

Who likes liver? How German speakers use prosody to mark questions as rhetorical

Jana Neitsch, Daniela Wochner, Katharina Zahner, Nicole Dehé, Bettina Braun
University of Konstanz

While an *information-seeking question* (ISQ) elicits information [1, 2], the implied answer to a *rhetorical question* (RQ) is usually part of the common ground (e.g. [1, 3]). RQs have as yet mainly been the subject of pragmatic and semantic investigations, but research on their prosodic realisation is still rare. In a first study on the production of RQs in German, [4] found that RQs differed prosodically from ISQs in *final boundary tone*, *pitch accent type*, *duration*, *voice quality*, and *initial pitch*. Compared to [4], the present study controls the discourse status of the syntactic object, such that it is always discourse given, as well as the inherent rhetoricity between *wh*- and polar questions by keeping the contexts the same.

We constructed 22 context-question quadruplets (Tab. 1), manipulating question type (*wh* vs. polar; half of the contexts were presented with polar and half with *wh*-questions) and illocution type of context (ISQ vs. RQ; distributed within-subjects). Forty participants were randomly assigned to one of two lists and tested individually. After having read the context silently from a screen, they produced the target question as naturally as possible. So far, data of 12 participants (\bar{X} =21.7 years, SD=2.27 years, 2 male) have been analysed (n=512, *wh*: 129 ISQs, 130 RQs; polar: 127 ISQs, 126 RQs). Using *Praat*, two trained annotators labelled the words (Tier 1), accented syllable(s) (Tier 4), pitch accents and boundary tones (following GToBI [5], Tier 5; see Fig. 1). In case of disagreement, a consensus label was found. We also extracted vocal effort (H1*-A3*, [6]) at three vowels of interest (Tier 3, Fig. 1).

Boundary tones: For polar questions, RQs were predominantly realised with a high plateau (H-%: 74%), while ISQs were mostly produced with a high rise (H-^H%: 87%). For *wh*-questions, RQs mostly show low boundary tones (L-%: 94%), while ISQs were more variable: high rises (33%), low rises (19%), and low boundary tones (45%); Tab. 2.

Nuclear pitch accents: For polar questions, both ISQs and RQs were most often realised with a low-pitched accent (L*, ISQ: 84%, RQ: 77%), followed by a rise (high rise or high plateau). For *wh*-questions, there is a difference in nuclear pitch accents, however. RQs show a higher proportion of L*+H nuclear accents (55%) than ISQs (3%), while ISQs show a higher proportion of L+H* nuclear accents (43%) than RQs (22%). Moreover, in ISQs, L* (28%) and H+!H* (15%) are more common than in RQs (L*: 9%, H+!H*: 2%); Tab. 2.

Acoustic measurements: Tab. 2 summarizes the means of RQs and ISQs in the acoustic variables *initial pitch*, *duration*, and *vocal effort*, listed by question type (p-values in brackets; significance levels were corrected according to [7]). Our findings show that RQs, compared to ISQs, have a longer utterance duration (for both *wh*- and polar questions), longer normalised durations of the object (for both *wh*- and polar questions), but shorter durations of the first word (for *wh*-questions only); RQs show a tendency for lower initial pitch (for *wh*-questions only) and breathier voice quality in the first measured vowel (for *wh*-questions only).

These results are comparable to [4] (except for *voice quality* and *initial pitch*), corroborating [4]'s findings in two respects: First, our results suggest that the observed differences can be attributed to illocution type and are not merely an artefact of information-structure. Second, our results support [4]'s findings that the differences between illocution types are realised more clearly in *wh*-questions than in polar questions. Thus, speakers provide clear prosodic cues to mark interrogatives as rhetorical, at least in *wh*-questions. Likewise, in a recent perception study (investigating *accent type* and *voice quality*), we show that these cues are used by listeners in order to identify RQs. We are currently analysing more data to shed more light on the reported findings regarding *initial pitch* and *voice quality* on the one hand, and the difference in the realisation of rhetoricity across question types on the other. We also focus on the production-perception link, testing cues, which have been identified as markers of rhetoricity in production, also in perception.

	ISQ	RQ
polar	“You serve your relatives a meat platter with sauerkraut, liver and many other things. You want to know whether someone would like to have some of it or not.” Mag denn jemand Leber? “Does anyone like liver?”	“During coffee your aunt offers your relatives a tray with liver. However, obviously no one wants this, because it does not go with coffee and cake.” Mag denn jemand Leber? “Does anyone like liver?”
wh	“You serve your relatives a meat platter with sauerkraut, liver and many other things. You want to know which of them would like to have some of it.” Wer mag denn Leber? “Who likes liver?”	“During coffee your aunt offers your relatives a tray with liver. However, obviously no one wants this, because it does not go with coffee and cake.” Wer mag denn Leber? “Who likes liver?”

Tab. 1: Example context-question quadruplet in an information-seeking and rhetorical context. Wh-questions always contained the wh-element ‘who’ and polar questions always included ‘anyone’.

