QUESTIONS AS PROSODIC CONFIGURATIONS: HOW PROSODY AND CONTEXT SHAPE THE MULTIPARAMETRIC ACOUSTIC NATURE OF RHETORICAL QUESTIONS IN GERMAN

Jana Neitsch & Oliver Niebuhr

Centre for Industrial Electronics, Mads Clausen Institute, University of Southern Denmark, Sønderborg/DK
jana@neitsch.com, oniebuhr@mci.sdu.dk

ABSTRACT

Contrary to information-seeking questions (ISQs), rhetorical questions (RQs) occur in non-neutral contexts, e.g., to criticize, challenge, or persuade the addressee. We study the influence of two attitudinal contexts (*disgust* and *mockery*) on the prosodic realization of *wh*-RQs (*Who likes lavender?*) in German relative to a non-attitudinal control condition of ISQs realized in a *sincere interest* context. RQs show context-specific acoustic and phonological differences, but are, overall, more strongly separated from ISQ realizations. Moreover, the way in which RQs and ISQs differ from each other as well as from statements in German suggests that German questions represent a coherent bundle of parameters – a so-called "prosodic construction" – that is constituted to a larger extent by gradual acoustic parameters than by categorical phonological ones.

Keywords: Rhetorical question, attitude, prosodic construction, acoustic analysis, German.

1. INTRODUCTION

In contrast to information-seeking questions (ISQs) with which the interlocutor receives knowledge from the addressee [1,2], rhetorical questions (RQs) imply answers that are already known to all interlocutors [3] and, on that basis, seek the addressee's commitment with respect to the underlying proposition [4].

Isolated from their context, questions like *Who likes playing soccer?* can basically be both an ISQ or an RQ [5]. This also applies to German [6]. Some researchers consider context the most salient determiner of an RQ [4,5,6]. It facilitates the "understanding of the question as not doing questioning" [7,p.55]. Exploiting this role of context, a recent study [8] investigated the prosodic realization of German polar and *wh*-questions in contexts that triggered either a rhetorical or an information-seeking reading of the same question. Compared to ISQs, RQs in this study showed longer utterance durations (esp. of the sentence-final object noun), a breathier voice quality, and a different phonological make-up of their nuclear tunes (polar: ISQ L* H-^H% vs. RQ L*+H H-%; *wh*: ISQ L+H* L-H% and L+H* L-%

vs. RQ L*+H L-%). These findings parallel those of [9] by showing that RQs are, just like ISQs, multiparametric prosodic entities.

Furthermore, RQs and ISQs share the same morpho-syntactic question markers and they both differ from statements along the same prosodic parameters. However, unlike often intuitively assumed by language users, RQs are not prosodically closer to statements than ISQs. The opposite is true. There are several prosodic features (like duration and voice quality) in which RQs differ even more from statements than ISQs [8,9]. One reason for this could be the ironic flavor that characterizes many RQs. Several researchers consider irony a major function of RQs [10,11,12] or state that RQs are frequently used to express ironic remarks [5,11,12] – and exaggerated prosody can be one way of expressing irony. In [13], it is noted that exaggerated pitch patterns "indicate clearly fake enthusiasm" (p.250). Irony is not the only contextual coloring which with RQs occur, though. RQs also enable speakers to criticize or persuade the addressee, or to create humor [12,14,15,16,17,18,19,20].

To sum up, RQs are on the one hand fundamentally *similar* to ISQs. Both show a multi-parametric prosodic marking of the communicative function 'question' at the acoustic level, in addition to using the same morpho-syntactic means. On the other hand, RQs and ISQs are fundamentally *different*. In contrast to ISQs, RQs are a type of question that cannot simply be posed out of the blue and with the neutral intention of information transfer. Rather, RQs are strongly attitudinally and/or expressively colored, depending on the contexts in which they occur, and this also shapes the prosody of RQs.

