Prosody and breathing

A comparison between rhetorical and informationseeking questions in German and Brazilian Portuguese

Neitsch, Jana; Barbosa, Plinio; Niebuhr, Oliver

Published in:
Proceedings of Interspeech 2020

DOI:
10.21437/Interspeech.2020-1607

Publication date:
2020

Document version:
Final published version

Citation for pulished version (APA):
Neitsch, J., Barbosa, P., & Niebuhr, O. (2020). Prosody and breathing: A comparison between rhetorical and informationseeking questions in German and Brazilian Portuguese. In *Proceedings of Interspeech 2020* (pp. 1863-1867). International Speech Communication Association (ISCA). Proceedings of the International Conference on Spoken Language Processing https://doi.org/10.21437/Interspeech.2020-1607

INTERSPEECH 2020
October 25–29, 2020, Shanghai, China

Prosody and breathing: A comparison between rhetorical and information-seeking questions in German and Brazilian Portuguese

Jana Neitsch[1], *Plinio A. Barbosa*[2], *Oliver Niebuhr*[1]

[1] Centre for Industrial Electronics, Mads Clausen Institute, University of Southern Denmark, Sønderborg, Denmark
[2] Instituto de Estudos da Linguagem, Unicamp, Campinas, Brazil
neitsch@mci.sdu.dk, pabarbosa.unicampbr@gmail.com, olni@sdu.dk

Abstract

Several studies have shown that rhetorical *wh*-questions (RQs) and string-identical information-seeking *wh*-questions (ISQs) are realized with different prosodic characteristics. In contrast to ISQs, RQs have been shown to be phonetically realized with a breathier (i.e., softer) voice quality (e.g., German and English) and longer constituent durations (e.g., German, English, Icelandic). Based on similar results found for different languages, we investigate *wh*-RQs and sting-identical *wh*-ISQs in Brazilian Portuguese (BP) and German (G). We analyze (i) whether specific duration and voice-quality patterns characterize and separate the two illocution types (RQ and ISQ) in BP, and (ii) if direct measures of the respiratory sub-system reveal differences between illocution types, given that breathiness involves greater transglottal air flow which can be observed in the speakers' chest and/or abdomen movement.

Our data suggest that, similar to G, English, and Icelandic, duration and voice quality patterns play a role in the realization of RQs compared to ISQs in BP, reinforcing the assumption that there are cross-linguistically similar phonetic features in the realization of RQs compared to ISQs. We also find that speakers of G breathe in more deeply and dynamically than speakers of BP, suggesting a link between breathing and voice quality.

Index Terms: Rhetorical questions, German, Brazilian Portuguese, voice quality, duration, breathing

1. Introduction

In contrast to information-seeking questions (ISQs) that are generally realized in order to elicit information from the hearer [1,2], rhetorical questions (RQs) provide the hearer with information that is usually common knowledge for all interlocutors [3]. Even if RQs and ISQs are string-identical, they are realized differently as it has been shown in previous studies investigating different languages [4,5,6]. The aim of the present production study is twofold and specifically addresses (i) duration and voice quality in the realization of both *wh*-RQs and string-identical *wh*-ISQs and (ii) the respiratory subsystem during the realization of the two illocution types (i.e., RQ and ISQ).

Regarding aim (i), this paper investigates the phonetic realization of the two illocution types in Brazilian Portuguese (BP) compared to German (G), with particular focus on RQs. Since previous production studies suggest similarities with respect to the phonetic characteristics of RQs in various Germanic languages, the next step is to analyze a Romance language such as BP to further examine whether or not there are cross-linguistic phonetic similarities regarding the realization of RQs. For instance, for G, English and Icelandic, [4,5,6] found that *wh*-RQs such as *Wer mag denn Sellerie?* (Who likes celery?) are produced with significantly longer constituent durations than string-identical *wh*-ISQs. That is particularly true for the sentence-final object noun in G and English – resulting in a longer overall sentence duration for RQs. The authors of [4,5,6] argue that the difference in duration between RQs and ISQs is in line with durational differences between questions (e.g., declarative questions) that tend to be realized with shorter durations than non-questions (e.g., declarative statements, see [7,8]). Given the similar results in different languages, there is reason to assume that BP also shows longer constituent durations in RQs than in ISQs.

