Effects of Exhaustivity and Uncertainty on Audiovisual Focus Production

Charlotte Wollermann and Bernhard Schröder

Institute of Communication Sciences, University of Bonn, Germany
Institute of German Linguistics, University of Duisburg-Essen, Germany
cwo@ifk.uni-bonn.de, bernhard.schroeder@uni-due.de

Abstract

This paper presents an experimental study on the role of uncertainty and exhaustivity for audiovisual pragmatic focus production. We investigate whether audiovisual cues of uncertainty appear more often for the production of non-exhaustive answers than for exhaustive answers. Results for the audio modality suggest that rising intonation occurs more often in connection with non-exhaustivity as with exhaustive answers, but no effects on fillers and pauses are observed. For the visual channel eyebrow and head movements accompanying the speech signal are found, but effects are relatively weak.

Index Terms: audiovisual prosody, pragmatic focus production, uncertainty, exhaustivity, question-answering

1 Introduction

The current paper presents an experimental study on the role of uncertainty and exhaustivity on audiovisual pragmatic focus production.[1] Our research questions are the following: 1) Which audiovisual cues do speakers use when uttering (non-)exhaustive answers? 2) Is there a correlation between audiovisual cues of non-exhaustivity and audiovisual cues of uncertainty?

First, a brief overview of the theory of *focus* is given (see section 1.1) and afterwards of previous studies of *uncertainty* (see section 1.2) and also of *audivisual prosody* (see section 1.3). In section 1.4 we establish a connection between these fields and give a motivation for the current study. Section 2 reports on our previous studies and in section 3 we describe our experimental study. In section 4 we discuss our findings.

1.1 Research on focus

In this section we present the notation of *focus* and the closely linked concept of *exhaustivity*.

1.1.1 Definition

Focus is a concept that refers to the fact that pitch accent correlates with salient information in utterances, whereas given information is often deaccented (e.g. [1, 2]). This assumption holds for West Germanic languages like English, German or Dutch. Various focus phenomena occur in natural language.[2] Even though the no-

tation *focus* is used in a variety of disciplines incorporating many different intuitions it is widely accepted that focus can be defined in terms of question-answering: "I believe that the intuition that focus is the answer to the question being addressed is a basic and crucial one." [5: 261] For our study we use this concept of focus, which is also referred to as *pragmatic focus*.

1.1.2 Exhaustivity

In the framework of semantic-pragmatic focus theories it is often suggested that focus licenses an exhaustive interpretation of the focussed constituent(s) (e.g. [6]). Example 1 taken from [7] should make this clear.

(1a) Who passed the examination?

(1b) John and Mary.

(1b) serves as answer to (1a). If the hearer concludes from (1b) that John and Mary are the only persons who passed the examination, the answer is interpreted exhaustively. If there are (or may be) also other people who passed the examination, the interpretation is non-exhaustive. The crucial factor for exhaustivity is the knowledge about the situation under discussion, which is ascribed to the speaker by the hearer. Semantic-pragmatic focus theories (e.g. [6, 8]) expect that in a context of a question, accent is highly correlated with focus. This focus detection should facilitate an exhaustive interpretation. In these theories, contextual influence is generally not considered, or the immediate context, i.e. the type of question, is discussed exclusively.

In [4] these predictions about prosody were tested empirically for the audio channel. It was found that not only accent, but also the expectancies of the hearer, the sensitization for focus phenomena and contextual factors are relevant for focus interpretation.

However, it is less clear to what extent these theoretical predictions hold for the *audiovisual* channel.

1.2 Research on uncertainty

The question of how speakers and listeners signal and detect *uncertainty* in face to face communication has been investigated in several studies. The work of [9] investigated memory processes in question–answering situations. Results showed that uncertainty is expressed by using phrases like "I guess" and also by means of prosody. *Rising intonation* and *delay* were found as such relevant cues. Further, [10] showed that *intonation*, *form of answers*, *pauses* and *fillers* contribute to the perception of uncertainty. *Pauses* and *fillers* have also been found as relevant cues with respect to self-repair in speech, also referred to as *c-repairs* (e.g. [11]). These repairs occur if the speaker recognises and corrects the slip of the tongue, but is not certain what she is going

---

[2]A terminological overview is for instance given in [3] and [4].

to produce. A connectionist model of such repairs can be found in [12]. Moreover, in [13] the role of visual information was investigated with respect to both production and perception of uncertainty. *Smiles* and *funny faces*, i.e. marked facial expressions, were found as relevant cues.

