



# Variation of Accent Type and of Context – Influences on Pragmatic Focus Interpretation

Charlotte Wollermann<sup>1,2</sup>, Ulrich Schade<sup>1,3</sup>, and Bernhard Schröder<sup>2</sup>

<sup>1</sup>Speech and Communication, University of Bonn, Germany

<sup>2</sup>German Linguistics, University of Duisburg-Essen, Germany

<sup>3</sup>Fraunhofer Institute for Communication, Information Processing and Ergonomics FKIE, Germany

cwo@sk.uni-bonn.de, ulrich.schade@fkie.fraunhofer.de, bernhard.schroeder@uni-due.de

## Abstract

We present an empirical study on the variation of accent type and of context on pragmatic focus interpretation. The material was based on audio-recordings of nine German speakers who were instructed to read dialogues with embedded question-answer pairs in which the answers constituted the pragmatic focus of the utterance. Different accent types occurred for marking the focus constituent. The audio-material was presented to 53 subjects. Interpretation was tested by using pictures intended to illustrate the (non-)exhaustive reading. When presenting the picture illustrating the non-exhaustive reading, the results show in general a significant influence of both context and prosody, but the contextual influence is stronger.

**Index Terms:** prosody, uncertainty, exhaustivity, focus interpretation

## 1. Introduction

In this paper we present an experimental study on the role of the variation of accent type and of context on pragmatic focus interpretation.<sup>1</sup> Our research questions are as follows: Is there an effect of prosodic cues of uncertainty on the exhaustivity of answers? Does accent type variation and context variation influence the hearer's assumptions on the speaker's certainty on the issue under discussion?

### 1.1. The concept of pragmatic focus

The term *focus* often refers to the intuition that pitch accent correlates with new information in utterances, whereas old information is deaccented. This phenomenon can be observed in West Germanic languages like English, German or Dutch [e.g. 1]. Even though focus is a complex phenomenon and labelled differently [2], most theories agree that focus can be defined as the answer to an explicitly or implicitly given question, e.g. [3: 261]. This phenomenon is also referred to as *pragmatic focus*. Semantic-pragmatic focus theories [4, 5] expect that in the context of a question, pitch accent is highly correlated with focus. If the focus is detected, the interpretation should be biased towards exhaustivity.

If the background question is interpreted as a *mention-all* question, the precondition for an *exhaustive* interpretation is given. If the hearer of (1b) concludes that *only* John kissed Mary, she interprets the answer exhaustively. In the case of non-exhaustivity there might be also other persons who kissed

Mary. Exhaustivity depends on the knowledge about the answer which is ascribed to the speaker by the hearer [6].

(1a) Who kissed Mary?

(1b) [John]<sub>F</sub> kissed Mary.

We assume that if the speaker signals (un)certainly she is (not) completely informed; we only test exhaustive interpretation which is epistemically strong in the terms of [6].

The findings of [2] suggest that not only accentuation effects focus interpretation, but also the hearer's expectations, the sensitization for focus phenomena and in particular contextual factors. For written speech, the results of [7] show also evidence that question and answer type affects the exhaustivity.

### 1.2. Expression of uncertainty

[8, 9] reported that uncertainty is expressed and also perceived by different prosodic cues: rising intonation, delays, fillers and lexical cues. According to [10] fall-rise intonation contributes to a context-independent meaning of utterance interpretation conveying the speaker's uncertainty in English. However, for German there are barely any empirical data on the expression of uncertainty by prosodic cues regarding the pragmatic level.

### 1.3. Assumption

We assume that if the speaker signals (un)certainly with respect to the answer the hearer perceives this prosodic information for decoding the utterance. The hearer will assume that the speaker is (un)certain regarding the answer, then a (non-)exhaustive interpretation should be preferred.

## 2. Related work

In [11] and [12] we tested the influence of intonation and context on the exhaustivity of answers. Results show that the exhaustive reading is generally preferred. However, the interpretation is biased towards the non-exhaustive interpretation when the micro and macro context are biased towards uncertainty and non-exhaustivity. Compared to the contextual influence the prosodic influence on exhaustivity is relatively weak. From these findings we derived a model of pragmatic focus interpretation [13]: the context is relevant for raising the hearer's expectations; these expectations are top-down influences and the prosodic information influences the focus interpretation bottom-up.