Additionally, compared to ISQs, RQs have been shown to be realized with a breathier voice quality in initial position in G and English *wh*-questions (see [4,5]). These findings are based on both the investigation of HNR (Harmonics-to-Noise ratio) and the manual classification of voice quality according to the perception of a trained annotator. With respect to the perception of the two illocution types, voice quality has also been shown to be relevant for the distinction between RQs and ISQs. On the basis of previous findings gained from a German production study (see [4]), [9] conducted a series of perception studies focusing on *wh*-questions in German and specifically investigated the relevance of the nuclear tune and voice quality for the interpretation either as *wh*-RQ or *wh*-ISQ. The results indicate that participants were able to identify RQs on the basis of sentence-initial breathiness and a nuclear late peak followed by a low boundary tone (L*+H L%). In contrast, a modal voice quality in combination with a nuclear early peak followed by a low boundary tone primarily triggered ISQ interpretations (H+L* / H+!H* L-%, following [10]).

With respect to (ii) and based on previous findings, the first question is if direct measures of the respiratory subsystem can reveal differences between both illocution types during their realization. Given that breathiness involves greater transglottal air flow due to glottal air leakage ([11]) leading to a higher level of air consumption, the thorax volume decreases faster and speakers need to breathe in more deeply both before and after breathy voice. In other words, if movements of the speakers' chest and/or abdomen are involved in this process, is it possible to detect them by using a respiratory inductance plethysmography (RIP) system? The advantage of this system as a direct bio-signal measurement – compared to acoustic measurements of voice quality – is that it is less sensitive to factors such as overlying interferences or the or the quality of the microphone. Hence, the present study also investigates the usefulness of the RIP system for measuring breathing patterns that are related to changes in the voice quality of a speaker and could therefore further support acoustic measurements.

The second question concerning aim (ii) addresses irony, which has previously been defined as a main function of RQs

http://dx.doi.org/10.21437/Interspeech.2020-1607

(e.g., [12,13]) and as a way of expressing an ironic remark (e.g., [14,15]), laughter, sighs or any other nonverbal expressions might be detectable in participants' breathing behavior during the realization of RQs. So far, there are hardly any investigations that analyze a speaker's respiration during the realization of RQs or ironic utterances. Nevertheless, from investigations on the breathing behavior during emotional states we know that changes in both breathing rhythm and pattern (e.g., [16,17]) are involved and can have consequences for speech breathing. For instance, in their review, [17] report a slightly deeper breathing in the expression of disgust in comparison to neutral speech. As for vocal cues during the expression of irony, breathiness and vocal effort were also found to be higher in certain types of irony, such as sarcasm: [18] found not only lower HNR values (which are associated with a breathier voice) for sarcasm, but also a different behavior across languages in terms of vocal cues as reported for English and Cantonese [19]. In BP, sarcasm and amused irony were realized with more vocal effort reflected in less sharp long-term average spectrum (LTAS, see [20]). These results additionally motivated us to investigate possible changes in breathing patterns during the production of RQs in G and BP.

Taken together, the present production experiment investigates the realization of string identical RQs and ISQs in G and BP. This study compares the phonetic characteristics in terms of duration and voice quality as well as the breathing patterns during the realization of the two illocution types.

2. Experiment

2.1. Method

2.1.1. Materials

The materials that were designed for the present production experiment are illustrated in Table 1.

