |

**Table 2:** What time is it according to B's answer?

(2) A: Wie spät ist es? (What time is it?)

- a. B: Es ist 5 nach 3, aber meine Uhr geht 5 Minuten vor. (It's 5 past 3, but my watch is 5 minutes fast.)
- b. B: Es ist 5 nach 3, und meine Uhr geht 5 Minuten vor. (It's 5 past 3, and my watch is 5 minutes fast.)
- c. B: Es ist 5 nach 3. Meine Uhr geht 5 Minuten vor. (It's 5 past 3. My watch is 5 minutes fast.)

These experiments show, firstly, that the data are reliable and, secondly, that modal enrichment is not an effect of the junctor “aber” (*but*) alone. There is no significant difference in the interpretations of the examples (2a–2c). Nevertheless, we think that the occurrence of “aber” might still contribute to modal enrichment. We will discuss this, empirically not well-founded, hypothesis in section 3.

The test subjects derived a *non*-literal interpretation of B's answer. The question arises whether we have to assume the existence of a conventionalized operation of modal enrichment to properly derive this interpretation or whether we can just assume that the *non*-literal interpretation is the result of some non-conventionalized ‘modulation’ of the answer meaning. The idea behind modulation, as we understand it here, is that a given utterance is somehow – it does not matter much how – made compliant with conversational maxims.<sup>3</sup> Of course, a conventionalized operation can serve to make an utterance compliant with conversational maxims, too. However, it exists independently of the maxims, and it can also be applied without reaching compliance. Thus, the assumption of such an operation becomes plausible if it can be shown that test subjects reliably apply it even in the interpretation of utterances that cannot be made compliant with the conversational maxims.

To solve the issue, we let test subjects interpret the undoubtedly odd answers of examples (3).

(3) A: Wie spät ist es? (What time is it?)

- a. B: Es ist 5 nach 3. Meine Uhr geht 3 Stunden vor. (It's 5 past 3. My watch is 3 hours fast.)
- b. B: Es ist 5 nach 3. Meine Uhr geht 6 Stunden vor. (It's 5 past 3. My watch is 6 hours fast.)
- c. B: Es ist 5 nach 3. Meine Uhr geht 9 Stunden vor. (It's 5 past 3. My watch is 9 hours fast.)

Assuming that a conventionalized operation of meaning enrichment exists, we expect that the subjects apply this operation and derive the interpretation that it is in fact five past 12, five past 9 or five past 6, respectively (15:05–3/6/9 hours). We expect a clear tendency towards a *non*-literal interpretation in each example, although the answer is not made compliant with the conversational maxims. (It still seems anomalous even if interpreted *non*-literally.) If the *non*-literal interpretation was derived by free modulation we would not expect such a clear result. Since the answer cannot be made fully adequate, we would expect the test subjects not to show a clear tendency towards *non*-literal interpretation. Rather, we expected them to interpret the answers literally, to reply that they just could not properly interpret them or to derive different *non*-literal interpretations.

The results of the experiment are given in Table 3: as can be seen, there is a clear tendency towards the answers predicted by the hypothesis that modal enrichment is a conventionalized operation.<sup>4</sup>

Even ridiculous answers like that of example (4) are interpreted in a foreseeable manner:

(4) Dialog im Oktober 2007 (dialogue taking place in October 2007):

A: Was für ein Wochentag war der 11. Januar? (Which weekday was January 11th?)

B: Der 11. Januar war ein Mittwoch, aber mein Kalender ist von 2005. (January 11th was a Wednesday, but my calendar is from 2005.)

|          | -3/6/9 hours | +3/6/9 hours | literal (15:00) | other     |
|----------|--------------|--------------|-----------------|-----------|
| exp. 3a: | 14 (82.4%)   | 1 (5.9%)     | 1 (5.9%)        | 1 (5.9%)  |
| exp. 3b: | 14 (82.4%)   |              | 2 (11.8%)       | 1 (5.9%)  |
| exp. 3c: | 13 (76.5%)   |              | 1 (5.9%)        | 3 (17.6%) |

**Table 3:** What time is it according to B's answer? (17 test subjects)

We asked test subjects which weekday January 11th was according to B's answer.<sup>5</sup> (The experiment took place in July 2007. In fact, January 11th 2007 was a Thursday, January 11th 2005 was a Tuesday.) A number of subjects' first reaction upon being asked was to act amused, confused or show mild signs of aggression. Still, as can be seen from the results given in Table 4,<sup>6</sup> they interpreted the example in accordance with the hypothesis that they would apply an operation of modal enrichment. (If January 11th was a Wednesday in 2005 then it must have been a Friday in 2007: two years have 730 days, 730 modulo 7 (a week) is 2. Thus, two days must be added. Two days after a Wednesday is a Friday.)<sup>7</sup>

|        | - 2 days (Mon) | + 2 days (Fri) | literal (Wed) | other    | '?'       |
|--------|----------------|----------------|---------------|----------|-----------|
| exp. 4 | 7 (14.9%)      | 23 (48.9%)     | 8 (17.0%)     | 3 (6.4%) | 6 (12.8%) |

**Table 4:** Which weekday was Jan 11th according to B's answer? (47 subjects)

Both experiments show that subjects have an operation of modal enrichment available and that they can foreseeably apply it even in the interpretation of odd examples.

## 2.3 Discussion

Let us discuss the experiments and our interpretation of the results: an opponent might object that the results do not prove the existence of a conventionalized operation. It might still be that the subjects chose their interpretations by freely modulating the answer meanings in order to get the best interpretation possible. This objection, however, presupposes that there is a 'best' interpretation of the examples and, further, that the *non*-literal interpretation of (4), e.g., is 'better' than the literal one. We do not find this very convincing.

It can still be objected that examples (3) and (4) are far from natural. We would not expect such answers in real world situations, and if we were confronted with them we would rather ask what the speakers means than just 'understand' him. Most probably, the subjects reacted as if the experiments were meant as tests not of linguistic examples but their intelligence. (Of course, we told them there were no wrong answers and that we were not testing them but the examples. However, even if they believed us they may still have tried to be 'right'.)

While this is true, it does not affect our interpretation. Even if the subjects reacted like in an intelligence test, they proved that they can perform modal enrichment. It does not matter for our argument that they performed this opera-

tion in a non-natural setting since the original example (1) shows that they do not *only* perform it non-natural settings. (Reminder: the original example is a real world example.) The important point is that they performed this operation but they did not reach compliance with conversational maxims.<sup>8</sup>

Our opponent may concede that the original example is a real-world example, but he may insist that the experimental situation is not – not even for the first experiments. The interpretation of B's answer takes place in a lab situation, without reference to an action context in which the interpreter for some reason *has to* know what time it is. It might be that in such a context where the information on the time is crucial, B's answer is interpreted differently.

When we first heard the original answer we needed to know the time – that is, we interpreted the answer in a context as named by our opponent – and still we understood it properly. So, at least we were able to enrich the answer 'in the wild'. Nevertheless, it is a justified demand to test this with a broader group of subjects.

## 2.4 Robustness and commitment

To meet the critique of our potential opponent, we performed the following experiment: we gave sheets with the text of example (5) to test subjects. We asked them to read the text and choose one of the options O-1 to O-6. (The numbering of the options from O-1 to O-6 was not on the original sheets, we added it here to be able to refer to each option individually.)

- (5) Anweisung: Lesen Sie bitte folgenden Text. Bitte wählen Sie dann die Fortsetzung, die Ihres Erachtens am besten passt: „Sie sind auf dem Weg zum Bahnhof, um einen Zug um 15:07 zu nehmen und haben den Bahnhof beinahe erreicht. Dieser Zug ist, wie Sie wissen, stets pünktlich; er verspätet sich nie. Eigentlich haben Sie einen Zeitpuffer eingeplant, aber Sie sind ein wenig ins Träumen geraten und wissen nicht, wie viel Uhr es ist und ob Sie Ihren Zug schon verpasst haben. Daher sprechen Sie einen Passanten an, der Ihnen antwortet: „Es ist 10 nach 3, meine Uhr geht aber 5 Minuten vor.““

(Instruction: Please read the following text. After that, please choose the response which is, according to your intuition, most appropriate: “You are on your way to the train station to take a train at 3:07 and have nearly reached the station. The train is, as you know, always on time. It is never late. According to your plan, you should have a time buffer, but you started to daydream on the way and now, you are not sure what time it is and whether you have

already missed your train. Thus, you ask a passer-by, who answers: ‘It’s 3:10, but my watch is 5 minutes fast.’”

Was tun Sie nun? (What do you do now?)

- O-1: Sie bedanken sich bei dem freundlichen Passanten und gehen zügig los, damit Sie den Zug noch erreichen.  
(You thank the friendly passer-by and start to walk fast to reach the train.)
- O-2: Sie bedanken sich bei dem freundlichen Passanten; anschließend gehen sie fluchend nach Hause.  
(You thank the friendly passer-by; then, you go home cursing.)
- O-3: Bevor Sie zum Bahnhof gehen, versichern Sie sich noch einmal bei dem freundlichen Passanten, dass Sie noch genug Zeit haben.  
(Before you go to the station, you make sure that you have enough time, asking the friendly passer-by again.)
- O-4: Sie rüsten sich innerlich für den Nachhauseweg, versichern sich aber noch einmal bei dem freundlichen Passanten.  
(You prepare to go home, but ask the friendly passer-by again.)
- O-5: Sie fragen den freundlichen Passanten: „Entschuldigen Sie, wie viel Uhr ist es jetzt bitte tatsächlich?“  
(You ask the friendly passer-by: ‘I’m sorry, what time is it now, really?’)
- O-6: Sie fragen einen anderen Passanten, dessen Uhr vielleicht genauer geht.  
(You ask another passer-by, whose watch may be more exact.)

The six options the subjects had to choose from can be (partially) ordered on a scale:

1. We assume that a subject chooses O-1 if he chooses a *non*-literal interpretation of the answer due to modal enrichment.
2. Option O-3 demands a *non*-literal interpretation, too. However, we assume that the subjects who choose this option are not completely sure whether the *non*-literal interpretation is right.
3. We assume that the subjects who choose options O-5 and O-6 are indecisive about which interpretation is right.
4. Option O-4 demands a preference for the literal interpretation.
5. Finally, option O-2 demands a clear decision in favour of the literal interpretation.