It is for this reason that RQs are an ideal *testbed* for studying the multi-parametric prosodic nature of the communicative function 'question'. First, by analyzing RQs, we can expect that prosodic differences to statements show up more clearly and, thus, can be analyzed more reliably than for ISQs. Second, comparing RQ realizations across attitudinal contexts to ISQ realizations will reveal which prosodic parameters – acoustic-phonetic as well as phonological ones – vary as a function of the question-statement difference itself and which parameters show a separate variation pattern that is associated with the contextu-

al/attitudinal embedding of RQs rather than with the question-statement difference itself.

The present paper specifically focuses on the second point and represents the authors' initial step into investigating whether the communicative function 'question' represents a *prosodic construction* in German, i.e. a consistently "recurring temporal pattern of prosodic activity" [21,p.2]; and if so, which prosodic parameters are part of this construction. Those parameters that form a prosodic construction would act as a coherent bundle. That is, if one parameter shifts along its phonological or physical axis, then all other parameters would show a parallel shift so that ISQs and RQ types only differ in how strong these parallel shifts are pronounced.

Addressing the issue of questions as prosodic constructions will additionally help figuring out whether RQs and ISQs are clearly separated across contexts and along each prosodic parameter and whether further parameters separating RQs and ISQs emerge when context is taken into account properly.

We analyze the multi-parametric prosodic variation of string-identical questions produced in different contexts that trigger *sincere interest* on the one hand and either *disgust* or *mockery* on the other. Disgust and mockery represent two contextual embeddings for the realization of RQs. Sincere interest represents the contextual reference condition. That is, it elicits ISQs whose realizations are compared to both RQ subtypes. Disgust and mockery were selected, because they represent not only typical but also very different attitudinal/expressive realization contexts for RQs. Our selection of acoustic and phonological prosodic parameters was geared to those parameters that turned out to be relevant in producing and identifying German ISQ and RQ in previous empirical studies (e.g., [8,9]).

Note that due to the pilot nature of our study and the small dataset we restrict ourselves to a descriptive analysis of the acoustic data and report only those differences that appear consistent enough to become statistically significant in the dataset of a larger follow-up study.

2. EXPERIMENT

2.1. Methods

2.1.1. Materials

Six questions were designed. Their wording was equally compatible with RQs and ISQs. Each question started with the *wh*-word *wer* ('who') followed by a verb and the modal particle *denn* (that occurs with both illocution types in German [22]). Questions are concluded by a sentence-final object noun. The object noun was mainly sonorous and consisted

of three syllables with primary lexical stress on the penultimate syllable (e.g., *Wer mag denn Lavendel?* 'Who likes PRT lavender?').

All the six questions were paired with short contexts for *disgust*, *mockery* and *sincere interest*, see Table 1. Each context already introduced the sentence-final object noun so as to make sure that it was prosodically realized as given information in the subsequently elicited question. All contexts were tested in previous production and perception studies and proved to be effective and reliable in eliciting the ISQs and RQs with their targeted attitudes. Contexts were presented as texts.

Table 1: Example contexts for RQ/ISQ elicitation.

Mockery	Disgust	Sincere interest
Your mother tells you that her neighbour read in the newspaper that **lavender** can be eaten. The other day she observed him sitting in the garden and eating the blossoms which **you both find extremely funny.** You say:	You and your friend walk into a perfumery where a woman instantly offers you a new scent with **lavender**. But you and your friend find the scent so gross that you quickly continue walking because **you are feeling nauseous.** You say:	You and your room mates want to plant a small flowerbed in your garden and you have always dreamt of **lavender**. **You are very interested** in whether the others agree with that. You say:
Who likes lavender?		

2.1.2. Participants

So far, five voluntary monolingual native speakers of German have been recorded in a sound attenuated booth (ø = 22.6 years, 2 male).

2.1.3. Procedure

Participants received the context-question pairs in the form of a Power Point presentation. Context and question were displayed on subsequent PPT slides.

Each participant realized each question in all three attitudes. Occurrences of the same question were separated by at least two other context-question pairs. Participants were asked to read the given context silently. Then, upon button press, the corresponding question was presented on the next slide. Participants had to realize the question aloud as naturally as possible and such that it would sound appropriate in the given situational context frame. The experiment was self-paced.