Table 1: *Example contexts for RQ and ISQ in G and BP with an English (E) translation below.*

	RQ	ISQ
G	Du und deine Freunde seid zum Grillen eingeladen und ihr habt alle einen riesen Hunger. Leider dauert es noch eine Weile bis zum Essen und die Gastgeberin stellt lediglich ein paar winzige Oliven auf den Tisch, worüber ihr euch lustig macht. Du sagst:	Du und deine besten Freundinnen seid in einem Restaurant zum Essen und auf der Speisekarte stehen köstliche Oliven, über die du vor einiger Zeit gelesen hast. Dich interessiert schon immer, wie sie schmecken und du möchtest wissen, ob jemand mit dir probieren möchte. Du sagst:
BP	Você e seus amigos são convidados para churrasco e todos vocês estão com fome. Infelizmente, leva um tempo para comer ea anfitriã só coloca algumas azeitonas minúsculas sobre a mesa, o que você se divertir. Você diz:	Você e seus melhores amigos estão comendo em um restaurante e o cardápio apresenta deliciosas azeitonas que você leu sobre algum tempo atrás. Você sempre se interessou em como eles saboreiam e você quer saber se alguém quer tentar com você. Você diz:
E	You and your friends are invited to a barbecue and everyone is very hungry. Unfortunately, it still takes a while till food is ready and the hostess only puts a few tiny olives on the table, which you find really funny. You say:	You and your best friends are in a restaurant for dinner and the menu offers delicious olives that you read about a while ago. You have always been interested in how they taste, and you want to know if someone wants to try them with you. You say:
	Wer isst denn Oliven? / Quem come azeitonas? Who eats olives?	

Overall, we designed 12 target *wh*-questions for the study. The wording of these target questions was equally compatible with both a rhetorical and an information-seeking reading. Each target interrogative started with the *wh*-word "who" (G: *wer*, BP: *quem*) followed by the verb (e.g., "eats", "likes", "wants"). In the German stimuli, the verb was followed by the modal particle *denn* (e.g. *Wer mag denn Oliven*, "Who likes PRT olives"), which can occur in both illocution types ([21]) and has been discussed as creating a more familiar and casual speaking style ([22]). The sentence was concluded by a sentence-final object noun (e.g., "wasabi", "lavender", "Nutella") which consisted of three to four syllables. The object nouns were sonorant almost throughout and carried the lexical stress on the penultimate syllable in both languages. For each target interrogative, two short written contexts were designed, one triggering an RQ reading and one triggering an ISQ reading (see Table 1).

The sentence-final object noun was already introduced in each context to make sure that the speakers realized them prosodically in the subsequently presented target interrogatives as given information. Both the contexts as well as the respective target interrogatives were translated into BP by a native speaker on the basis of the materials designed for the study in G.

2.1.2. Participants

So far, 12 native speakers of G (7 female, Ø = 30.5 years, SD = 4.3 years) and 12 native speakers of BP (5 female, Ø = 25.8 years, SD = 2.6 years) voluntarily participated in the experiment. They were tested individually and were naïve as to the purpose of the study.

2.1.3. Procedure

The procedure was similar for all participants across both languages (cf. [4]). A self-paced elicitation task was designed and conducted by participants in a sound-attenuated booth. They sat comfortably at a table with a laptop computer on a desk in front of them. Contexts (triggering an RQ reading or an ISQ reading) and the respective target interrogatives were presented to participants in the form of a Power Point presentation.

Each trial started with the presentation of a short context (Table 1) and participants were asked to carefully read it silently in their own tempo. Then, they were presented with a target interrogative which was displayed in isolation on a subsequent slide. Participants' task was to read the respective target interrogative out loud and as naturally as possible in the given context. Participants were recorded with a headset microphone (JK MIC-J 071S).

Each participant produced each interrogative two times, i.e. in each of the two contexts (RQ and ISQ). Occurrences of the same interrogative were separated by at least two other context-question pairs. Prior to the actual experiment, participants were presented with a written instruction that was printed on a sheet of paper. Per language, 144 target interrogatives were recorded (12 speakers x 6 *wh*-questions x 2 illocution types). For each language, we additionally monitored the chest and abdominal breathing for 50% of our participants (i.e. 6 out of the 12) with a belt system that was connected to a RespTrack monitor developed by [23]. The elicitation study took about 12-15 minutes.