We ran the experiment in two trials. The results of both trials are given in Table 5.<sup>9</sup> As can be seen, even in the action context a very clear majority of test subjects chose the interpretation derived by modal enrichment.<sup>10</sup>

|                    | O-1        | O-2      | O-3      | O-4 | O-5    | O-6 |
|--------------------|------------|----------|----------|-----|--------|-----|
| trial 1 (50 subj.) | 44 (88%)   |          | 2 (4%)   |     | 4 (8%) |     |
| trial 2 (30 subj.) | 29 (96.7%) | 1 (3.3%) |          |     |        |     |
| total (80 subj.)   | 73 (91.3%) | 1 (1.3%) | 2 (2.5%) |     | 4 (5%) |     |

**Table 5:** Answers of experiment 5 (two trials)

Let us take stock: the experiments show that recipients can perform operations of modal enrichment and that in our examples it can be foreseen that they do. Thus, a speaker who makes an utterance as in our examples can expect that his utterance will be interpreted *non-literally*. He can anticipate modal enrichment.

Is the speaker even committed to the *non-literal* interpretation of his utterance? To answer this question we performed a further experiment. We constructed a situation in which the standard watch-answer is given (as we know, it can be expected that it will be interpreted *non-literally*). Then it turns out that the *non-literal* interpretation of this answer is false while the literal interpretation is right. In the example, the recipient (Felix) interprets the answer *non-literally* and then complains that he was misled. We asked the test subjects whether he is right in complaining and, thus, whether the speaker (Benedikt) is committed to the *non-literal* meaning of his answer.

We presented the test-subjects with a two-sheet questionnaire. The text of example (6a) was on the first sheet. The subjects had to read the text and then interpret and note down what time it is according to Benedikt's answer. After that, they had to read the text of example (6b) that was on the second sheet and they had to decide whether Felix' complaint was justified. To this end, we gave them three options to choose from. (Again, the numbering of the options from O-1 to O-3 was not on the original sheets. We added it here to be able to refer to each option individually.)

- (6) a. Felix (F) und Benedikt (B) spazieren durch einen Park, der ungefähr fünf Minuten Fußweg vom Bahnhof entfernt ist, und unterhalten sich. Felix möchte einen Zug nehmen, der um Punkt 16:15 h vom Bahnhof abfährt. Der Zug ist stets pünktlich. Benedikt weiß davon.

F: Wie spät ist es?

B: Es ist 10 nach 4. Meine Uhr geht 10 Minuten nach.

(Felix (F) and Benedikt (B) are on a walk in a park, from which it is a walk of about five minutes to the station; they are having a conversation. Felix wants to take a train that starts at 4:15 sharp from the station. The train is always on time. Benedikt knows about this.)

F: What time is it?

B: It's 10 past 4. My watch is 10 minutes late.)

Wie spät ist es der Antwort zufolge? (What time is it according to Benedikt's answer?)

- b. Felix vermutet, den Zug nicht mehr erreichen zu können, und beschließt (durchaus verärgert) einen späteren Zug zu nehmen. Nach 5 Minuten kommen Felix und Benedikt an einer Uhr vorbei, die 16:15 h anzeigt. Felix beschwert sich bei Benedikt.

F: Was redest Du denn für einen Quatsch! Es ist ja erst 4 Uhr 15!

B: Warum? Ich habe vorhin gesagt, es sei 10 nach 4.

(Felix assumes that he will not be able to reach the train and decides (quite annoyed) to take a later train. Ten minutes later, Felix and Benedikt pass a clock, which shows 4:15. Felix complains to Benedikt.)

F: You are talking complete nonsense! It's only a quarter past 4!

B: Why? I've said that it's 10 past 4.)

Beschwert sich Felix bei Benedikt zu Recht? (Is Felix justified in complaining to Benedikt?)

- O-1: Benedikt hat Felix getäuscht. Felix musste ihn so verstehen, dass es bereits 20 nach 4 war. Felix ärgert sich zu Recht. (Benedikt has misled Felix. Felix had to understand Benedikt's words to mean that it was already 20 past 4. Felix is justified in being angry.)
- O-2: Benedikts Äußerung war vielleicht missverständlich, aber durchaus richtig. Felix hätte nachfragen können. (Benedikt's statement might have been mistakable, but was in principle right. Felix could have asked.)
- O-3: Benedikt hat die Uhrzeit korrekt genannt. Dass Felix ihn falsch verstanden hat, ist Felix' Fehler.

(Benedikt gave the time correctly. It is Felix's mistake that he misunderstood him.)

As in the previous experiment, the options can be ordered on a scale, namely from O-1 (misleading, fraud) via O-2 (mistakable, but in principle right) to O-3 (acceptable, correct).

The results are given in Table 6. In each column, the numbers of subjects who chose a certain interpretation are given: 20% of the subjects interpreted Benedikt's answer literally, 80% (74% + 6%) chose a *non*-literal interpretation. The 16:00-interpretations (6%) might have been due to modal enrichment and a miscalculation, that is, an error.<sup>11</sup> In the rows, the numbers of choices for one of the options O-1 to O-3 are given: according to 50% of the subjects, Benedikt was committed to the *non*-literal interpretation of his answer (O-1), only 4% found the answer completely acceptable (O-3). The second percentage within the brackets refers to the ratio of the subjects who evaluated Felix' complaint in the respective way (row) and the subjects who chose the respective interpretation (column) – e.g., 67.6% of those who chose the *non*-literal interpretation 16:20h (column) found that Felix was right in complaining to Benedikt (row: O-1).

|       | 16:00         | 16:10          | 16:20           | total    |
|-------|---------------|----------------|-----------------|----------|
| O-1   |               |                | 25 (50%, 67.6%) | 25 (50%) |
| O-2   | 1 (2%, 33,3%) | 10 (20%, 100%) | 12 (24%, 32.4%) | 23 (46%) |
| O-3   | 2 (4%, 66.7%) |                |                 | 2 (4%)   |
| total | 3 (6%)        | 10 (20%)       | 37 (74%)        |          |

**Table 6:** experiment 6 (50 test subjects)

Two thirds of those who chose the foreseeable *non*-literal interpretation 16:20 agreed with the complaint to Benedikt and made him responsible for the misunderstanding; and even those subjects who interpreted the answer literally found it mistakable. We are not surprised that those who interpreted the answer as meaning that it was 16:00 – most probably due to a calculation error – refrained from complaining (or agreeing to a complaint). The data show that under circumstances a speaker not only *can* foresee a *non*-literal interpretation but *must* foresee it and accommodate his utterance accordingly. He can be accused of non-cooperativeness if he does not; it is not just on the recipient to assure that the right *non*-literal interpretation is chosen.

## 2.5 Enrichment beyond the watch

Until now, we have shown very clear data on modal enrichment for the intial watch-example and variations of it. One might suspect that this is a singular example insofar that only this example or examples of this very particular type demand modal enrichment. To prove that this is not the case, Schmitz (2011) has already provided examples. Further examples for modal enrichment, not reported so far, include (7), among others. (In contrast to the intial watch-examples, these are not real-world examples, but have been made up.)

- (7) a. A: Was wird Peter wohl sagen, wann er kommt? (When does Peter say he will arrive?)  
 B: Er kommt um 3 Uhr, er verspätet sich aber aus Prinzip immer um eine Stunde. (He will arrive at three o'clock, but as a matter of principle he will be one hour late.)
- b. A: Wann kommt eigentlich Peter? (By the way, when will Peter arrive?)  
 B: Er kommt um 3 Uhr, er verspätet sich wie immer aber um eine Stunde. (He will arrive at three o'clock, but as always he will be one hour late.)

We presented these example to test subjects and asked them when they thought Peter would arrive. The results are given in Table 7: the clear majority expected Peter to arrive at 4 o'clock, although literally understood, B says that he will arrive at 3 o'clock.<sup>12</sup> We asked subjects how they arrived at their expectations. Those who expected Peter to arrive at 4 o'clock replied that they interpreted B's answers in the sense of 'Peter will say that he will arrive at 3 o'clock, but as a matter of principle/as always he will be one hour late'. They concluded that in fact Peter will arrive at 4 o'clock. That is, they performed an operation of modal enrichment by deriving a modal operator 'Peter said that' and applying it to the first answer clause.

There is a peculiarity in example (7a): A asks for the content of Peter's utterance rather than for his time of arrival. Conventionally, however, the latter is the underlying purpose of asking the question.<sup>13</sup> And, as (7b) shows, we can also omit the reference to speaking with Peter and virtually get the same interpretations.

|                        | 15 h      | 16 h       |
|------------------------|-----------|------------|
| exp. 7a (17 subjects): | 2 (11.8%) | 15 (88.2%) |
| exp. 7b (10 subjects)  | 3 (30%)   | 7 (70%)    |

**Table 7:** When do you think Peter will be here?

A final example that proves that various forms of modal enrichment – that is, enrichment with various form of modal operators – are possible, is (8). We did not collect experimental data on the interpretation of this example but leave it to the intuitions of the reader. (*We would expect to pay 5 €.*)

(8) A: Wieviel kostet das Buch? (How much is the book?)

B: 10 €, aber es gibt 50 % Rabatt. (10 €, but there is a discount of 50%.)

## 2.6 Examples without enrichment

We have seen that subjects perform operations of meaning enrichment with various examples. The following experiment raises the issue whether modal enrichment is really a kind of ‘repair’, that is, an operation that is not applied by default but only if something goes wrong:

(9) A: Wie spät ist es? (What time is it?)

B: *points to a clock according to which it is 15:10h.* Es ist 5 nach 3, meine Uhr geht aber 5 Minuten vor. (It’s 5 past 3, but my clock<sup>14</sup> is 5 minutes fast.)