The mentioned studies focused on the role of uncertainty for human-human interaction. On the side of human-machine interaction, [14] modelled various degrees of uncertainty by means of articulatory speech synthesis. Results showed that listeners are generally able to differentiate between the different intended levels. With respect to visual speech synthesis, e.g. [15] expressed uncertainty by variation of head and eyebrow movement with a Talking Head.

However, to our knowledge there is barely any empirical data and research with respect to the role of uncertainty for *pragmatic focus production* in *natural* speech.

1.3   Research on audiovisual prosody

The experimental study of [16] showed that visual information influences speech perception in a previously unrecognised manner. Since then, the McGurk effect could have been observed in a large number of studies on different fields like psychology, neurology, linguistics etc.. However, there is strong empirical evidence that visual information plays an important role for speech perception in different languages.

1.3.1   Audiovisual speech synthesis

It has been found that visual information is helpful for the development of *Talking Heads*. An overview of the main approaches for synthesizing faces is for instance given in [17, 18]. Providing additional visual information for communication is helpful when background noise is present, e.g. at airports, train stations, museums or shopping centers. Further it can help hearing-impaired people as a tool for interactive and adaptive training of speechreading (cf. [19]: 54). According to [20: 260] Talking Heads might also be used as a tool for research on audiovisual speech perception.

However, for creating believable Talking Heads the study of dialogic interaction in human-human communication constitutes an important foundation. This basic research is necessary in order to develop dialogue systems with deep linguistic expressive capabilities on the syntactic, semantic and *pragmatic* level.

1.3.2   Visual prosody and information structure

Several studies investigated the interaction between intonation and head/eyebrow movement. According to the *metaphor of up and down* (e.g. [21]: 202 ff) rising or falling of pitch is accompanied by rising or falling of gesture and facial expression, e.g. eyebrow movement. Empirical evidence for this assumption was found in [22]: The authors report an interplay between rising of eyebrows and rising of pitch which is linguistically and communicationally motivated. Further, the data of [23] suggest a correlation between head movement and pitch. It has also been shown that visual cues effect the perception of prominence (e.g. [24]).

Contrary, there is limited research on the role of audiovisual prosody with respect to *focus*. For instance [25] tested the relevance of audiovisual cues for production and perception of *contrastive focus* in French. Results suggest that visual prosodic cues being identified by means of an articulatory analysis are relevant

for perception. On the side of visual speech synthesis, [26] investigated whether pitch accents and eyebrow movements expressed by a Talking Head effect the perception of *focus of attention*. Results show evidence for eyebrow movements, but the impact of accent is stronger. However, to our knowledge, there is a lack of empirical investigation on the interaction of *audiovisual prosody* and again *pragmatic focus production* although it might have direct impact on interpretation.

1.4   Connecting exhaustivity, uncertainty and audiovisual prosody

As noted in section 1.1.2, exhaustive interpretation depends on the knowledge about the situation in question. Following [27] we say that a sentence s is interpreted exhaustively if it is implicated that the speaker does not know of any other entities having the property in question than those mentioned in s. We speak of a strong reading, if we furthermore implicate that the speaker knows that no other entities have the property in question, while in a weak reading does not know if there are other entities with the property in question.

In our approach [28, 29] the degree of certainty refers to the exhaustive interpretation of an utterance, which might be intended as epistemically strong or weak implicature. With respect to production, and the current study, we expect that non-exhaustivity given by the context conducts the speaker to use audiovisual cues of uncertainty, whereas an intended exhaustive context rather favours realisation of audiovisual cues of certainty.

2   Previous studies

In our previous study on *pragmatic focus interpretation* [28] we tested whether *rising intonation* as prosodic indicator of uncertainty effects exhaustive interpretation of answers and what role type of question plays. Focus interpretation was tested by using pictures illustrating the different intended readings. From the subjects' choice of the picture we inferred the preference for interpretation. Results suggested that the exhaustive interpretation was strongly preferred in our scenario; only weak effects of both intonation and immediate context could be observed.