Since our data suggest only a weak influence of prosody on focus interpretation, we conducted a production study [14, 15]

<sup>1</sup>We would like to thank Till Krempel for helping us with the experimental set-up.

in order to see what prosodic cues speakers realize when producing pragmatic focus utterances. The following accent types for realizing the focus constituent were found: L+H\* was realized most frequently for the marking of the focus constituent, followed by H\* or no accentuation. A few realizations of L\* also occurred. The detailed analysis of the analysis of the frequencies of accent types can be found in [14, 15]. However, H\* appeared more often for the contextual variant intended to be biased towards uncertainty and non-exhaustivity than for the contextual variant intended to be biased towards certainty and exhaustivity. We interpret this as a manifestation of the biological codes [16]: high pitch expresses uncertainty and continuation on the pragmatical level, whereas low pitch expresses certainty and finality. We use the material elicited in [14, 15] for our current study.

### 3. Interpretation study

#### 3.1. Goal

The goal of this study is to test to what extent the variation of accent type and of context influences the exhaustivity of answers.

#### 3.2. Material

The stimuli were elicited in our production study [14, 15] and based on the material used in [11, 12]. The material consisted of six question-answer pairs embedded in different dialogues. The scenario was a fictitious party where different groups of students acted differently. For every action, there was a question asking about the agent and an answer providing the information. The focus exponent in the answer was either a noun phrase (NP) referring to one group of students (dialogue 1, 4 and 6) or a coordinated NP referring to two groups of students (dialogue 2, 3 and 5). We refer to one group by *focus sentence with one NP* and to two groups by *focus sentence with (a coordination of) two NPs*.

Two variants of context were generated. i) The variant I was intended to have a bias towards certainty and exhaustivity. It was characterized by *contextual congruity*. One student group was salient during the whole dialogue, a question followed which was congruent to the focus utterance (see 2a + 2c). No alternatives were given in the context. Further, a sentence indicating certainty about the answer followed (see 2e). ii) The variant II was characterized to have a bias towards uncertainty and non-exhaustivity. It was marked by *contextual incongruity*: one “competing” discourse entity was introduced at the beginning of the dialogue (*usually the linguists are late for parties*). Thus the alternatives were explicitly given. A general question followed (see 2b) attached to broad focus and the respective students group as focus of the answer (see 2d) was different from the salient discourse entity. A sentence indicating uncertainty about the answer followed (see 2f). Each student group was characterized by an unique accessory, i.e. mathematicians were wearing green glasses, linguists blue hats etc.

(2a) Wer ist zu spät gekommen? *Who was late?*

(2b) Was ist passiert? *What happened?*

(2c) [Die Mathematiker]<sub>F</sub> sind zu spät gekommen. [*The mathematicians*]<sub>F</sub> were late.

(2d) [Die Mathematiker sind zu spät gekommen.]<sub>F</sub> [*The mathematicians were late.*]<sub>F</sub>

(2e) Das waren die Einzigen, die nicht pünktlich waren... *They were the only ones not being in time...*

(2f) Wenn ich mich nicht täusche, waren die Linguisten auch nicht pünktlich... *If I am not wrong, the linguists were also not in time...*

As already mentioned, there were 6 dialogues per speaker. Each speaker realized for three dialogues the contextual variant I and for the other three dialogues the contextual variant II. In contrast to our production study [14], in our current study we removed the sentence indicating (un)certainly about the answer for each dialogue from our data for testing the influence of uncertainty as a paralinguistic expression on focus interpretation.

There were 54 dialogues with embedded question answer pairs (9 speakers x 6 dialogues) and 18 filler-dialogues (9 speakers x 2 dialogues); the two filler-dialogues were always the same.