Example (9) is identical to the initial watch-example except that we add that B points to a clock that displays 15:10. We drew a picture of such a clock on a blackboard and pointed to it while reading the example to the test subjects. Again the subjects had to indicate what time it is according to B’s answer. We now expected that most subjects would interpret the answer literally and take its second clause to be an explanation why B’s answer differs from the display of the clock. The results are given in Table 8: in fact, most of the subjects interpreted the answer literally, but also nearly half of them did not (though, only one test person came to the interpretation expected if modal enrichment were demanded). That is, the results are not very clear, and it should make us wonder why so many test subjects still chose a *non*-literal interpretation.

|         | 15:00    | 15:05     | 15:10     | 15:15    | ‘?’       |
|---------|----------|-----------|-----------|----------|-----------|
| exp. 9: | 1 (7.1%) | 8 (57.1%) | 2 (14.3%) | 1 (7.1%) | 2 (14.3%) |

**Table 8:** What time is it according to B’s answer? (14 test subjects and native speakers of German)

Nevertheless, modal enrichment does not ‘always’ take place. Schmitz (2011) proves that, firstly, the replacement of the conjunctions – so far: “aber” (*but*) or “und” (*and*) – can change the interpretations significantly: the vast majority of test subjects interpreted example (10a) literally, and subjects were undecided about the interpretation of example (10b) – about half of them interpreted it literally and about half expanded its meaning. Secondly, a strong accent on “ist” (*is*) like in example (10c) triggers the literal interpretation. Some might explain this as an effect of a so-called ‘verum-focus’. Thirdly, by inserting “tatsächlich” (*in fact*) modal enrichment is blocked, so that example (10d) is interpreted literally.

- (10) a. Es ist 5 nach 3, obwohl meine Uhr Minuten vorgeht. (It’s 5 past 3, although my watch is 5 minutes fast.)
- b. Es ist 5 nach 3, denn meine Uhr geht 5 Minuten vor. (It’s 5 past 3, because my watch is 5 minutes fast.)
- c. Es *IST* 5 nach 3, meine Uhr geht aber 5 Minuten vor. (It *IS* 5 past 3, but my watch is 5 minutes fast.)
- d. Tatsächlich ist es 5 nach 3, meine Uhr geht aber 5 Minuten vor. (In fact, it’s 5 past 3, but my watch is 5 minutes fast.)

The fourth – and for our concern here: most interesting – way of avoiding modal enrichment is to change the order of the answer clauses. Example (11a) is a modification of the initial watch-example which we already discussed above (cf. (2c)); example (11b) is derived from (11a) by changing the order of the answer clauses:

- (11) A: Wie spät ist es? (What time is it?)
- a. B: Es ist 5 nach 3. Meine Uhr geht 5 Minuten vor. (It’s 5 past 3. My watch is 5 minutes fast.)
- b. B: Meine Uhr geht 5 Minuten vor. Es ist 5 nach 3. (My watch is 5 minutes fast. It’s 5 past 3.)

We performed an experiment by letting two groups of test subjects interpret the examples and note down what time it is according to the answers provided. (Each test group interpreted only one example.) The results are given in Table 9: while example (11a) was, as we have already seen, mainly interpreted *non-literally*, (11b) was interpreted literally by most subjects. The results exhibit a significant

difference between the examples which can be traced back to the order of the answer clauses – the two-sided Fisher-test for exact data yields a p-value of  $< 0.001$ .<sup>15</sup>

|                         | 15:00      | 15:05      | 15:10    |
|-------------------------|------------|------------|----------|
| exp. 11a (48 subjects): | 39 (81.3%) | 8 (16.7%)  | 1 (2.1%) |
| exp. 11b (22 subjects): | 7 (31.8%)  | 15 (68.2%) |          |

**Table 9:** What time is it according to B's answer?

The results of this watch-experiment can be replicated with the Peter-example (7a). The example is repeated below as (12a), (12b) is a modification with the answer clauses in opposite order. Again, we let two different test groups interpret the examples and note down when, they think, Peter will arrive. (The order of answer clauses is not the only difference between the examples: (12a) contains “aber” while (12b) does not. This, however, should not affect the results, since we have already shown with example (7b) that even without “aber” we can expect the answer of (12a) to be modally enriched.)

- (12) A: Was wird Peter wohl sagen, wann er kommt? (When does Peter say he will arrive?)
- a. B: Er kommt um 3 Uhr, er verspätet sich aber aus Prinzip immer um eine Stunde. (He will arrive at three o'clock, but as a matter of principle he will be one hour late.)
- b. B: Er verspätet sich aus Prinzip immer um eine Stunde. Er kommt um 3 Uhr. (As a matter of principle he will be one hour late. He will arrive at three o'clock.)

The results of the Peter-experiment are given in Table 10: as we expected after the respective watch-experiment, the vast majority of subjects interpreted (12b) literally, while (12a) was interpreted *non-literally* (as we have seen before). As in the watch-example, the results exhibit a significant difference in the interpretations due to the order of the answer clauses – the two-sided Fisher-test for exact data yields a p-value of  $< 0.001$ .<sup>16</sup>

The order-effect is an interesting phenomenon regarding modal enrichment. It has to be taken into account by an explanation of *why* recipients perform modal enrichment, or *why not*, respectively. We will elaborate on this issue in section 4.

|                         | 14h      | 15h        | 16h        |
|-------------------------|----------|------------|------------|
| exp. 12a (17 subjects): |          | 2 (11.8%)  | 15 (88.2%) |
| exp. 12b (14 subjects): | 1 (7.1%) | 13 (92.9%) |            |

**Table 10:** When do you think Peter will be here?

## 2.7 Conclusion

Let us sum up so far: The *non*-literal interpretations of the examples discussed in this section are derived by the application of a modal operator that refers to an information source. In the examples, the speaker can expect the hearer to choose the *non*-literal interpretations. (The data are very robust.) The recipient in turn can expect the speaker to intend the expanded, *non*-literal meanings of his utterances; the speaker is to some degree committed to the *non*-literal meanings. However, the *non*-literal interpretation can be blocked by diverse means, including intonation. Finally, the interpretations of the examples depend on the order of their answer clauses – reversing the order in the original answer ‘turns off’ modal enrichment.

## 3 The influence of “aber” (but)

In this section, we will investigate the impact that “aber” (*but*) might have on the interpretations of the examples in which it occurs, in particular the initial watch-example (1).

We start with a brief excursus on default inferences and the meaning of “aber” (*but*) with respect to such inferences. Veltman (1996) provides us with the following examples:

- (13) a. If it rains, normally the temperature is below 15°C. ( $p \rightsquigarrow r$ )  
 b. If there is an easterly wind, normally the temperature is 15°C or higher.  
 ( $q \rightsquigarrow \neg r$ )

Let it be given that a person knows the default rule (13a) according to which the temperature is *normally* below 15°C if it rains. If this person gains the additional information that it rains then he can conclude that *presumably* the temperature is below 15°C. Analogously, if the person knows that the temperature is *normally* at least 15°C if there is an easterly wind, and he learns that there is an easterly

wind, then he can conclude that *presumably* the temperature is at least 15°C. The temperature cannot be below 15°C and 15°C or higher at the same time. Thus, if one finds out that both it rains and there is an easterly wind, then one can neither conclude that *presumably* the temperature is below 15°C nor that it is *presumably* at least 15°C. If two contradictory default inferences are licensed then none of them applies:

$$O[p \leadsto r] [q \leadsto \neg r] [p \text{ and } q] \not\vdash \textit{presumably}(r) \quad (1)$$

$$O[p \leadsto r] [q \leadsto \neg r] [p \text{ and } q] \not\vdash \textit{presumably}(\neg r) \quad (2)$$

According to Winter and Rimon (1994) the conjunction “but” (corresponding to *aber*) blocks default-conclusions of its first conjunct. Therefore, default-conclusions of the second conjunct can apply even if they contradict the (blocked) conclusions of the first conjunct. Winter and Rimon provide the following examples, among others:

- (14) a. I love Venice but I do not want to live there.  
       b. I love Venice and I do not want to live there.

Let the default rule be given, that if someone loves a city, then normally he would also want to live there. Thus, from the first clauses of the examples in (14) one can infer that presumably the speaker would like to live in Venice. The second clauses, however, explicitly contradict this conclusion – the speaker actually does not want to live in Venice. In (14a) the false default conclusion is blocked by “but”. Therefore, a conflict between the conclusion and the information given by the second clause is avoided. In (14b), in contrast, the blocking does not take place, so that there is a conflict between the default conclusion and the contradictory information of the second clause. Winter and Rimon (1994) argue that this is the reason why (14b) is less acceptable than (14a).

The function of “but” (*aber*) is to block default conclusions that are in conflict with what the speaker wants to convey. That is, “but” is used to block conclusions that are in principle *relevant* regarding the discourse topic but not consistent with the intended meaning of the speaker’s utterance.

We can apply Winter’s and Rimon’s analysis to Veltman’s example:

- (15) a. It rains and there is an easterly wind.       (*p and q*)  
       b. It rains but there is an easterly wind.       (*p and q*)  
       c. There is an easterly wind but it rains.       (*q and p*)

Let a naïve information state be given that is updated with the default rules from (13) and then with one of the sentences from (15). After an update with (15a) it comes to a clash of contradictory default conclusions so that none of them applies (as described above). In (15b) and (15c), however, the default conclusions licensed by the first clauses are blocked due to the appearance of “but”. Therefore, the default conclusions licensed by the second clauses do not clash with contradictory conclusions; they apply:<sup>17</sup>

$$0[p \rightsquigarrow r] [q \rightsquigarrow \neg r] \quad [p \text{ but } q] \models \textit{presumably}(r) \quad (3)$$

$$0[p \rightsquigarrow r] [q \rightsquigarrow \neg r] \quad [q \text{ and } b] \models \textit{presumably}(\neg r) \quad (4)$$

(End of excursus)

Which role does “aber”/“but” play in the examples of modal enrichment? Let the following two plausible default rules be given:

- (16) a. If it is 5 past 3 by a given watch, then normally it is also in fact 5 past 3.  
 $[p \rightsquigarrow q]$
- b. If it is in fact 5 past 3, then normally it is also 5 past 3 by a given watch.  
 $[q \rightsquigarrow p]$

The initial watch-example – “It’s 5 past 3 but my watch is 5 minutes fast” – can either be modally enriched or interpreted literally. The two different interpretations are given in the following, where underlining indicates content added by enrichment (or rather: paraphrases of such content). We add “in fact” to the literal interpretation just to make it definite that this example cannot be further enriched:

- (17) What time is it?
- a. By my watch it’s 5 past 3, but my watch is 5 minutes fast.
- b. In fact it’s 5 past 3, but my watch is 5 minutes fast.