The goal of [29] was to find out which factors contribute to non-exhaustivity. We tested if stronger cues of uncertainty, i.e. *rising intonation* combined with *pauses* and/or richer variation of *macro context* facilitate/s non-exhaustivity. Results again showed that the exhaustive reading was generally preferred. However, we could observe that stimuli embedded into context intended to favour non-exhaustivity were judged as better suiting the picture illustrating non-exhaustivity than stimuli embedded into context intended to favour exhaustivity. By contrast the influence of prosody was relatively weak. We concluded that micro and macro context seem to play a more crucial role for exhaustivity than theoretically expected. Our findings confirm the results of [4] (see also section 1.1.2).

3   Experimental study

In this section, we present the goal, the method and the results of the current empirical study.

3.1 Goal

As noted in section 1 our research questions are the following: 1) Which audiovisual cues do speakers use when uttering (non-) exhaustive answers? 2) Is there a correlation between audiovisual cues of non-exhaustivity and audiovisual cues of uncertainty?

3.2 Method

In the following we report on the material, the experimental procedure and the analysis of the data.

3.2.1 Material

Our stimuli consist of six question-answer pairs embedded into different dialogues. The scenario is a fictitious party where different groups of students do different things. For every action, there is a question asking for the agent and an answer providing the information. The subject of the answer is either one group of students or two groups, which is/are also the focus of the answer. We refer to the first case as focus sentence with one NP (noun phrase) and to the second case as focus sentence with two NPs. The proportion is 50:50. Each NP consists of a determiner and a noun.

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

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

(2c) Die Mathematiker sind zu spät gekommen. *The mathematicians were late.*

(2d) Das waren die Einzigen, da bin ich mir sicher. *I am certain that they were the only ones.*

(2e) Wenn ich mich nicht täusche, waren die Linguisten auch zu spät. *If I am not wrong, the linguists were also late.*

For every dialogue there are two variants of context: (I) One student group is salient during the whole dialogue. The question-answer pair consists of a specific question to which the answer is congruent (see 2a). The answer (2c) is followed by an utterance indicating that the speaker is certain that no other group performed the action under discussion (2d). (II) At the beginning of the dialogue a "competing" student group usually carrying out the action under discussion is introduced and a general question follows (see 2b). The answer (2c) is followed by an utterance indicating that the speaker is uncertain whether another group also performed the relevant action (2e).

We assume that variant II more often correlates with audiovisual cues of uncertainty, i.e. more intonational *rises* for the focus of the answer and at intermediate phrase-final position and in general more *pauses* and *fillers*. We expect a correlation of *rising of pitch* and a *rising of eyebrow-/head movement*. Furthermore, we expect for the visual channel *smiles* and *funny faces*. Contrary, for variant I we hypothesize generally a *lowering of pitch* and of *eyebrow/head movement* for the focus of the answer and at phrase-final positions.

3.2.2 Procedure

Nine speakers (1 male, 8 female) were recorded, all of them are students from the University of Bonn and native speakers of German. We used prescripted dialogues in our scenario since we give the checkability and thus comparability of the data the top priority. Subjects were instructed to read the six stimulus-dialogues

and three further filler-dialogues loud, presented in a random order. We did not give any guidelines how to realise the text. For all speakers the dialogue partner was always the same female speaker. In sum, there were 81 dialogues.

3.2.3 Multimodal annotation and statistical analysis

For the annotation, the answer with the focus and the sentence indicating (un)certainty were considered. In particular we analysed NPs and end of intermediate phrases. The audio data were annotated by the first author and two phoneticians independently from each other to ensure intercoder validity with the help of Praat.[4] For these purposes the annotation scheme GToBI according to [30] was used. For the visual annotation we used the system ELAN.[5] Again there were three different coders. For the statistical analysis we employed the Fisher's exact test.

3.3 Results

In this section we present the results for both modalities.

3.3.1 Audio condition

Table 1 shows the accent types according to [30] which were found in our data. In addition we used acc for the absence of an accent and edge tone for the absence of an edge tone.