#### 3.3. Testing focus interpretation

We tested the focus interpretation by using a variant of the picture selection task [17]. There were two different types of pictures. For half of the dialogues (dialogue 2, 3 and 6) we used the pictures showing the exhaustive reading to test focus interpretation, i.e. there was only one group of students illustrated in the picture performing the question under discussion. In the case of example (2) the picture showed only the mathematicians being late. For the other half of the dialogues (dialogue 1, 4 and 5) the pictures illustrating the non-exhaustive interpretation were presented. Here the alternatives were also illustrated in the picture. From the subjects' choice of the picture we inferred the preference of interpretation. Thus influences of expectations of the subjects with respect to the goal of the experiment should be minimized.

#### 3.4. Hypothesis

Based on our previous studies [11, 12, 15] we assumed that the exhaustive interpretation is influenced by accentuation and context. We expected that the degree of congruency between the information conveyed by the prosody, context and picture effected the exhaustivity: a context with incongruent information (variant II) combined with H\* as accent type for realizing the focus constituent and the picture showing the intended non-exhaustive reading was expected to be biased towards non-exhaustivity, whereas a context with congruent information (variant I) combined with L+H\* for realizing the focus utterance and a picture intended to illustrate the exhaustive reading was expected to be biased towards exhaustivity. Based on our findings reported in [15] for two coordinated NPs it was expected that the congruity between the two accents, i.e. for realizing the first and second NP, was biased towards exhaustivity, whereas the incongruity between the two accents was biased towards non-exhaustivity.

#### 3.5. Procedure

The material was presented as an online experiment to 53 subjects; all of them students and German native speakers. Each subject had to listen to exactly one speaker uttering the eight dialogues (6 dialogues with the focus utterances and 2 dialogues functioning as distractors). After each dialogue was played, subjects had to judge on a 5-point Likert-Scale how well the picture suits the dialogue (1=very bad, 5=very good). For every picture there was a caption with the different students groups characterized by the appropriate accessory. Furthermore for each dialogue there was a multiple-choice-question asking for the subjects' personal opinion about an aspect of the dialogue functioning also as a distractor.

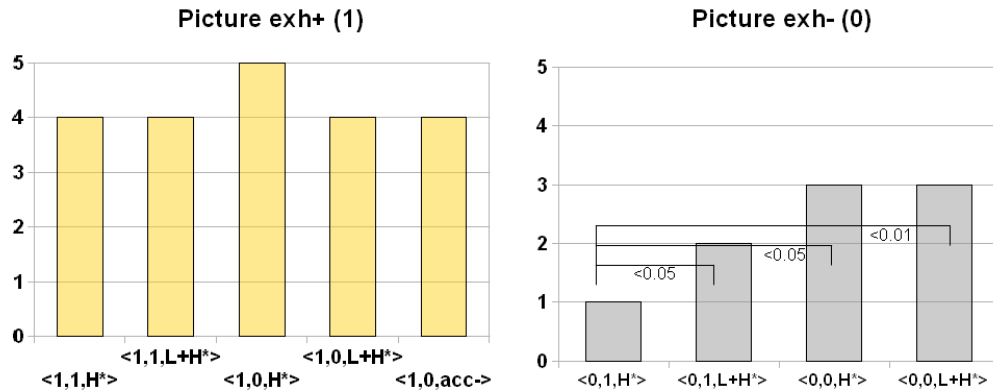


Figure 1: Recipients judgements for dialogues with focus sentences with one noun phrase; medians

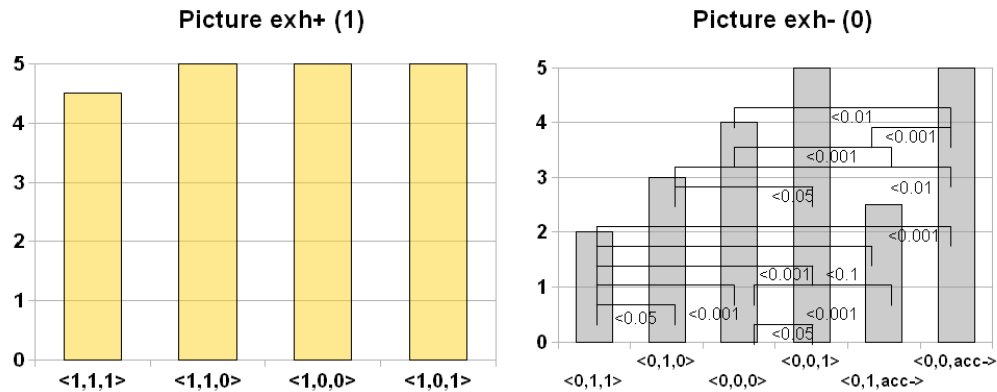


Figure 2: Recipients judgements for dialogues with focus sentences with two noun phrases; medians