We presume that Winter’s and Rimón’s analysis for the English “but” is also true for its German counterpart “aber”. Given the default rules (16), in (17a) the conclusion that presumably it is 3 past 5 is blocked, while in (17b) it is the other conclusion – according to which it is 3 past 5 by the speaker’s watch – that cannot apply. The question under discussion is what time it is. The questioner expects an answer to this question, he is not interested in the condition of Benedikt’s watch. Therefore, only in (17a), “aber”/“but” blocks a default conclusion that is *relevant*

regarding the question under discussion. It is the function of “aber”/“but” to block *relevant* conclusions that otherwise might be drawn. This function is only fulfilled if the answer is interpreted *non-literally*.

By interpreting the first clause of the answer *non-literally*, the recipient assures that “aber”/“but” can fulfill its function of blocking a relevant but potentially conflicting default conclusion. He establishes an adequate use context for “aber”/“but”. This might be a trigger for modal enrichment. (Modal enrichment is then a *repair* procedure.) Of course, it cannot be the only trigger since we have shown in section 2 that modal enrichment also takes place when the connector is not present.

In section 4, we will come back to the notion of polyphony. Let us already note here that the French connector corresponding to the German conjunction “aber”, “mais”, is one of the most prominent examples of polyphonic analysis.<sup>18</sup> The main point these theories add to the semantics presented above is that the speaker must accept and be responsible for the the *point of view* (POV) expressed by the second clause (*my watch is 5 minutes fast*) and the POV corresponding to the the default conclusion derived from this clause and the generally accepted rule  $q \rightsquigarrow \neg r$ , but does not present himself as responsible for the first clause and (consequentially) for the conclusion of the first clause and  $q \rightsquigarrow \neg r$ . In 4.3, we will discuss text-structural reasons why “It is 3 o’clock” and “Peter will arrive at 3 o’clock” may sensibly be assumed not be attributed to the speaker but someone or something else. Yet we see already here, that if we follow a polyphonic analysis, this may help to explain why “aber” should license such an interpretation more easily than no connector at all.

## 4 Relevance, reasoning, rhetorical structure

In this section, we will discuss approaches to explaining why modal enrichment occurs, or why it does not occur, respectively. Firstly, in 4.1, we will discuss the idea that modal enrichment is a repair procedure for making an utterance relevant. We will show that this attempt of an explanation fails if based merely on a ‘question under discussion’-related notion of relevance. Secondly, in 4.2, we will refer to scripts of deriving and conveying information. Modal enrichment is carried out in order to make an utterance match such a script; again, it can be considered as a repair operation. Thirdly, in 4.3, we discuss an approach which explains modal enrichment not as a repair but as a primary form of interpreting information in so-called ‘quotation contexts’. This explanation is elaborated with reference to theories of polyphony. The approaches discussed in 4.2 and 4.3 make

to some extent different predictions on meaning and acceptability. Finally, in 4.4, we will elaborate the approaches regarding these differences.

## 4.1 Relevance

In section 3, we have argued that “aber” triggers modal enrichment because only if the first answer clause is enriched can “aber” block a contextually relevant default conclusion. That is, enrichment is explained with reference to the relevance of B’s utterance regarding A’s question. “Aber” cannot be the only enrichment trigger. Yet, this suggests a question over whether, and if so how, we can further exploit the concept of ‘relevance’ for explaining why recipients perform modal enrichment.

A first attempt of explaining the initial watch-example might go as follows: without enrichment the answer is not entirely relevant. Relevance can be defined with respect to the question under discussion, namely A’s question what time it is: only information that at least partially answers this question is considered to be relevant. (Cf. respective definitions of Ginzburg 1996, Groenendijk 1999, among others.) The subjects enrich the meaning of the first answer clause in order to make the entire answer relevant. The information that the watch is five minutes fast is irrelevant unless it has an effect on the determination of the time. It does not have such an effect if “It’s 5 past 3” is interpreted literally. However, it has such an effect if “It’s 5 past 3” is interpreted in the sense of “By my watch it’s 5 past 3”.

This attempt is confronted with two problems: firstly, if the answer is interpreted *non-literally*, then the information on the watch has an effect on the determination of the time but is still irrelevant regarding the question under discussion since A asked only for the time and not for the accuracy of B’s watch. Thus, if relevance is defined exclusively with respect to answering a question under discussion, the *non-literal* interpretation does not make the entire answer relevant.

According to both the literal and the *non-literal* interpretation, B conveys information on the time and his watch. Only according to the *non-literal* interpretation, the information on the watch is crucial for determining the time, that is, for deriving the answer on A’s question. If we want to argue that by modal enrichment B’s answer is repaired to become relevant, then the concept of relevance must be so defined that it also covers the reasoning towards the sought-after information (the actual time): information can be relevant even if it is not an answer to the question under discussion but helps to derive such information.<sup>19</sup>

The second problem is that by our attempt we cannot explain why the interpretation depends on the order of the answer-clauses. We have seen that it makes a difference whether the information on the accuracy is given first or second – only if it is given second will the answer be enriched; reversing the order ‘turns off’ modal enrichment. (Cf. examples (11) and (12) above.)

On the first sight, this problem might be solved by referring to the discourse update potential of the answer clauses. It might be that due to the update potential the sequencing of the clauses can be constrained (as in Update Semantics). The idea can be illustrated with an example from Veltman (1996):

- (19) a. (knocking at the door) ... Maybe it’s Mary. ... It’s John.  
 b. (knocking at the door) ... It’s John. ... \* Maybe it’s Mary.

Let it be given that there is someone knocking at the door and that we know that it must be either Mary or John (not both). After we hear the knocking, we can utter that it might be Mary, since as far as we know it might be Mary who is knocking, even if we find out later that it is in fact John (example (19a)). However, as soon as we found out that it is John who is knocking we can no longer reasonably assume that it might be Mary. Therefore, the sequence given in (19b) is considered to be ‘bad’. This is explained with reference to the discourse update potential of [[It’s John]]. Once the common ground is updated with the information that it is John who is knocking, the assumption that it might be Mary is no longer consistent with the common ground and, thus, cannot be reasonably uttered.

We can try to transfer this approach to the watch-example: due to the update potentials of the clauses (not further specified), example (19a) is considered to be ‘good’ while (19b) is ruled out. (Remember that underlining indicates content added by enrichment.)

- (20) a. By my watch it’s 5 past 3. My watch is 5 minutes fast.  
 b. My watch is 5 minutes fast. By my watch it’s 5 past 3.

Update Semantics operates on a propositional level: one sequence of propositions can be accepted while the same propositions in reversed order are ruled out. For ruling out a sequence of propositions, it does not matter whether the propositions are derived by literal or *non*-literal interpretation. Therefore, if we rule out (20b) we also rule out the following example (21), in which the modal operator is explicitly given.

(21) My watch is 5 minutes fast. By my watch it's 5 past 3.

Example (21), however, seems to be acceptable and well interpretable; it should *not* be ruled out. Thus, the attempt to solve the order-problem by stating a constraint on sequences of propositions, like in Update Semantics, fails.

To take stock, let us reconsider example (11): we are looking for an explanation why the answer of (11a), repeated as (22a), is interpreted *non-literally* and why the answer of (11b), repeated as (22b), is interpreted literally.

(22) Wie spät ist es? (What time is it?)

- a. B: Es ist 5 nach 3. Meine Uhr geht 5 Minuten vor.  
(It's 5 past 3. My watch is 5 minutes fast.)
- b. B: Meine Uhr geht 5 Minuten vor. Es ist 5 nach 3.  
(My watch is 5 minutes fast. It's 5 past 3.)

One way to solve the issue is to assume that enrichment is in principle possible in both examples but that in (22b) the recipient has a *preference* for the literal interpretation – and in (22a) for the *non-literally* interpretation. Thus, one defines 'preferred patterns', or schemas, instead of *constraints* for sequences of propositions. These patterns have an effect if the recipient can follow his preferences and choose one interpretation. They do not have an effect in cases like (21) for which there is no nearby modal expansion.

Another way to solve the issue is to restrict the possibilities of deriving a *non-literally* interpretation, that is, of applying an operation of modal enrichment. Instead of making it possible to derive the *non-literally* and the literal interpretations of both examples and then filter out all but one interpretation in each case, this approach assumes that modal enrichment is only possible in example (22a). In (22b), the *non-literally* interpretation cannot be derived, therefore it need not be filtered out by additional rules.

We will follow the first approach in the next subsection, 4.2, and the second approach in subsection 4.3.

## 4.2 Reasoning

We assume that the scenario of the initial watch-example (1) and its variations suggests that time is read off a watch or some chronometer. This assumption is justified by the fact that most people's feeling for time is inaccurate. One can certainly also imagine a different course of events:

(23) A: What time is it?

B: (*thinking briefly*) It's 15:00. (*looking at his watch*) My watch is 5 minutes late.

But we suggest this scene would be rather comical than believable. Thus, we take reference to a chronometer for granted. This reference is part of a *script* of determining and telling the time. A recipient can presume that the speaker follows this script – which is supported by the fact that that oral utterances often show ‘traces of forming thoughts’ (“Spuren der Gedankenbildung”, Schwitalla 2006: 35).

In this subsection, we will follow the first approach mentioned above, at the end of 4.1, and explain the enrichment of (22a) and the non-enrichment of (22b) with reference to a script of determining and telling the time. We claim that modal enrichment is technically possible whenever a modal operator can be derived from the utterance context. This is the case in both examples, irrespective of their clause orders. The actual performance of modal enrichment is determined by additional rules of interpretation: modal enrichment only takes place if it helps to match the interpretation with the presupposed script (in case of the initial example, a time-telling script, in other cases other scripts of deriving information from a source before conveying it). It is in this respect a *repair* procedure.

Let us illustrate this approach on the basis of the initial example:

(24) It's 5 past 3. My watch is 5 minutes fast.

We claim that (23) can in principle be interpreted both literally and *non-literally*. However, the vast majority of test subjects chooses the *non-literal* interpretation. We reconstruct the *non-literal* interpretation as follows: the recipient learns that it is 5 past 3 by the speaker's watch. Moreover, he learns that the watch is 5 minutes fast; he concludes that it is in fact 3 o'clock. He follows the speaker's line of thought from data acquisition (reading the watch) via data interpretation to the actual answer on the question under discussion. This line of thought conforms to a script of telling the time.