Table 1: GToBI notation according to [30]

Abbreviation	Explanation
H*	peak accent
L*	low accent
L+H*	rising peak accent
L*+H	low rising accent
L-	low edge tone, intermediate phrase final
H-	high edge tone, intermediate phrase final

Results of the audio annotation are presented in figures 1-5. For the comparisons between accent types we used a 2x2 contingency table each time. Firstly, we differentiate between answers containing one or two NP/s as focus while looking at the first NP. For sentences with one NP the following can be observed (see fig. 1): Both L+H* and H* are found more often for variant I than for variant II, but the statistical testing yields no significant difference ($p > 0.05$). L* occurs once for variant II; no accent is found one time for variant I and four times for variant II ($p > 0.05$). For sentences with two NPs the following can be remarked for the first NP (see fig. 2): L+H* is the kind of accent which is found most often in both variants, followed by H* which occurs five times exclusively in variant II. This difference is significant with $p < 0.05$. No accent is found in two cases for variant I and in one case for variant II ($p > 0.05$).

In a further step we regard the cases in which we have two NPs and focus on the second NP (see fig. 3). L+H* can be found in five cases for variant I and seven cases for variant II; no accent occurs six times for variant I and in four cases for the second variant. This difference is not significant with $p > 0.05$. Furthermore, L* and also H* appear once for variant I, H* three times for variant II. This occurrence is not significant when compared to all other accent types ($p > 0.05$).
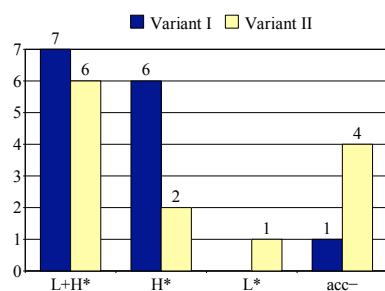
---

[4] http://www.fon.hum.uva.nl/praat
[5] http://www.lat-mpi.eu/tools/elan

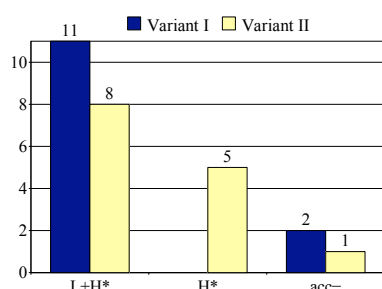Figure 1: Results for the answer with the focus, sentences with one NP; vI: N=14, vII: N=13.

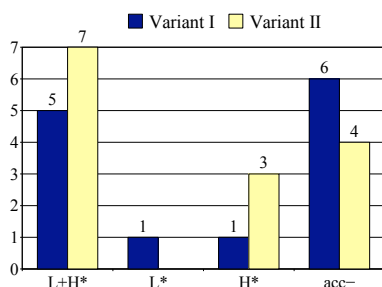Figure 2: Results for answer with the focus, sentences with two NPs, first NP; vI: N=13, vII: N=14.

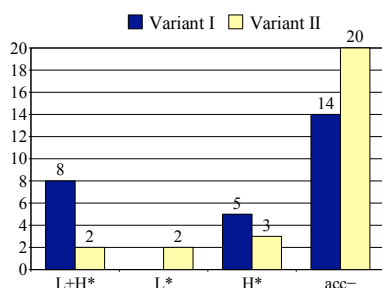Figure 3: Results for answer with the focus, sentences with two NPs, second NP; vI: N=13, vII: N=14.

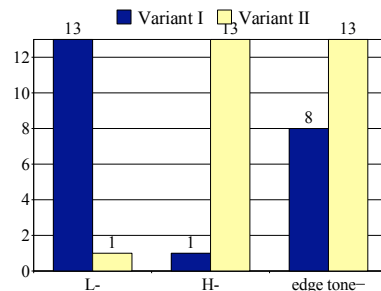Figure 4: Results for sentence indicating (un)certainty, NP, each variant: N=27.

Figure 5: Results for sentence indicating (un)certainty, intermediate phrase; vI: N=22; vII: N=27.