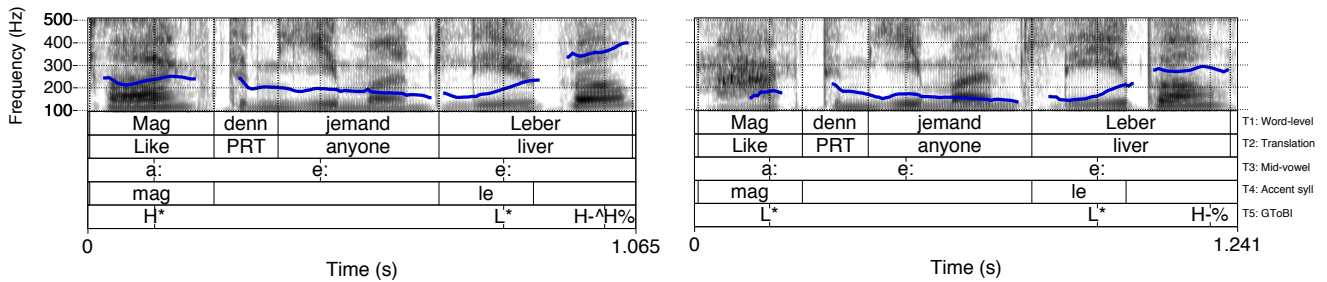


Fig. 1: Example polar question, ISQ (left) and RQ (right), showing the annotation layers.

GToBI labels	polar questions	wh-questions
<i>Final boundary:</i>	H-^H% RQ: 21% vs. ISQ: 87% H-% RQ: 74% vs. ISQ: 4% L-% RQ: 4% vs. ISQ: 4%	RQ: 4% vs. ISQ: 33% RQ: 2% vs. ISQ: 2% RQ: 94% vs. ISQ: 45%
<i>Nuclear accent:</i>	L*+H RQ: 3% vs. ISQ: 0% L+H* RQ: 6% vs. ISQ: 9% L* RQ: 77% vs. ISQ: 84% H+!H* RQ: 0% vs. ISQ: 0%	RQ: 55% vs. ISQ: 3% RQ: 22% vs. ISQ: 43% RQ: 9% vs. ISQ: 28% RQ: 2% vs. ISQ: 15%
Acoustic measurements	polar questions	wh-questions
<i>Initial pitch</i>	RQ: 213Hz vs. ISQ: 224Hz (p=0.16)	RQ: 207Hz vs. ISQ: 216Hz (p=0.08)
<i>Utterance duration</i>	RQ: 1397ms vs. ISQ: 1207ms (p<0.0001)	RQ: 1288ms vs. ISQ: 1085ms (p<0.0001)
<i>Norm. duration (final object) [4]</i>	RQ: 47.6% vs. ISQ: 46.2% (p<0.03)	RQ: 54.7% vs. ISQ: 52.3% (p<0.02)
<i>Norm. duration (first word) [4]</i>	RQ: 18.0% vs. ISQ: 17.4% (p=0.16)	RQ: 11.0% vs. ISQ: 11.7% (p<0.02)
<i>Vocal effort (H1*-A3*)</i>	vowel 1, RQ: 19.9dB vs. ISQ: 18.4dB (p=0.27) vowel 2, RQ: 19.5dB vs. ISQ: 20.2dB (p=0.41) vowel 3, RQ: 21.2dB vs. ISQ: 20.0dB (p=0.44)	vowel 1, RQ: 16.1dB vs. ISQ: 14.5dB (p=0.09) vowel 2, RQ: 17.6dB vs. ISQ: 17.3dB (p=0.65) vowel 3, RQ: 22.1dB vs. ISQ: 22.0dB (p=0.96)

Tab. 2: Mean frequency of occurrence of GToBI labels for RQs and ISQs and means for acoustic measurements (RQs and ISQs, Benjamini-Hochberg corrected significance level).

[1] Groenendijk, J. & M. Stokhof. 1984. *Studies on the semantics of questions and the pragmatics of answers*. Amsterdam: Universiteit van Amsterdam, PhD Thesis.

[2] Meibauer, J. 1986. *Rhetorische Fragen* (Linguistische Arbeiten 167). Berlin: De Gruyter.

[3] Caponigro, I. & J. Sprouse. 2007. Rhetorical questions as questions. *Proc. of Sinn und Bedeutung 11*, 121–133.

[4] Wochner, D., J. Schlegel, N. Dehé & B. Braun. 2015. The prosodic marking of rhetorical questions in German. *Proc. of Interspeech 2015*, 987–991.

[5] Grice, M., S. Baumann, & R. Benz Müller. 2005. German intonation in autosegmental-metrical phonology. In J. Sun-Ah (Ed.). *Prosodic Typology*. Oxford: OUP, 55-83.

[6] Mooshammer, C. 2010. Acoustic and laryngographic measures of the laryngeal reflexes of linguistic prominence and vocal effort in German. *JASA*, 127(2). 1047-1058.

[7] Benjamini, Y. & Y. Hochberg. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1). 289-300.