By contrast, the literal interpretation can be reconstructed like this: the recipient learns that it is 5 past. Additionally, he learns that the speaker's watch is 5 minutes fast; he can now conclude that it is 10 past 3 by the speaker's watch. The recipient does not directly follow the speaker's line of thought. Instead, he understands the information on the watch as part of a retrospective explanation about how the speaker came to his estimation of the time.

We assume that recipients prefer an immediate line of thought over a retrospective explanation and, therefore, modally enrich (24).

(25) My watch is 5 minutes fast. It's 5 past 3.

Let us now reconstruct the interpretation of (25), with the clauses in opposite order. Again, we claim that (25) can in principle be interpreted both literally and *non*-literally. In contrast to (24), however, the vast majority of test subjects choose the literal interpretation of this example. We reconstruct the *non*-literal interpretation as follows: the recipient learns that the speaker's watch is 5 minutes fast. Moreover, he learns that it is 5 past 3 by the watch; he concludes that it is in fact 3 o'clock. The recipient follows the speaker's line of thought. To arrive at an answer to his question he has to draw the crucial conclusion himself. (In this subsection, we do not demand that in the script of watch-reading and time-telling the information displayed by the watch comes before the information on the accuracy of the watch. The script is compatible with both orders.)

The literal interpretation is reconstructed like this: the recipient learns that the watch is 5 minutes fast. He adds an (underspecified) dummy premise on what time it is by the watch; then, the literal meaning of the second sentence – that it is in fact 5 past 3 – is interpreted as a conclusion of the first sentence and the dummy premise. Again, the recipient follows the speaker's line of thought. To arrive at an the answer to his question he does not draw the final conclusion himself;<sup>20</sup> he assumes that the speaker draws the conclusion for him.

For the interpretation of (25), modal enrichment is not necessary to match the answer with a script of telling the time. This might be reason enough to not perform modal enrichment but to rely on the literal interpretation. Another reason might be that recipients prefer interpretations that reduce their demand of own reasoning; they choose the literal interpretation because only according to this interpretation it is the speaker who draws the relevant conclusion.

A conclusion comes by default at the end of a reasoning process and, thus, at the end of an utterance that follows the script of such a process. After the conclusion, the answer is finished, and the voice goes down. Now, let us change the example, so that the voice stays high at the end of the second answer clause:

(26) A: Wie spät ist es? (What time is it?)

B: Meine Uhr geht Minuten vor. ↑ Es ist 5 nach 3. ↑

(My watch is 5 minutes fast. ↑ It's 5 past 3. ↑)

We performed an experiment by presenting example (26) to 16 subjects. As in the other experiments, we read the example aloud twice. Unlike in the other experiments, the voice stayed high at the end of the second answer clause. (The voice is indicated by the arrows.) If we are right in assuming that the answer can in

principle be interpreted both literally and *non*-literally, and that the literal answer is chosen because there is a tendency to interpret the second clause as the conclusion drawn by the speaker, then the change of intonation should convey a change of interpretation: the voice stays high; therefore, we are not at the end of the utterance and, thus, the reasoning process; the conclusion is still to be drawn. This is matched by the above mentioned reconstruction of the *non*-literal interpretation. The subjects are therefore expected to interpret the answer rather *non*-literally than literally.

|          | 15:00    | 15:05     | '?'      |
|----------|----------|-----------|----------|
| exp. 26: | 12 (75%) | 3 (18.8%) | 1 (6.3%) |

**Table 12:** What time is it according to B's answer? (16 test subjects)

The results of the experiment are given in Table 12: as can be seen, the majority of subjects chose the *non*-literal interpretation. If we compare the data with the interpretations of example (11b), given in Table 9, – (11b) is the same answer with the voice going down at the end – to test whether there is a correlation of voice and interpretation, the two-sided Fisher test yields a p-value of  $< 0.01$ .<sup>21</sup> That is, the data are highly significant regarding a correlation of voice and interpretation; they can be interpreted as confirming our reconstruction presented above.

Let us take stock: for the approach presented in this subsection, we assume that, firstly, recipients prefer an immediate line of thought over a retrospective explanation. Line-of-thought following can be described by scripts; a *non*-literal interpretation involving model enrichment can be seen as the result of a repair procedure to make an utterance match a script of deriving and conveying information. Secondly, we claim the hypothesis that recipients prefer interpretations that reduce demands on their reasoning. Therefore, they tend to choose the literal interpretation of example (25).

The approach presented in this subsection was illustrated only with the watch-examples. It is, however, fairly straightforward to apply it also to the Peter-examples (12).

### 4.3 Rhetorical structure

In this subsection, we will present an attempt at an explanation of modal enrichment that focuses on what one may call the rhetorical structure of the discourse. We assume that the textual structure of the discourse fragments discussed here

is such that it permits or exacts a modal enrichment. In all examples under discussion, it is plausible that the speaker first presents ‘quoted’ the unfiltered information she retrieved from an information source. This ensures a fast flow of information without interruptions, if not the maximally efficient one. The recipient cannot expect that the speaker will consider information that is plausibly unfiltered to be unconditionally true.

We assume that in giving the time, the speaker generally refers to her chronometer of choice and quotes it directly. The information provided by such an external source – the watch – is generally correct and the script by which the speaker acts is probably available to and similar for most speakers. Therefore the ‘quotation’ need not necessarily be indicated. (This line of argumentation can also be applied to the Peter-examples (7).)

On the one hand, in quoting somebody else or an external information source like a chronometer, the speaker is not necessarily committed to the truth of the quoted content. On the other hand, one normally does not quote nonsense and information one considers wrong, at least not without warning the hearer that one may be passing on misinformation. If information on the reliability of the information quoted is added afterwards, the relationship between quotation and assessment of reliability is readily inferred – in the examples, the reliability assessment cannot stem from the watch (and it would be self-mockery if it were a quote from Peter). By contrast, comments on the reliability of the quotation given beforehand make a direct quotation inappropriate, as, firstly, rhetorical structure only permits unmarked quotations in clear cases, i.e. immediately after the quote-inducing question and, secondly, the script is altered in a way that the cognitive advantage that may be the reason for quoting unfiltered information is evidently lost and the correction should then be factored into the following information.

(27) A: What time is it?

- a. B: It’s 5 past 3. My watch is 5 minutes fast.
- b. B: My watch is 5 minutes fast. It’s 5 past 3.

Let us illustrate the approach with our standard example (27): A’s question induces a quotation context, so that the first answer clause of (27a) can be interpreted as a quote (by B’s watch). Although in principle the quotation could be extended to the second clause, this is not possible in this particular case because a watch cannot give information on its own accuracy. In (27b), the qualifying clause, which cannot be a quotation, comes first and immediately closes the quotation context. Therefore, the following indication of the time cannot be considered quoted information; (27b) cannot be modally enriched.

For our explanation to be plausible, we have to assume that information generally comes with a *slot* for the information source from which it comes. Stipulating such a slot for an information source may seem a large overhead; theories of polyphony (v.i.), however, may provide independent motivation.

In the remainder of this subsection, we will discuss the ideas sketched above more precisely, especially the question when information sources may be left unspecified but reliably inferred to be different from the speaker. We will also briefly discuss the relationship of our account to polyphonic analyses of utterances.

As in the previous subsection, the approach presented in this subsection will only be simulated with the watch-examples. It is, again, straightforward to apply it to the Peter-examples (12).

#### 4.3.1 Quotations in quotation contexts

When we speak of quoting, this is short for ‘passing on unfiltered information’; one should construe this to be a metaphorical way of speaking, which does not correspond to the concept of indirect speech.

As modalisation in our experiments occurs reliably, an indication of the information source cannot be obligatory in all cases to (quite) reliably convey a certain interpretation. Similarly to cases of temporal or causal relationships, we assume that textual order can provide default interpretations;<sup>22</sup> these default interpretations may be justified by reference to scripts, as suggested above.

Let us consider the original case of giving the time. We can assume that (i) the time displayed by a chronometer is generally considered correct unless we have evidence to the contrary, and (ii) a speaker generally reads the time off the watch directly, and immediately gives the information to a hearer. Assumption (i) generally justifies (ii) and both together justify the omission of a restriction like “by my watch”. Reading time off a watch stating: “By my watch, it is 15:05” would generally be perceived as pedantic and awkward unless one doubted the correctness of the watch or has reason to make one’s announcement ‘fool-proof’.<sup>23</sup> It is therefore the deviation from (i) that must be signalled, not compliance with it.

In the examples discussed above, the default assumption (i) is not in order. The speaker’s watch does not display the correct current time, and the speaker knows this. He can, however, still rely on the hearer’s assuming that he read the time off some chronometer. An efficient way to inform the hearer without delay, as sketched above, is to follow the standard script, and then specify the deviation of the watch.

Based on these considerations, we can tentatively explain the core examples (we will slightly modify this explanation in the next paragraph): if to answer a question one must refer to an external source of information, this question creates a *quotation context*, i.e. a context in which the information may be passed on unfiltered. Information presented in a quotation context are by default considered to be ‘quotations’ or unfiltered information, to which the speaker is not committed. That is, the modal enrichment of the initial example is my no means a *repair* but a primary means of interpretation. Information on the reliability of an information source are by default attributed to a information source different from the one that gave the information. Thus, a reliability assessment ‘breaks’ the quotation context, and the information following it must therefore by default be attributed to the speaker.

### 4.3.2 Polyphony

We assume that speakers can present information as stemming from a different information source than themselves. (This is readily made plausible for (in)direct speech.) When searching for more general accounts of such phenomena, we encountered theories of polyphony, which can be viewed as theories of ‘diverging [rather than common] ground’ (Nemo 2007) trying to model how the attribution of information to information sources and different degrees of speaker commitment interact to influence interpretation.

The notion of linguistic polyphony has been worked out mainly in French linguistics, following seminal work by Oswald Ducrot.<sup>248</sup> We take our terminology from ‘the Scandinavian theory of Linguistic Polyphony’ (ScaPoLine<sup>25</sup>), as it seems to lend itself to the description of what is going on in our example.