After examining the answer with the focus we account on the sentence indicating (un)certainty whether no/a further group performed the action under discussion. When we do not make a distinction whether the (un)certainty sentence was preceded by a focus sentence containing one or two NP/s the following with respect to the realisation of the NP in the (un)certainty sentence is observable (see fig. 4): No accent is generally found most often for both variants, followed by L+H*. The opposition of these two accent types shows a marginal significant difference with p 0.1. L* occurs twice for variant II, but it is absent in variant I. The comparision between L+H* vs. L* shows a marginal significant difference with p 0.1. Further, H* is found five times for variant I and three times for variant II. This occurence is not significant when compared to the other prosodic realisations (p 0.05).

Regarding the intermediate phrase of the sentence indicating (un)certainty (see fig. 5) we observe that L- appears mostly in variant I, whereas H- can be found more often in variant II. The comparison shows a highly significant difference (p 0.0001). Further, absence of an edge tone occurs more often in variant II than in variant I in a (marginal) significant way (H- vs. edge tone : p 0.1; L- vs. edge tone : p 0.01). With respect to the presence of fillers and pauses no significant difference between variants is found for both the answer with the focus and the sentence indicating (un)certainty.

3.3.2 Visual condition

Table 2 shows our annotation scheme for the visual modality. Again, we use a 2x2 contingency table, but this time we oppose presence to absence of each prosodic realisation. As for the audio condition we differentiate between focus sentences with one or two NP/s. For those cases in which we have two NPs as focus of the answer our data show the following results: For variant I (see fig. 6) in eight of thirteen cases we find eyebrow or head movement on the NP: L+H* combined with a lowered head (L+H* & HL) and also together with raised eyebrows (L+H* & ER) appears each time twice in our data. Further, the following prosodic realisations are found once: L+H* accompanied by raised eyebrows and a lowered head (L+H* & ER & HL), L+H* in combination with a raised head (L+H* & HR), L+H* together with frowned eyebrows (L+H* & EF) and raised eyebrows per se (acc & ER).

The results for variant II are illustrated in figure 7. In nine of fourteen cases the speakers use eyebrow or head movement for realising the NP. H* joined with raised eyebrows (H* & ER) is found three times. The following combinations occur each time

Table 2: Annotation scheme for the visual information

Abbreviation	Explanation
E	eyebrows
H	head
R	raising
L	lowering
F	frown
S	shacking
N	nodding
E	absence of eyebrow movement
H	absence of head movement

once: L+H* accompanied by raised eyebrows and a lowered head (L+H* & ER & HL), the same accent type together with a lowered head (L+H* & HL), L+H* combined with a head nodding (L+H* & HN) and also with raised and then frowned eyebrows (L+H* & ER/F). Beyond that, H* paired with frown eyebrows and a raised head (H* & EF & HR) and also with raised eyebrows (H* & ER) is realised each time once. The statistical analysis shows that H* accompanied by either raised eyebrows (H* & ER) or a raised head (H* & EF & HR) occurs more often in variant II than in variant I. This difference is marginal significant with p 0.1. The other comparisons between the different prosodic realisations do not yield to a significant difference with p 0.05 each time.

For the sentence indicating (un)certainty our data show the following for the realisation of the NP, expressing the inclusion/exclusion of alternatives: For those sentences which are preceded by a focus sentence with two NPs effects on the production of head movement occur. For variant I (see fig. 8) we find in eight of thirteen cases head movement for the realisation of the NP. L+H* together with a lowered head (L+H* & HL), H* accompanied by a head shacking (H* & HS) and also an exclusive head lowering (acc & HL) are found each time twice. In addition, H* combined with a lowered head (H* & HL) and a head shacking per se (acc & HS) is realised once.

The results for variant II are shown in figure 9. In the vast majority the speakers do not realise a head movement when uttering the NP: Only in one case L* is accompanied by a head shacking (L* & HS). The statistical analysis brings to light that a head lowering regardless of the accompanying accent type (L+H* or H* or acc & HL) can be observed more often in variant I than in variant II. This difference is significant with p 0.05. Finally, it has to be noticed that we can neither observe a significant appearance of funny faces nor smiles in favour of variant II.
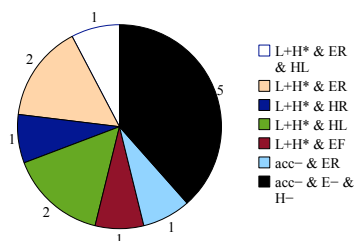

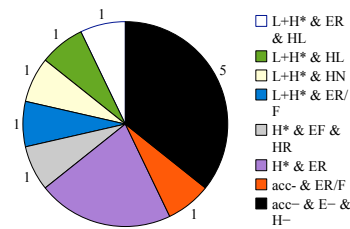Figure 6: Focus sentences with two NPs, first NP; variant I, N=13.