Prior to the experimental session, participants were presented with a written instruction on a sheet of paper. In total, 90 questions were recorded: 5 speakers x 6 questions x 3 attitudes.

Using Praat [23], a phonetically trained annotator labeled the word boundaries in all questions as well as the pitch-accent and boundary-tone categories following the GToBI conventions for German [24]. Word boundaries were used to take two duration

measurements, i.e. absolute utterance duration and absolute constituent duration (of the final object noun). For the analysis of voice quality, the Harmonics-to-Noise Ratio [HNR in dB, e.g., 25] was automatically extracted in the stressed syllables of the *wh*-word, the verb and the sentence-final object noun. HNR values were extracted at the vowel midpoints via Praat's Voice Report [30] in the frequency range between 100-600Hz. Note that vowels with modal voice quality show higher HNR-dB-values than breathy vowels. Finally, in order to determine how high-pitched a question started, the first pitch point of each utterance was automatically extracted and measured in Hz (which was possible because speaker was a within-subjects factor).

3. RESULTS

Figure 1 summarizes schematically – and proportionally for each analyzed parameter – in which way and to what extend the question realizations in the three context conditions differ from one another.

Figure 1: Value ranges of all analyzed parameters from the smallest level or frequency (left) to the highest level or frequency (right). The intermediate level or frequency is displayed proportionally within each value range. VQ means 'voice quality'.
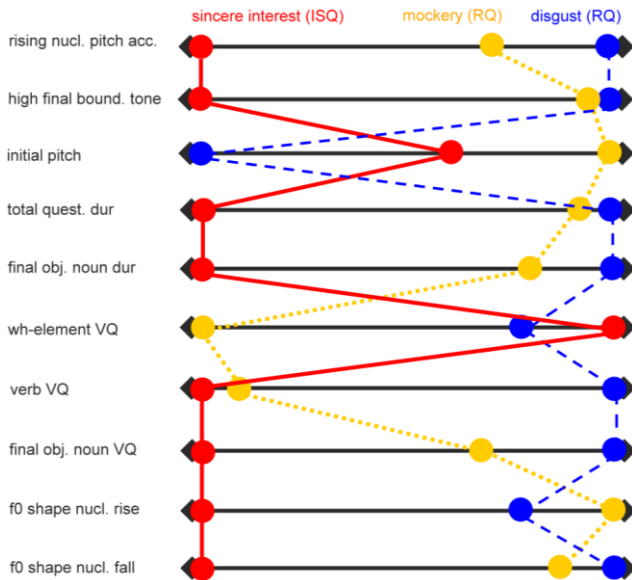
The analysis of the nuclear accents showed that sentence-final object nouns in *disgust* contexts were exclusively realized with rising pitch patterns, i.e. either L*+H (73%) or L+H* (23%). In contrast, more than every fourth object noun in *mockery* contexts had non-rising pitch patterns like L* (10%) and H* (10%). Rising L*+H was also the most frequent pattern in *mockery*, though (63%, L+H*: 10%). The variation in nuclear pitch accent patterns was largest in the *sincere-interest* contexts. Moreover, in these contexts it were the low or falling

pitch-accent patterns that prevailed (L*+H: 37%, L+H*: 10%, L*: 47%, H*: 3%).

The analysis of the final boundary tone showed that questions produced in *disgust* and *mockery* contexts were predominantly realized with a low boundary tone (L%: 90% and 97%), whereas questions from *sincere-interest* contexts mainly had a high boundary tone (H-^H%: 47%, L-H%: 37%).

Regarding pitch level at the beginning of questions, results reveal on average higher initial pitch values in *mockery* contexts (207 Hz) than in *disgust* contexts (195 Hz), while questions from the *sincere-interest* contexts started on average at an intermediate initial pitch level (203 Hz).