The main idea of research on polyphony is that in speaking, the speaker does often not directly and solely relate his own positions but ‘puts on stage’ different positions that he attributes to different sources which are called ‘discourse entities’ in ScaPoLine. This polyphony is indicated (‘coded’) by linguistic means (Noelke 2006: 146–147); ScaPoLine models this systematic coding.

We restrict ourselves to introducing those elements of the theory that are relevant to our description: points of view (*points de vue*, POV) and, to a lesser extent, discourse entities, in particular speaker images. Points of view can be seen as triples, which are often notated as:

[E](J(Φ))

(5)

E is initially a variable that must eventually be resolved to refer to (an image of) a discourse entity. It indicates the source of the POV, i.e. who is responsible for (which more or less corresponds to: ‘is committed to’) the POV as presented by the utterer; according to ScaPoLine, defaults of different strengths apply.<sup>26</sup>  $\Phi$  is the content of the POV, and is either a proposition (in the simple case) or (a reference to) another POV (other variants are possible, but will be ignored here). J is a judgement which indicates the attitude of the POV-holder to the POV-content. Examples are: TRUE, which implies responsibility by the POV-holder, AGREE, which ‘declines’ responsibility for the POV-content, while not refuting it, and finally UNJUSTIFIED, the opposite of taking responsibility for a POV.

Discourse entities E are considered only images of potentially real entities created by the speaker; the speaker may also depict himself or other entities in different roles at different times. For our purposes, we may need and take from Nølke et al. (2004) (§ 2.3) the following:<sup>27</sup>

- LOC, the speaker’s image presented as the one who constructs the ‘polyphonic configuration’ and sense (the ‘locutor as constructor’, ‘constructeur de sens’) and is responsible for the whole discourse fragment;
- L, the textual speaker, i.e. an image of the speaker who has not uttered a point of view (POV), but is presented as holding it;
- $l_i$ , the image of the speaker at time  $t_i$  and who is presented by LOC as uttering a POV;
- images of ON (general opinion, named for the French pronoun; can include or exclude the locutor) and potentially (groups) of individual third parties.<sup>28</sup>

Let us interpret a simple example:

(28) This example is not trivial.

In (28), ScaPoLiners would generally suspect a ‘polemic negation’, which contradicts a salient positive point of view, (Nølke 2004: 51) and would analyse it as follows:<sup>29</sup>

POV1 : [X]TRUE([[This example is trivial]]) (6)

POV2 : [ $l_0$ ]UNJUSTIFIED (POV1) (7)

The variable X representing some discourse entity need not be resolved; POV1 can be left underspecified. The resolution may be constrained, however; in the case at hand, it is evident that:  $X := l_0$  would not be possible (different ‘images of the speaker’ are discussed below).

The concept of responsibility is more elaborate than we have suggested until now: there are ‘links of (non-)responsibility’ between discourse entities and points of view.<sup>30</sup> Responsibility of X for a POV ‘[X]J( $\Phi$ )’ entails that J( $\Phi$ ) is true or realized for X; X has had the idea to put forward P and X accepts all consequences of P, including its argumentativity.<sup>31</sup> A refutative link from Y to ‘P : [X]J( $\Phi$ )’ means that Y considers J( $\Phi$ ) false or erroneous. Responsibility and refutation may be considered extreme points on (at least one) scale,<sup>32</sup> between which lie different degrees of acceptance. For our argument, the ‘classical’ ScaPoline tripartition between responsibility, non-responsibility (Nølke et al. 2004: 2.4.3) and refutation is sufficient. It is important to state that links between discourse entities and POVs may be established in the course of interpretation, thus yielding the model of ‘diverging grounds’.

Note that a speaker can agree with a POV without being responsible for it; *refutation* is a special kind of non-responsibility. The speaker – not his image; even though some abstraction is in order, especially if one is dealing with a literary text – is supposed to be responsible for the overall utterance.

Let us now discuss whether/how the context suggests to resolve the POV source to the watch in our initial example (1) and why the resolution does not work out this way in case the order of clauses is reversed.

POV1 : [X]TRUE([[It’s 5 past 3]]) (8)

POV2 : [Y]TRUE([[My watch is 5 minutes fast]]) (9)

In cases where “aber” (*but*) is present, we could assume that the speaker is responsible for POV2, i.e. we set  $Y := l_1$  (assuming that  $t_1$  is the time of uttering the second sentence). We know that  $l_0$  is not necessarily responsible for POV1, which is encoded in the meaning of “aber” (cf. section 3).

In cases without “aber”, the first clause being uttered in a quotation context, POV1, can be attributed to the watch or the speaker. POV2 must sensibly be attributed to (an image of) the speaker. This transcends the ‘module of polyphony’ as described in ScaPoLine and would need an inferential module as suggested by Kratschmer (2009), which could then refer to the scripts, as discussed above. Thus, we can describe the process of reasoning as follows: POV1 activates the script *Telling Time*, which contains the custom of immediately pronouncing the time read.

There are two possibilities to continue the analysis: firstly, we can resolve the variable X to the watch, which would stretch the metaphoric meaning of polyphony quite some (yet not necessarily the idea of discourse entities in ScaPoLine<sup>33</sup>). Secondly, we can resolve it to the image of the speaker at time  $t_0$ , when he just

reads the time from his watch innocently and pronounces it. He is corrected<sup>34</sup> by his own image at time  $t_1$ . In the latter case, ScaPoLiners would speak of ‘inner polyphony’ (namely between two images of the speaker). Arguing with images of the speaker has the advantage that we can stipulate that it is more plausible that one corrects oneself (or one’s image) afterwards than before.

Even though this is not foreseen in the works on polyphony that we have consulted, we suggest that there are structural positions in discourse at which it is possible for the speaker to utter (fragments of) statements for which he need not assume responsibility. This can be viewed as an application of free indirect speech. In discussion of indirect speech of Nølke et al. (2004) (§3, coining the term ‘hidden polyphony’, *polyphonie dissimulée*, 64) and extended ScaPoLine for longer (and literary) text fragments (§5), this idea is not discussed, as far as we can see; and the concrete extension of the resolution algorithm for POV sources seems to be of minor importance in these chapters. Regardless, the restriction that reliability assessments may occur unqualified only after the information concerned has been presented may be commonsensically attributed to the fact that quoting uncorrected information is faster, as argued above. Our preliminary rules regarding quotation contexts, stated at the end of the previous paragraph, can thus be assumed to be an extension of the tentative algorithm to infer the source of a POV and establish enunciative links incrementally and partially specified by Nølke et al. (2004) (§2.3.4 and onwards).

We can leave our notion of a quotation context from the last paragraph unmodified, but information sources are now construed to be sources of POVs. The notions of quotations in quotation context and of reliability assessments must therefore be adapted to include the notion of a POV and discourse entities: if to answer a question one must refer to an external source of information, this question creates a *quotation context*, i.e. a context in which the information may be passed on unfiltered. POVs expressed in a quotation context are by default considered to be ‘quotations’ or unfiltered information, and a link of responsibility to the speaker’s image that takes responsibility for the utterance will not be established by default. It will, however, ultimately be established if there is no evidence that the speaker declines responsibility, as he may do by uttering a reliability assessment, namely that information on the reliability of the source of a POV is by default attributed to (an image of) a discourse entity different from it. As before, a reliability assessment ‘breaks’ the quotation context, and the information following it must therefore by default be attributed to the speaker (or his image), which is the default in the absence of evidence to the contrary.

#### 4.4 Testing competing hypotheses

The two approaches introduced in the previous subsections give different explanations about why the initial watch-example (1), without “aber”, is interpreted *non-literally* and why the example with the answer clauses in reversed order is, or even must be, interpreted literally. The first approach (Reasoning) sees modal enrichment as a kind of repair; it focusses on the issue of why and under which circumstances enrichment takes place. The second approach (Rhetorical Structure) defines quotation contexts in which modal enrichment (‘quotation’) occurs by default; it therefore focuses rather on the issue of why and under which circumstances enrichment does *not* take place.

The approaches open different views on the same phenomena. Since they also make different predictions regarding interpretations and acceptability assessments they can be evaluated with regard to these predictions.

(29) My watch is 5 minutes fast. It’s 5 past 3.

The crucial example is (28), which is interpreted literally. According to the Reasoning approach, this example can in principle also be interpreted *non-literally*, while (at least a strong version of) the Rhetorical Structure approach excludes the *non-literal* interpretation.

We have already seen that the example can in fact be interpreted *non-literally* if the voice stays high at the end of the second clause. (Cf. example (26) and the results of the respective experiment given in Table 12.) That is, a strong version of the Rhetorical Structures approach that even denies the adaption of quotation contexts by intonation cannot be upheld.

Furthermore, if modal enrichment of (29) with neutral intonation is ruled out, then the following sequence of sentences should be considered inconsistent and, thus, unacceptable:

(30) Meine Uhr geht 5 Minuten vor. Es ist 5 nach 3. Tatsächlich ist es also 3 Uhr.  
(My watch is 5 minutes fast. It’s 5 past 3. So it’s in fact 3 o’clock.)

The three examples (31a–31c), by contrast, are considered to be consistent and acceptable according to both the Reasoning and the Rhetorical Structures approach.

- (31) a. Meine Uhr geht 5 Minuten vor. Es ist 5 nach 3. Wundere Dich also nicht, dass es auf meiner Uhr 10 nach 3 ist.  
(My watch is 5 minutes fast. It's 5 past 3. So don't be surprised that it's 10 past 3 by my watch.)
- b. Es ist 5 nach 3. Meine Uhr geht 5 Minuten vor. Tatsächlich ist es also 3 Uhr.  
(It's 5 past 3. My watch is 5 minutes fast. So it's in fact 3 o'clock.)
- c. Es ist 5 nach 3. Meine Uhr geht 5 Minuten vor. Wundere Dich also nicht, dass es auf meiner Uhr 10 nach 3 ist.  
(It's 5 past 3. My watch is 5 minutes fast. So don't be surprised that it's 10 past 3 by my watch.)