Figure 7: Focus sentences with two NPs, first NP; variant II, N=14.
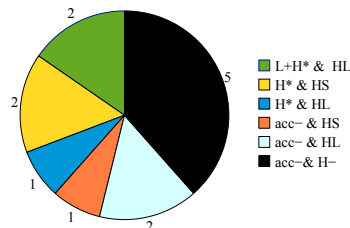

Figure 8: Sentences indicating (un)certainty preceded by focus sentence with two NPs; NP; variant I, N=13.
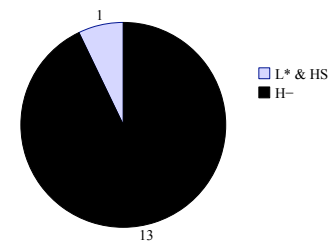

Figure 9: Sentences indicating (un)certainty preceded by focus sentence with two NPs; NP; variant II, N=14.

4   Conclusions

In our study we tested to what extent exhaustivity and uncertainty effect audiovisual pragmatic focus production. Our data suggest in general that there are *intonational* variations with respect to the production of *(non-)exhaustive* answers which correlate with prosodic contours of *(un)certainty* and *continuation* and *finality* respectively. These findings can be explained by the *frequency code* and *production code* according to [31]: High pitch corresponds with uncertainty and also indicates continuation on the informational level, whereas low pitch corresponds to certainty and finality. But we could not find an effect on *pauses* and *fillers*. Regarding the visual channel, we could observe impact of *non-exhaustivity* on *raising of eyebrows* or *head* accompanying high pitch. Further, our data suggest that a lowering of head per se

might serve as an indicator of finality. But our visual data do not show evidence for signals of uncertainty, i.e. *smiles* and *funny faces*.

For our future work we consider it as important to use a scenario with spontaneous speech so that speakers use (un)certainty as paralinguistic expression when producing pragmatic focus. Stimuli used in previous studies on the production of uncertainty consisted typically of one-word-sentences. In order to investigate pragmatic focus constructions adequately and to evoke a contrast between new and given information, it would be necessary that the speaker utters a whole sentence. Here, it would be important that comparability of the stimuli is ensured even though the speech is spontaneous.

References

[1] D. Bolinger, "A theory of pitch accent in English", in: *Word*, 14, pp. 109-149, 1958.

[2] D. Büring, *The meaning of topic and focus: The 59th street bridge accent*, London: Routledge, 1997.

[3] M. Krifka, "Basic Notions of Information Structure", in *Working Papers of the SFB632, Interdisciplinary Studies on Information Structure (ISIS) 6*, C. Féry, G. Fanselow and M. Krifka, Eds., Potsdam: Universitätsverlag Potsdam, pp. 13-56, 2007.

[4] B. Fisseni, *Focus: Interpretation*, Dissertation, University of Duisburg-Essen, to appear.

[5] N. Kadmon, *Formal Pragmatics. Semantics, Pragmatics, Presupposition, and Focus*, Oxford: Blackwell, 2001.

[6] J. Groenendijk and M. Stokhof, *Studies on the Semantics of Questions and the Pragmatics of Answers*, Dissertation, University of Amsterdam, 1984.

[7] R. van Rooij and K. Schulz, "Pragmatic Meaning and Non-monotonic Reasoning: The Case of Exhaustive Interpretation", in *Linguistics and Philosophy*, 29, pp. 205-250, 2006.

[8] M. Rooth, "A theory of focus interpretation", in *Natural Language Semantics*, pp. 75-116, 1992.

[9] S.E. Brennan and M.Williams, "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, pp. 383-398, 1995.

[10] V. Smith and H. Clark, "On the course of answering questions", in *Journal of Memory and Language*, 32, pp. 25-38, 1993.