Results concerning total question durations show that questions produced in *disgust* contexts were on average realized longer (1384 ms) than those produced in *mockery* contexts (1351 ms). Yet, both *mockery* and *disgust* questions had longer total durations than the questions from the *sincere-interest* contexts (1141 ms). Similar to the findings in [8], context-related duration differences between questions were most strongly pronounced in the sentence-final object noun. The noun was longest in *disgust* contexts (807 ms), had an intermediate duration in the *mockery* contexts (772 ms), and was shortest in the *sincere-interest* contexts (620 ms).

The analysis of voice quality (based on vowel midpoints) yielded the following results: lexical *wh*-elements were produced breathier in the *sincere interest* (13.3 dB) and *disgust* contexts (13.5 dB) than in the *mockery* contexts (14.1 dB). On the verb, however, voice quality was breathiest in the *disgust* contexts (11.6 dB), less breathy in the *mockery* contexts (13.3 dB) and least breathy in the *sincere interest* contexts (13.6 dB). This tripartite voice quality difference became stronger at the end of the questions, i.e. within the sentence-final object noun (*disgust* contexts: 14.5 dB, *mockery* contexts: 15.1 dB, *sincere interest* contexts 16.4 dB).

As a supplement to the above eight parameters that were selected and analyzed on the basis of previous studies [8,9], we used the tonal-target labels L and H from the phonological analysis in order to determine and compare the f0 shapes of the rising and falling slopes in the question-final nuclear tunes. For the sake of comparability, only those nuclear tunes were taken into account that consisted of tri-tonal LHL sequences. The three f0 values of L, H, and L were measured as well as the two f0 values halfway in between the three tones. On this basis, the range proportion measure (R_{prop}) was determined, following [26]. R_{prop} is the F0 range from onset to mid of the rising or falling slope divided by the total range of that slope. The results of our f0-shape analysis showed that nuclear question tunes in *disgust* and

mockery contexts were characterized by clearly concavely shaped f0 rises with R_{prop} values well below 0.5 (0.40 and 0.36), followed by more convexly shaped f0 falls (0.69 and 0.76). The opposite was true for the nuclear question tune in the *sincere-interest* context whose f0 rise was convexly shaped with a R_{prop} value of 0.59, followed by a strongly concavely shaped f0 fall of $R_{prop} = 0.96$.

4. DISCUSSION

Regarding the general differences between RQs (*mockery/disgust*) and ISQs (*sincere interest*), our results are consistent with those of previous studies, e.g. [8], which adds to the validity of the presented data, despite the small sample. In terms of phonological parameters, we found, like [8], that the nuclear pitch accent is mainly rising in RQs, but falling in ISQs. With respect to the final boundary tone, RQs had a low and ISQs mainly a high boundary tone. This supports what was stressed by [9]: final boundary tones have a specific meaning, and this specific meaning is not simply 'question'. That is, there is no straightforward link between sentence mode and final boundary tone. This single fact implies already that also the prosodic differences between RQs and ISQs do not originate from signaling illocution type, but from expressing different attitudinal stances.

Regarding the gradual acoustic-prosodic parameters, we replicated the result of [8] for duration: Both the total durations and the durations of the final object nouns were greater in RQs than in their ISQ counterparts. It also accords with [8] that breathiness dominated in RQs as opposed to ISQs. RQ breathiness was strongest on the verb, but the difference to ISQs was strongest on the final object noun.

Against the background of these empirical agreements with other studies, the present pilot study also sheds initial light on the prosodic-construction nature of German questions. First, our results show that the realization of RQs in German is clearly context-specific. That is, there is not a single stable prosodic RQ profile. Yet, RQs still differ from ISQs along the entire prosodic profile that we analyzed; and the differences between RQs and ISQs were larger than those between the two RQ subtypes *mockery* and *disgust*. Thus, RQs are also not simply "real" questions (ISQs) with one or two minor prosodic differences. Rather, they are prosodically an entirely different phenomenon.