Our independent variables were the *context* and *accent type*; our dependent variable was the *recipients' judgement*. The results were statistically analysed using the Wilcoxon Rank Sum Test since our data were ordinalscaled and since the group sizes differed. Our level of significance was 5%.

## 4. Results

In figure 1 the medians are shown for dialogues with *focus sentences with one NP*. The left diagram illustrates the medians when presenting the picture illustrating the exhaustive reading, the right diagram illustrates the medians presenting the picture illustrating the non-exhaustive reading. On the x-axis the variant of the picture, of the context and of the accent type is indicated. Using the ordered triple  $\langle x,y,z \rangle$   $x$  refers to the picture and has either the value 1 (the picture illustrating exhaustivity) or 0 (the picture illustrating non-exhaustivity),  $y$  refers to the context and has also either the value 1 (variant I) or 0 (variant II),  $z$  refers to the accent type: here we distinguish between  $L+H^*$ ,  $H^*$  and  $acc-$  (no accentuation).

When we present the picture illustrating the exhaustive reading we observe the following:  $\langle 1,1,H^* \rangle$ ,  $\langle 1,1,L+H^* \rangle$ ,  $\langle 1,0,L+H^* \rangle$  and  $\langle 1,0,acc- \rangle$  is ranked with a median of 4. The median for  $\langle 1,0,H^* \rangle$  is 5. The statistical analysis does

not show significant between the judgements,  $p$  each is each time  $>0.05$ . We infer that neither the context nor the accent influences the focus interpretation when presenting the picture illustrating exhaustivity.

The following can be observed when we present the picture showing the non-exhaustive reading:  $\langle 0,1,H^* \rangle$  is judged with a median of 1 and  $\langle 0,1,L+H^* \rangle$  with a median of 2. Furthermore, both  $\langle 0,0,H^* \rangle$  and  $\langle 0,0,L+H^* \rangle$  have median of 3. The comparison  $\langle 0,1,H^* \rangle$  vs.  $\langle 0,1,L+H^* \rangle$  shows a significant difference with  $p<0.05$ . Here the accent type is varied. We infer that the judgement is influenced by the context, such that  $L+H^*$  combined with the context being biased towards exhaustivity is judged as more adequate than the same context combined with  $H^*$ . The difference between  $\langle 0,1,H^* \rangle$  vs.  $\langle 0,0,H^* \rangle$  is also significant with  $p<0.05$ . Here the context influences the interpretation in favour of non-exhaustivity. Finally,  $\langle 0,0,L+H^* \rangle$  is rated as better suiting the dialogue than  $\langle 0,1,H^* \rangle$  with  $p<0.01$ . It is assumed that the context has the stronger impact in favour of non-exhaustivity in this case given the fact that the context being biased towards non-exhaustivity presented with the picture illustrating the non-exhaustive reading is rated as more adequate than the same picture combined with  $H^*$ .

In figure 2 the medians are shown for dialogues with *focus sentences with two NPs*. The left diagram illustrates the medians when presenting the picture illustrating the exhaustive reading, the right diagram illustrates the medians presenting the picture illustrating the non-exhaustive reading. On the x-axis the variant of the picture, the context and the congruity of accent type is indicated. By accent congruity we refer to those cases where the same accent is realized for both the first and second noun phrase, i.e. either L+H\* occurs both times or H\*. Using the ordered triple  $\langle x, y, z \rangle$  x refers to the picture and has also either the value 1 (the picture illustrating exhaustivity) or 0 (the picture illustrating non-exhaustivity), y refers to the context and has either the value 1 (context variant I) or 0 (context variant II), z refers to the congruity of accent type and has the value 1 (congruity), 0 (incongruity) or acc- (no accentuation for both NPs).