Thus, if we follow the Rhetorical Structures approach we firstly expect that the acceptability of example (30) is rated low by test subjects while the acceptability of (31a–31c) is rated high. Secondly, we expect that there is a significant difference between the ratings of (30) and (31a–31c). According to the Reasoning approach, however, neither a low rating of (30) nor a significant difference in the ratings of (30) and (31a–31c) is to be expected.

We let 19 test subjects rate the comprehensibility, coherence and naturalness of the examples on three 6-point Likert-scales, each ranging from 1 (*ganz eindeutig/verständlich* (fully comprehensible); *zusammenhängend* (consistent); *sehr natürlich* (very natural)) to 6 (*ganz unklar/unverständlich* (not comprehensible at all); *unzusammenhängend* (inconsistent); *sehr unnatürlich* (very unnatural)). We explained the coherence- and naturalness-criteria as follows: “*zusammenhängend: der Text fließt gut*” (*coherent: good text flow*) vs “*unzusammenhängend: es gibt mindestens einen Bruch im Text*” (*incoherent: there is at least one break within the text*); and “*sehr natürlich: würde ich so sagen; habe ich schon gehört*” (*very natural: I would make such an utterance; I have heard such an utterance*) vs “*sehr unnatürlich: würde ich nie sagen; habe ich noch nie so gehört*” (*very unnatural: I would never make such an utterance; I have never heard such an utterance*). The results of the experiment are given in Table 13.<sup>35</sup> In each table cell the numbers of subjects who chose the particular rating are named – e.g., two subjects rated the coherence of example (30) very high (‘1’). (As there is no significant effect, we give no cross-tabulation even though the data are dependent.)

We detected neither a significantly low rating of (30) nor a significant dependency of ratings and examples ((30) vs (31a–31c)).<sup>36</sup> Even in a further experiment in which we slightly changed the conditions we were not able to confirm the Rhetorical Structure approach.

|   | comprehensible |       |       |       | coherent |       |       |       | natural |       |       |       |
|---|----------------|-------|-------|-------|----------|-------|-------|-------|---------|-------|-------|-------|
|   | (30)           | (31a) | (31b) | (31c) | (30)     | (31a) | (31b) | (31c) | (30)    | (31a) | (31b) | (31c) |
| 1 | 2              | 3     | 6     | 4     | 3        | 2     | 5     | 4     | 2       | 0     | 3     | 2     |
| 2 | 6              | 4     | 4     | 8     | 3        | 3     | 3     | 7     | 7       | 8     | 3     | 9     |
| 3 | 6              | 3     | 4     | 4     | 4        | 4     | 3     | 2     | 4       | 3     | 6     | 4     |
| 4 | 2              | 4     | 3     | 1     | 4        | 4     | 5     | 5     | 2       | 3     | 3     | 1     |
| 5 | 3              | 5     | 1     | 2     | 4        | 5     | 2     | 1     | 4       | 5     | 4     | 3     |
| 6 | 0              | 0     | 1     | 0     | 1        | 1     | 1     | 0     | 0       | 0     | 0     | 0     |

**Table 13:** Rating of examples (30) and (31a–31c) (19 test subjects)

That is, the results are in favour of the Reasoning approach. However, we refrain from taking these data as hard evidence against Rhetorical Structure. Competing hypotheses regarding the possibility of *non*-literal interpretations are hard to evaluate. The interpretation of the experiment described in this section is not straightforward: firstly, it could be that our questions were not clear enough and test subjects did not ‘know what we wanted’.<sup>37</sup> Secondly, it could be that our examples could be ‘repaired’ easily enough, and that this need for repair did not change the overall quality of the stimuli very much. This would mean that we have to establish some more suitable test sentences (ones that are ‘better’ overall, so that the differences matter more) and a baseline to distinguish when effects of factors are linguistically interesting.

It becomes obvious that it is still necessary to explain what the status of the experimental data is in formal pragmatics. If we accept empirical data as the gold standard, we probably need a quantitative paradigm for formal pragmatics which is similar to the sociolinguistic paradigm developed by William Labov.<sup>38</sup>

## 5 Conclusions

Let us summarize: by an operation of modal enrichment, a modal operator is constructed and applied to the (literal<sup>39</sup>) meaning of a sentence. We have shown that recipients can perform operations of modal enrichment, that these operations are conventionalized, that speakers can foresee whether their utterances will be enriched and that they are to some degree committed to an enriched, *non*-literal interpretation.

One, albeit unnecessary, trigger for modal enrichment is the occurrence of “aber”/“but”: given that a literal interpretation makes the use of “aber”/“but” not fully adequate, modal enrichment is applied to change this interpretation

and thereby provide an adequate use context for the conjunction.<sup>40</sup> Modal enrichment is thus a *repair* procedure.

We discussed three further approaches to explaining why modal enrichment takes place. According to the first approach (Relevance), modal enrichment is a repair procedure to make an utterance relevant. We have shown that the approach fails if merely based on a ‘question under discussion’-related notion of relevance. The second approach (Reasoning) also explains modal enrichment as a repair procedure, namely for making an utterance fit a given script of deriving and conveying information. We have shown that by this approach we can explain the phenomena and make correct predictions on interpretation and acceptability. The third approach (Rhetorical Structure) assumes that rhetorical structure provides a slot for a modal operator. Modal enrichment is not the free application of a modal operator but an instantiation of this slot; it is hence *not* a repair procedure. This approach also explains the phenomena. However, it differs from the Reasoning approach in some predictions regarding acceptability. We have evaluated the different approaches. The results confirm Reasoning but they do not seem clear enough to consider them as decisive evidence against Rhetorical Structure.

Open issues still to be investigated are, firstly, an overview on the operators of modal enrichment that can be constructed and applied, and, secondly, the relation and interaction of modal enrichment with other forms of meaning expansion. This demands formal modelling. Thirdly, as we remarked at the end of 4.4, methodological concerns regarding experimental data in formal, truth-conditional pragmatics are to be solved.

## Notes

\* We thank all test subjects who participated in the experiments. We also thank to Benedikt Löwe for providing us with the paradigmatic watch-example, and Patrick Brandt, Eric Fuß, Uwe Kirschenmann, Udo Klein, Marcus Kracht, Joost Kremers, Manfred Kupffer, Charlotte Matheson, Cécile Meier, Jens Michaelis, Katja Niemann, Steffen Pappert, Christian Plunze, Bernhard Schröder, Magda Schwager, Julia Steube, Petra Wagner, Martin Wolpers, Henk Zeevat, Ede Zimmermann for helpful comments, letting us borrow test subjects and stimulating discussions. All mistakes are ours.

1 In total, 42 test subjects participated in this experiment, among them 26 native speakers and 16 non-native speakers of German. The preference for enrichment was significant:  $\chi^2_{df=1} = 21.429$ ;  $p = 3.673 \cdot 10^{-06}$ . There was no significant difference between the groups of native and non-native speakers regarding their answers. (All non-native speakers replied that it was 15:00h according to the answer; the 6 subjects who interpreted the answer literally were native speakers of German.)

2 If one compares enriched versus non-enriched interpretations, the difference is significant

( $\chi^2_{df=1} = 27.000$ ;  $p = 2.035 \cdot 10^{-07}$ ;  $\chi^2_{df=1} = 11.636$ ;  $p = 0.001$ ;  $\chi^2_{df=1} = 21.333$ ;  $p = 3.860 \cdot 10^{-06}$ ), similarly so if one compares the preference for the intended interpretation ( $\chi^2_{df=1} = 24.083$ ;  $p = 9.226 \cdot 10^{-07}$ ;  $\chi^2_{df=1} = 8.909$ ;  $p = 0.003$ ;  $\chi^2_{df=1} = 18.750$ ;  $p = 1.490 \cdot 10^{-05}$ ). All significances 'survive' a Bonferroni correction, evidently.

**3** Cf. Recanati (2010) on modulation. It might be that Recanati would not accept our strong distinction between conventionalized operations and free modulation.

**4** There is evidently no significant difference between (3a), (3b) and (3c). If we compare the preferences for the 'correctly' enriched versus the literal interpretation, the difference is significant for all three experiments ( $\chi^2_{df=1} = 11.267$ ;  $p = 0.001$ ;  $\chi^2_{df=1} = 14$ ;  $p = 0.000$ ;  $\chi^2_{df=1} = 13.000$ ;  $p = 0.000$ ). Further testing the intended vs. other interpretations, the third experiment is only marginally significant due to Bonferroni correction ( $\chi^2_{df=1} = 7.118$ ;  $p = 0.008$ ;  $\chi^2_{df=1} = 7.118$ ;  $p = 0.008$ ;  $\chi^2_{df=1} = 4.765$ ;  $p = 0.029$ ).

**5** A similar, but different, experiment has been described by Schmitz (2008a) (37–38, footn. 8).

**6** The 'other' days (fifth column) were Tuesday and Saturday.

**7** If we test the preference for Friday vs. other solutions (omitting '?'), the difference is not significant ( $\chi^2_{df=1} = 0.610$ ;  $p = 0.435$ ). It is significant, however, if one tests the preference for Friday vs. the literal interpretation ( $\chi^2_{df=1} = 7.258$ ;  $p = 0.007$ ) or the preference for enrichment (Monday, Friday) vs. literal interpretation ( $\chi^2_{df=1} = 12.737$ ;  $p = 0.000$ ).

**8** They did not reach compliance with conversational maxims but with the demands of an intelligence test. – How are these demands defined if not by the performance of a given operation? Reaching compliance in this way is not an argument for free modulation.

**9** Different statistical evaluations are possible; the least favorable to enrichment is comparing 0-1 to all other options. The difference is still significant: (Trial 1:  $\chi^2_{df=1} = 28.880$ ;  $p = 7.700 \cdot 10^{-08}$ ; Trial 2:  $\chi^2_{df=1} = 26.133$ ;  $p = 3.186 \cdot 10^{-07}$ ; Total:  $\chi^2_{df=1} = 54.450$ ;  $p = 1.595 \cdot 10^{-13}$ ).

**10** Isn't it totally unnatural that a train is always on time and one can know this? – No, it is not.

**11** 16:10–10 minutes instead of 16:10 + 10 minutes.

**12** The preference for enrichment is significant for (7a) ( $\chi^2_{df=1} = 9.941$ ;  $p = 0.002$ ), while for (7b) due to the small number of subjects, the preference is non-significant ( $\chi^2_{df=1} = 1.600$ ;  $p = 0.206$ ); there is no significant difference between both experiments ( $p = 0.3261$  according to Fisher's Exact Test).