As for the specific nature of this phenomenon, note that, firstly, RQs differ from ISQs along the same prosodic parameters that also distinguish ISQs from regular statements in German [9]. Compared to the latter, ISQs are realized longer, breathier, and with stronger concave rises and convex falls in the LHL sequences of nuclear tunes [27, 28]. RQs seem to exaggerate these differences. That is, RQs seem to be even more question-like than ISQs. This exaggeratedness is also observed in ironic utterances [10-13] and when people use a "borrowed voice" instead of their own voice. A similar mechanism could be at work here, for example, in order to attenuate the expressed *mockery* or *disgust* – in addition to the morphosyntactic question frame itself. That (esp. *mockery*) RQs started with a higher initial pitch than ISQs also supports the idea of an exaggerated RQ prosody, see [9].

Secondly, note with respect to *mockery* and *disgust* that the prosodic profiles of the elicited RQs are not simply one-to-one reflections of the respective related emotions, see, e.g., [29]. Instead, the corresponding related emotion profiles seem to be modified and translated into gradual parallel changes along a parameter profile that distinguishes questions from statements in German.

In conclusion, in all analyzed aspects, RQs and ISQs clearly differ from each other and so do contextual subtypes of RQs. Simultaneously, however, these differences manifest themselves as parallel gradual parameter shifts along the same prosodic profile. Gradual parameter shifts along this profile can also turn questions into statements in German.

We interpret these findings as initial evidence that the analyzed parameters are not changed independently of each other by speakers, but *coherently* in the form of a bundle whose features are all shifted in parallel to a variable extent. In this sense, we found initial evidence for questions as prosodic constructions, i.e. "recurring temporal patterns of prosodic activity" [21,p.2]. The construction involves at least duration, voice quality, and f0 shape. It does not seem to involve the final boundary tone (statements do not differ systematically in this respect from ISQs) and the initial pitch level (no co-variation relative to other parameters, see Fig. 1). The role of the nuclear pitch accent is unclear, but it is generally reasonable to assume that *gradual acoustic* parameters contribute more to a prosodic construction than *categorical phonological* features.

More data, more RQ subtypes, and 'statement' as a second reference condition besides ISQs will help us further support our preliminary results. In a next step the results for German will also be compared to those of an analogous study in Brazilian Portuguese. This study is currently ongoing. Moreover, perception studies are needed in the future to test whether the fine-grained prosodic distinctions between RQ subtypes are perceivable, and whether listeners systematically associate these distinctions with the corresponding attitudinal stances (like *mockery* and *disgust*) and the right illocution type (RQ).

5. REFERENCES

[1] Groenendijk, J., Stokhof, M., 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., Sprouse, J. 2007. Rhetorical Questions as Questions, In: Puig-Waldmüller, E. (ed.), *Proc. of Sinn und Bedeutung* 11, Barcelona, Spain: Universitat Pompeu Fabra, 121-133.

[4] Biezma, M., Rawlins, K. 2016. Rhetorical Questions: Severing asking from questioning. In *Semantics and Linguistic Theory* 27, 302-322.

[5] Frank, J. 1990. You call that a rhetorical question?: Forms and functions of rhetorical questions in conversation. *Journal of Pragmatics, 14*(5), 723-738.

[6] Jung, V., Schrott, A. 2002. A question of time? Question types and speech act shifts from a historical-contrastive perspective: Some examples from Old Spanish and Middle English. In: *PRAGMATICS AND BEYOND NEW SERIES*, 345-372.

[7] Koshik, I. 2003. Wh-questions used as challenges. *Discourse Studies, 5*(1), 51-77.

[8] Braun, B., Dehé, N., Neitsch, J., Wochner, D., Zahner, K. accepted. The prosody of rhetorical and information-seeking questions in German. *Language and Speech.*

[9] Niebuhr, O., Bergherr, J., Huth, S., Lill, C., Neuschulz, J. 2010. Intonationsfragen hinterfragt – Die Vielschichtigkeit der prosodischen Unterschiede zwischen Aussage- und Fragesätzen mit deklarativer Syntax. *Zeitschrift für Dialektologie und Linguistik, 77*(3), 304-346.