With respect to the picture illustrating the exhaustive interpretation our data reveal the following:  $\langle 1, 1, 0 \rangle$ ,  $\langle 1, 0, 0 \rangle$  and  $\langle 1, 0, 1 \rangle$  is ranked with a median of 5;  $\langle 1, 1, 1 \rangle$  is judged with a median of 4.5. Against our expectation the recipients' judgements do not differ in a significant way, p is each time  $> 0.05$ . Our data do neither show evidence for a contextual effect nor for an effect of prosody when presenting the picture illustrating the exhaustive interpretation.

When presenting the picture showing the non-exhaustive reading it can be observed that both  $\langle 0, 0, 1 \rangle$  and  $\langle 0, 0, \text{acc-} \rangle$  have a median of 5. The median for  $\langle 0, 0, 0 \rangle$  is 4, the median for  $\langle 0, 1, 0 \rangle$  is 3, the median for  $\langle 0, 1, \text{acc-} \rangle$  is 2.5 and the median for  $\langle 0, 1, 1 \rangle$  is 2. The statistical analysis shows significant differences between the judgements for the following comparisons with  $p < 0.001$ :  $\langle 0, 0, 0 \rangle$  is rated as better suiting the dialogue than  $\langle 0, 1, 1 \rangle$ . Here the context and the accent influence the interpretation to be biased towards non-exhaustivity.  $\langle 0, 0, 1 \rangle$  is judged as more adequate than  $\langle 0, 1, 1 \rangle$ . Here we observe an influence of context to be biased towards non-exhaustivity. Further  $\langle 0, 0, \text{acc-} \rangle$  is rated as better suiting the dialogue than  $\langle 0, 1, 1 \rangle$ . Again, we conclude that the context effects the non-exhaustive interpretation. In a similar way  $\langle 0, 0, \text{acc-} \rangle$  is judged as more adequate than  $\langle 0, 1, \text{acc-} \rangle$ . We also interpret this as evidence for a contextual influence on non-exhaustivity. Further,  $\langle 0, 0, 0 \rangle$  obtains higher rankings than  $\langle 0, 1, \text{acc-} \rangle$ . Here the context and the prosodic information effect the interpretation in favour of non-exhaustivity.

For the following cases the statistical analysis shows p-values  $< 0.01$ :  $\langle 0, 0, \text{acc-} \rangle$  obtains higher rating than  $\langle 0, 1, 0 \rangle$ . Here the context being biased towards non-exhaustivity combined with the picture illustrating the non-exhaustive reading is rated as more adequate than the same picture in combination with accent incongruity. We interpret this as evidence for a contextual influence. Furthermore,  $\langle 0, 0, \text{acc-} \rangle$  is judged as better suiting the dialogue than  $\langle 0, 0, 0 \rangle$ . Against our expectation the prosodic information does not effect the non-exhaustive interpretation in this case.

The following comparisons show a p-value of  $< 0.05$ :  $\langle 0, 0, 1 \rangle$  vs.  $\langle 0, 1, 0 \rangle$ . In this case the context being biased towards non-exhaustivity combined with the picture showing the non-exhaustive reading obtains higher rankings than the same picture in combination with accent incongruity. Here the context influences the non-exhaustivity. Against our expectation,  $\langle 0, 0, 1 \rangle$  is judged as better suiting the dialogue than  $\langle 0, 0, 0 \rangle$ . Furthermore,  $\langle 0, 1, 0 \rangle$  obtains higher rankings than  $\langle 0, 1, 1 \rangle$ . In this case the prosodic information influences the recipients' judgement in favour of non-exhaustivity. Finally  $\langle 0, 1, \text{acc-} \rangle$  is judged as better suiting the dialogue than  $\langle 0, 1, 1 \rangle$  in a

marginally significant way with  $p < 0.1$ . This result does not reflect our expectations.