**13** A case in point where one would want to know what Peter said rather than when he arrives would be if one had had some experience with Peter's being late, but cannot be sure about the interlocutor's experiences. Yet note that in this case, one would expect the interlocutor to not make a difference between the arrival time announced by Peter and the time the interlocutor expects Peter to arrive.

**14** The German "Uhr" can be both translated by "watch" and "clock". For the experiment, a chronometer is drawn on the blackboard, so that "clock" seems more adequate. The relevant property of the chronometer is that the speaker has at least as much knowledge (regarding its precision) as the hearer.

**15**  $p = 5.645 \cdot 10^{-05}$  for a two-sided test enriched vs. non-enriched (whether or not we differentiate between the two enrichments);  $p = 9.932 \cdot 10^{-05}$  for 15:00 vs. the other readings.

**16**  $p = 6.844 \cdot 10^{-06}$  if we take both enrichments together.

**17** The blocking function of "but" and other words like, e.g., the German adverb "eigentlich" can be modelled within Veltman's framework of 'Defaults in Update Semantics'. Schmitz and Schröder (2004) make a proposal of a respective formalization; cf. also Schmitz (2008b).

**18** Cf. e.g. Nølle et al. (2004), § 4.3 and 4.4 and Nølle (2009), § 3.2, also for different nuances of 'contradiction'.

**19** This is not the concept of relevance defined by Ginzburg (1996) or Groenendijk (1999), but it has surely the flavour of Relevance Theory (Sperber and Wilson 1996).

**20** He even *cannot* draw this conclusion because the dummy premise does not give actual information on the watch-display.

**21**  $p = 0.005$  for a one-sided test.

**22** Temporal and clausal relations are the standard examples for explaining how order can affect the interpretation of an utterance. Cf., e.g., the example from Horn (2006) “If they get married and have a baby, their financial situation will be better than if they have a baby and marry.”

**23** If one does not doubt that one’s watch displays the correct time, identifying the information source this way could only be warranted in very unusual cases, where one looks at the watch, but uses another information source for giving the time, *and* where the information source is relevant to the success of communication. It is quite difficult to imagine such a case. Furthermore, the hearer would normally notice the speaker’s information source anyway, e.g. if he refers to a wall clock or mobile phone.

**24** The roots go further back to Mikhail Bakhtin; see Dendale (2007) for a brief sketch of the term’s history and a comparison of different approaches to polyphony in the sense mentioned above. Note that there are various notions of polyphony, e.g. in sociolinguistics. We restrict ourselves to the one we found most helpful in analysing our examples, though.

**25** See Nølke et al. (2004), or more briefly Nølke (2009) and Nølke (2006) (in English) for an introduction; Dendale (2007) (§3) gives a short overview of ScaPoLine in English. The presentation of Nølke (2006) is more detailed, and gives some background, both historical and in contrast to ‘the Anglo-Saxon tradition’ (137–138; §2). Kratschmer (2009) gives an introduction and some background information in German; she also provides (§5) a Danish-German-English glossary of terms.

**26** Cf. Nølke et al. (2004), §2.1.4; Nølke (2006), 142–143.

**27** See also Nølke (2006) (§4.2) and Nølke (2009).

**28** One point about this distinctions between different images of the speaker and other real-world persons as represented by the speaker in the discourse (and hence called ‘discourse entities’ or ‘êtres discursifs’ in ScaPoLine) is important, namely that the persons mentioned in discourse may well be misrepresented and that – e.g. in cases of irony – the speaker may play with different images of himself. Yet, this clearly exceeds the scope of this article. Similarly, we are not concerned with discourse entities which are purely technically necessary.

**29** Following Nølke et al. (2004), 29–30; Nølke (2006), 144–145.

**30** The relationship between the links and the judgment is not entirely clear as acknowledged by Nølke (2009), who promises a clarifying paper.

**31** Cf. Nølke (2009: 26), which we paraphrase here. Remember that discourse entities are ‘created’ by the speaker L.

**32** Cf. still Nølke 2009: 26.

**33** Cf. the example by Nølke 2006: (145) with the fruit juice ‘locutor’ of: “Drink me without sugar!” written on a bottle.

**34** Speaking of ‘correction’ draws on the polysemy of the term. One had probably better speak of ‘instructions that suggest that a correction to the information given previously would be in order’.

**35** Friedman tests for overall coherence and naturalness are significant ( $p < 0.049$ ) and marginally significant ( $p < 0.084$ ), respectively. This suggests that ratings do not depend very much on the stimuli, even more so because comparing the remaining stimuli pairwise with

(30) there is no significant difference, only (31a) is marginally significantly rated more natural ( $p = 0.094$ ) – and of course, we have to apply a correction.

36 See fn. 39.

37 Yet of course, they shouldn't have to know exactly what we wanted to investigate to understand our questionnaires.

38 Cf. Bayley (2002) and Fisseni (2011), 2.2.2.

39 Of course, it might be that also other operations of meaning enrichment are carried out and that therefore the modal operator is not applied to the literal but an already expanded meaning of the sentence.

40 Or, adverb, in the case of “aber”.

## 6 References

- Bayley, Robert. 2002. The quantitative paradigm. In: *The Handbook of Language Variation And Change*, J. K. Chambers, Peter Trudgill and Natalie Schilling-Estes (eds.), 117–141. Malden, Oxford, Carlton: Blackwell.
- Dendale, Patrick. 2007. A critical survey and comparison of French and Scandinavian frameworks for the description of linguistic polyphony and dialogism. In: Rita Therkelsen, Nina Møller Andersen and Henning Nølke (eds.), *Sproglig Polifoni. Tekster om Bachtin og ScaPoLine* [Linguistic Polyphony. Texts on Bakhtin and ScaPoLine], 109–144. Aarhus: Aarhus Universitetsforlag.
- Fisseni, Bernhard. 2011. *Focus: Interpretation? Empirical Investigations on Focus Interpretation*. Duisburg: Universitätsverlag Rhein-Ruhr.
- Ginzburg, Jonathan. 1996. Interrogatives: Questions, Facts, and Dialogue. In: Shalom Lappin (ed.), *Handbook of Contemporary Semantic Theory*, 385–422. Amsterdam: Elsevier.
- Groenendijk, Jeroen. 1999. The logic of interrogation. Classical Version. In: Tanya Matthews and Devon Strolovitch (eds.), *SALT IX: Semantics and Linguistic Theory*, 109–126. Ithaca: Cornell University.
- Horn, Laurence R. 2006. The border wars: a neo-Gricean perspective. In: Klaus von Heusinger and Ken Turner (eds.), *Where Semantics Meets Pragmatics*, 21–48. Amsterdam: Elsevier.
- Kratschmer, Alexandra. 2006. Polyphoniethoretische Formalisierung von historischen und Drittpersonsdiskursindividuen in narrativen Texten. In: *Arbejdsrapporter 5, Sprogligt Polyfoninetværk*, 33–63. Roskilde: RUC.
- Kratschmer, Alexandra. 2009. La decodifica del narratore implicito: polifonia – conoscenze enciclopediche – ragionamento probabilistico. Proposta di una pragmatica formale «scapolinizzante» [The decoding of the implicit narrator: polyphony – encyclopedic knowledge – probabilistic reasoning. Proposal of a “ScaPoLinizing” formal pragmatics], In: Alexandra Kratschmer, Merete Birkelund and Rita Therkelsen (eds.), *La polyphonie: outil heuristique linguistique, littéraire et culturel* [Polyphony: linguistic, literary and cultural heuristic tool], 219–260. Berlin: Thimme.
- Nemo, François. 2007. The pragmatics of common ground: from common knowledge to shared attention and social referencing. In: Anita Fetzer and Kerstin Fischer (eds.), *Lexical Markers of Common Grounds*, Amsterdam: Elsevier.

- Nølke, Henning. 2006. The semantics of polyphony (and the pragmatics of realization). *Acta linguistica Hafniensia* [Copenhagen Linguistic Documents] 38: 137–160.
- Nølke, Henning. 2009. La polyphonie de la ScaPoLine 2008 [The Polyphony of ScaPoLine 2008]. In: Alexandra Kratschmer, Merete Birkelund and Rita Therkelsen (eds.), *La polyphonie: outil heuristique linguistique, littéraire et culturel* [Polyphony: linguistic, literary and cultural heuristic tool], 11–42. Berlin: Thimme.
- Nølke, Henning, Kjersti Fløttum, Coco Norén. 2004. *ScaPoLine – La théorie scandinave de la polyphonie linguistique* [ScaPoLine – the Scandinavian theory of linguistic polyphony]. Paris: Kimé.
- Recanati, François. 2004. *Literal Meaning*. Cambridge: Cambridge University Press.
- Recanati, François. 2010. *Truth-Conditional Pragmatics*. Oxford: Oxford University Press.
- Schmitz, Hans-Christian. 2008a. *Accentuation and Interpretation*. London: Palgrave Macmillan.
- Schmitz, Hans-Christian. 2008b. 'eigentlich' again. In: Atle Grønn (ed.), *Proceedings of SuB12*, 567–581. Oslo: ILOS.
- Schmitz, Hans-Christian. 2011. Blocking modal enrichment (tatsächlich). In: Markus Steinbach and Jörg Meibauer (eds.), *Experimental Semantics/Pragmatics*, 177–197. Amsterdam: John Benjamins.
- Schmitz, Hans-Christian and Bernhard Schröder. 2004. Updates with "eigentlich". *Sprache und Datenverarbeitung* 28(1): 87–96.
- Schwitalla, Johannes. 2006. *Gesprochenes Deutsch. Eine Einführung* [Spoken German. An Introduction]. 3rd edition. Berlin: Erich Schmidt Verlag.
- Sperber, Dan and Deirdre Wilson. 1996. *Relevance. Communication and Cognition*. 2nd edition. Oxford: Blackwell.
- Veltman, Frank. 1996. Defaults in Update Semantics. *Journal of Philosophical Logic* 25: 221–261.
- Winter, Yoad and Mori Rimon. 1994. Contrast and implication in natural language. *Journal of Semantics* 11: 365–406.