[10] Freed, A. F. 1994. The form and function of questions in informal dyadic conversation. *Journal of Pragmatics, 21*(6), 621-644.

[11] Hudson, R. A. 1975. The meaning of questions. *Language*, 1-31.

[12] Oraby, S., Harrison, V., Misra, A., Riloff, E., Walker, M. 2017. Are you serious?: Rhetorical Questions and Sarcasm in Social Media Dialog.

[13] Attardo, S., Eisterhold, J., Hay, J., Poggi, I. 2003. Multimodal markers of irony and sarcasm. *Humor, 16*(2), 243-260.

[14] Cacioppo, J. T., Petty, R. E. 1982. Language variables, attitudes, and persuasion. *Attitudes towards language variation*, 189-207.

[15] Cantor, J. R. (1979). Grammatical variations in persuasion: Effectiveness of four forms of request in door‐to‐door solicitations for funds. *Communications Monographs, 46*(4), 296-305.

[16] Gibbs, R. W. 2000. Irony in talk among friends. *Metaphor and Symbol, 15*(1-2), 5-27.

[17] Ilie, C. 2015. Questions and questioning. *The International Encyclopedia of Language and Social Interaction*. 1-15.

[18] Petty, R. E., Cacioppo, J. T., Heesacker, M. 1981. Effects of rhetorical questions on persuasion: A cognitive response analysis. *Journal of Personality and Social Psychology, 40*(3), 432-440.

[19] Swasy, J. L., Munch, J. M. 1985. Examining the target of receiver elaborations: Rhetorical question effects on source processing and persuasion. *Journal of consumer research, 11*(4), 877-886.

[20] Schaffer, D. 2005. Can rhetorical questions function as retorts?: Is the Pope Catholic? *Journal of Pragmatics, 37*(4), 433-460.

[21] Ward, N. G., Gallardo, P. 2017. Non-native differences in prosodic-construction use. *Dialogue & Discourse, 8*(1), 1-30.

[22] Thurmair, M. 1991. Zum Gebrauch der Modalpartikel 'denn' in Fragesätzen. Eine korpusbasierte Untersuchung. In: E. Klein (ed.), *Betriebslinguistik und Linguistikbetrieb: Akten des 24. Linguistischen Kolloquiums* vol. 1. Linguistische Arbeiten 260. Tübingen: Niemeyer, 377-387.

[23] Boersma, P., Weenink, D. 2017. Praat - doing phonetics by computer (Version Version 5.3.64). Amsterdam: Institute of Phonetics. Retrieved from http://www.praat.org.

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

[25] de Krom, G. 1993. A cepstrum-based technique for determining a harmonics-to-noise ratio in speech signals. *Journal of Speech, Language, and Hearing Research, 36*(2), 254-266.

[26] Dombrowski, E., Niebuhr, O. 2005. Acoustic patterns and communicative functions of phrase-final F0 rises in German: Activating and restricting contours. *Phonetica, 62*(2-4), 176-195.

[27] Petrone, C., Niebuhr, O. 2014. On the intonation of German intonation questions: The role of the prenuclear region. *Language and Speech, 57*(1), 108-146.

[28] Niebuhr, O. 2012. Das ist (k)eine Frage: Phonetische Merkmale in der Identifikation standarddeutscher Deklarativfragen. In: I L. Anderwald (eds.), *Sprachmythen: Fiktion oder Wirklichkeit*. Frankfurt: Peter Lang, 203-222.

[29] Murray, I. R., Arnott, J. L. 1993. Towards the Simulation of Emotion in Synthetic Speech: A Review of the Literature of Human Vocal Emotion. *Journal of Acourstic Society of America 93* (2), 1097-1198.

[30] Batalla, F.N., Márquez, R.G., Belén Peláez González, M. et al. 2014. Acoustic Voice Analysis Using the "Praat" programme: Comparative Study with the "Dr. Speech" Programme. *Acta Otorrinolaringológica Española 65 (3)*, 141-218.