[11] W.J.M. Levelt, "Monitoring and self-repair in speech", in *Cognition*, 14, pp. 41-104, 1983.

[12] U. Schade and H.-J. Eikmeyer, "wahrscheinlich sind meine Beispiele soo sprunghaft und und und eh ehm zu zu telegraph - Konnektionistische Modellierung von "covert repairs"", in *GWAI-91 1. Fachtagung für Künstliche Intelligenz*, Th. Christaller, Ed., Berlin: Springer Verlag, 1991, pp. 264-272.

[13] M. Swerts and E. Krahmer, "Audiovisual prosody and feeling of knowing", in *Journal of Memory and Language*, 53(1), pp. 81-94, 2005.

[14] C. Wollermann and E. Lasarcyk, "Modeling and Perceiving of (Un)Certainty in Articulatory Speech Synthesis", in *Proceedings of the 6th ISCA Workshop on Speech Synthesis*, Bonn, Germany, 2007, pp. 40-45.

[15] E. Marsi and F. van Rooden, "Expressing Uncertainty with a Talking Head in a Multimodal Question-Answering System", in *Proceedings of the Workshop on Multimodal Output Generation*, Aberdeen, UK, 2007, pp. 105-116.

[16] H. McGurk and J. McDonald, "Hearing lips and seeing voices", in: *Nature*, 264, pp. 746-749, 1976.

[17] G. Bailly, M. Bérar, F. Elisei and M. Odisio, "Audiovisual speech synthesis", in *International Journal of Speech Technology*, 6, pp. 331-346, 2003.

[18] J. Beskow, *Talking heads - Models and applications for multimodal speech synthesis*. Doctoral dissertation, KTH, Stockholm, Sweden, 2003.

[19] J. Beskow, J., "Talking heads - communication, articulation and animation", in *TMH-QPSR*, 37(2), 1996, pp. 053-056.

[20] D.W. Massaro and M.M. Cohen, "Perception of synthesized audible and visible speech", in *Psychological Science*, 1, 55-63, 1990.

[21] D. Bolinger, *Intonation and its parts - Melody in spoken English*, Stanford, CA: Stanford University Press, 1986.

[22] C. Cavé, I. Guaitella, R. Bertrand, S. Santi, F. Harlay and R. Espesser, "About the relationship between eyebrow movements and F variations", in *Proceedings of the ICSLP*, Philadelphia, PA, 1996, pp. 2175-2179.

[23] K. G. Munhall, J. A. Jones, D. E. Callan, T. Kuratate and E. Vatikiotis-Bateson, "Visual prosody and speech intelligibility: Head movement improves auditory speech perception", in *Psychological Science*, 15, pp. 133-137, 2004.

[24] E. Krahmer and M. Swerts, "Testing the effects of audiovisual cues to prominence via a reaction time experiment", in *Proceedings of Interspeech*, Pittsburgh, PA, 2006.

[25] M. Dohen, H. Loevenbruck, M.-A. Cathiard, J.-L. Schwartz, "Visual perception of contrastive focus in reiterant French speech", in *Speech Communication*, 44, pp. 155-172, 2004.

[26] E. Krahmer, Zs. Ruttkay, M. Swerts and W. Wesselink, "Pitch, Eyebrows and the Perception of Focus", in: *Proceedings of Speech Prosody 2002*, Aix en Provence, France, 2002, pp. 443-446.

[27] R. van Rooij and K. Schulz, "Exhaustive interpretation of complex sentences", in *Journal of Logic, Language and Information*, 13, pp. 491-519, 2004.

[28] C. Wollermann and B. Schröder, "Does Uncertainty Effect the Case of Exhaustive Interpretation?", in *Proceedings of the ISCA Tutorial and Research Workshop on Experimental Linguistics*, Athens, Greece, 2008, pp. 233-236.

[29] C. Wollermann and B. Schröder, "Certainty, context and exhaustivity of answers" presented at the *Workshop on Speech and Face to Face communication*, Grenoble, France, 2008.

[30] M. Grice and S. Baumann, "Deutsche Intonation und GToBI", in *Linguistische Berichte*, 191, 267-298, 2002.

[31] C. Gussenhoven, *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press, 2004.