## 5. Conclusion

We presented an empirical study on the role of the variation of accent type and of context on the exhaustive interpretation of answers. For these purposes we used pictures intended to show either the exhaustive or non-exhaustive reading. It was assumed that the degree of the congruity between the contextual information, the picture and the prosodic information effects the pragmatic focus interpretation. When we present the picture illustrating the exhaustive interpretation our data do neither show a contextual influence nor a prosodic effect in a significant way. When presenting the picture showing the non-exhaustive reading significant effects of both the context and prosody can be observed, but our data suggest more evidence for the contextual influence.

In future work it would be interesting to broaden the investigation of pragmatic focus interpretation to the visual modality. Since we found for pragmatic focus production that visual cues accompany the acoustic cues [14], it would be interesting to test which role audiovisual prosody plays for pragmatic focus interpretation.

## 6. References

- [1] Gussenhoven, C. (2004). *The Phonology of Tone and Intonation*. Cambridge, UK: Cambridge University Press.
- [2] Fisseni, B. (2011). *Focus: Interpretation? Empirical Investigations on Focus Interpretation*. Duisburg: Universitätsverlag Rhein-Ruhr.
- [3] Kadmon, N. (2001). *Formal Pragmatics. Semantics, Pragmatics, Presupposition, and Focus*. Oxford, UK: Blackwell.
- [4] Groenendijk, J., Stokhof, M. (1984). *Studies on the Semantics of Questions and the Pragmatics of Answers*. Dissertation, University of Amsterdam.
- [5] Rooth, M. (1992). A theory of focus interpretation, in: *Natural Language Semantics*, 1, 75-116.
- [6] van Rooij, R., Schulz, K. (2006). Pragmatic Meaning and Non-monotonic Reasoning: The Case of Exhaustive Interpretation, in: *Linguistics and Philosophy*, 29, 205-250.
- [7] Skopeteas, S., Fanselow, G. (2009). Exhaustive identification of focused referents and contextual conditions. *Paper presented at DGfS 2009*, Osnabrück, Germany.
- [8] Smith, V., Clark, H. (1993). On the course of answering questions, in: *Journal of Memory and Language*, 32, 25-38.
- [9] Brennan, S.E., Williams, M. (1995). The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers, in: *Journal of Memory and Language*, 34, 383-398.
- [10] Ward, G., Hirschberg, J. (1985). Implicating Uncertainty: The Pragmatics of Fall-Rise Intonation, in: *Language*, 61(4), 747-776.
- [11] Wollermann, C., Schröder, B. (2008). Does Uncertainty Effect the Case of Exhaustive Interpretation?, in: *Proceedings of the ISCA Tutorial and Research Workshop on Experimental Linguistics*, Athens, Greece, 233-236.
- [12] Wollermann, C., Schröder, B. (2008). Certainty, Context and Exhaustivity of Answers. *Paper presented at the workshop on Speech and Face to Face communication*, Grenoble, France.
- [13] Wollermann, C., Schade, U., Fisseni, B., Schröder, B. (2010). Accentuation, Uncertainty and Exhaustivity - Towards a Model of Pragmatic Focus Interpretation, in: *Proceedings of Speech Prosody 2010*, 100063:1-4, Chicago, IL.
- [14] Wollermann, C., Schröder, B. (2009). Effects of Exhaustivity and Uncertainty on Audiovisual Focus Production, in: *Proceedings of the International Conference on Auditory-Visual Speech Processing*, Norwich, UK, 145-150.
- [15] Wollermann, C., Schröder, B., Schade, U. (2010). Audiovisual Congruence and Pragmatic Focus Marking, in: *Proceedings of Interspeech 2010*, Makhuari, Japan, 1137-1140.
- [16] Gussenhoven, C. (2002). Intonation and interpretation: Phonetics and Phonology, in: *Proceedings of Speech Prosody 2002*. Aix-en-Provence, France, 47-57.
- [17] Gerken, L.A., Shady, M.E. (1996). The picture selection task, in: McDaniel, D., McKee, C., Cairns, H.S. (eds.): *Methods for assessing childrens syntax*. Cambridge, MA: MIT Press, 125-145.