2.1.4. Measured variables

The measured variables of the present investigation are twofold. First, we analyzed and compared participants' respiratory behavior during the realization of RQs and their respective ISQ. More specifically, we analyzed participants' inhalation and exhalation phase. Breathing signals for both chest and abdomen

were filtered at a 1-Hz cut-off frequency to allow automatic detection of local maxima and minima. A local maximum just before the onset of the speech signal was taken as the inhalation peak and the corresponding breath cycle was then delimited by the minimum values just before and just after that peak. For each breathing cycle, we computed the following variables: amplitude of inhalation (difference between peak and previous minimum value); duration of inhalation (difference between time at the peak and time at the previous minimum); inhalation slope (ratio between amplitude and duration of inhalation); duration of exhalation (difference between time at following minimum and time at peak); exhalation slope (ratio between amplitude and duration of exhalation); duration of the breath cycle (the sum of the inhalation and exhalation durations).

Second, for the investigation of the prosodic components characterizing RQs and string identical ISQs in both languages, we compared participants' voice quality and overall sentence duration since both have been already established as reliable parameters for the distinction between RQs and ISQs. To this end, a phonetically trained annotator labeled the utterance boundaries to determine the overall sentence duration of all target utterances as well as word boundaries to analyze the duration for the initial question word, the verb and the sentence-final object noun. For the analysis of voice quality, the middle of the stressed vowel of the first constituent "who" (v1), the verb (v2) and the sentence-final object noun (v3) were manually labelled as b(reathy), m(odal), or g(lottalized) according to the auditory analysis of the phonetically trained annotator. Furthermore, F0 maximum and minimum (both measured in Hz) were automatically extracted by using Praat scripts [24].

3. Results

All variables were statistically analyzed by using R [25]. The breathing-related statistical mean differences across the question types were calculated using a Kruskal-Wallis test followed by Bonferroni-corrected Wilcoxon tests for each variable. Concerning voice quality, we validated the auditory classification of voice quality (*HNR* as dependent variable, *manual classification* and *language* as predictor variables, cf. [4]) to test if b- and m-labels differed with respect to HNR values for each of the vowel positions (v1, v2, v3). For G, both b- and g-labels differed from m-labels, both indicated by lower HNR values in all three positions (all p-values < 0.05). For BP, HNR values for b-labels were lower than HNR values for m-labels only in the third position (p = 0.02). Values for g-labels showed lower HNR values than m-labels in all positions (all p-values < 0.02). The proportion of the breathy and glottalized vowels was analyzed by categorizing b-labels as 1 and all others as 0 with *illocution type* and *language* as fixed factors and *subjects* and *items* as crossed random factors, allowing for random adjustments of intercepts ([26]). *P*-values were calculated using the Satterthwaite approximation in the R-package lmerTest ([27]). For the analysis of the overall sentence duration and the duration of the initial question word, the verb and the final object noun of the target sentences, linear mixed effects regression models were calculated with *sentence duration* as dependent variable, *illocution type* and *language* as fixed factors and *subjects* and *items* as crossed random factors.

Regarding the variables for breathing-related factors, our results clearly indicate that there were no significant differences between RQs and ISQs in either of the languages (G: all p-values > 0.28, BP: all p-values > 0.37). Generally, speakers of G showed a higher inhalation amplitude (p < 0.0001), which in turn yields steeper slopes (in absolute terms) of both inhalation (p < 0.0001) and exhalation (p<0.0001). The investigation of abdomen and chest movements that were caused by laughter or sighs during the realization of the RQ was inconclusive: There were not enough data points for an established analysis.

Results for the proportion of b-labels in v1 and v2 showed an interaction between *illocution type* and *language* (v1: β=3.1, *SE*=0.7, p<0.0001; v2: β=4.1, *SE*=1.3, p=0.002) indicating a significantly higher amount of b-vowels in both positions for RQs than for ISQs in G (v1: β=3.5, *SE*=0.7, p<0.0001; v2: β=3.2, *SE*=0.8, p<0.0001), but not so for BP (both p-values>0.3; see Figures 1 and 2). For v3, there was also an interaction between *illocution type* and *language* (β=1.7, *SE*=0.002, p<0.0001) indicating that RQs were produced with more breathy vowels in the third position, with this effect being stronger for G (β=2.5, *SE*=0.6, p<0.0001) than for BP (β=0.7, *SE*=0.002, p<0.0001). The proportion of g-labels showed no difference in the three positions in BP (all p-values > 0.07).

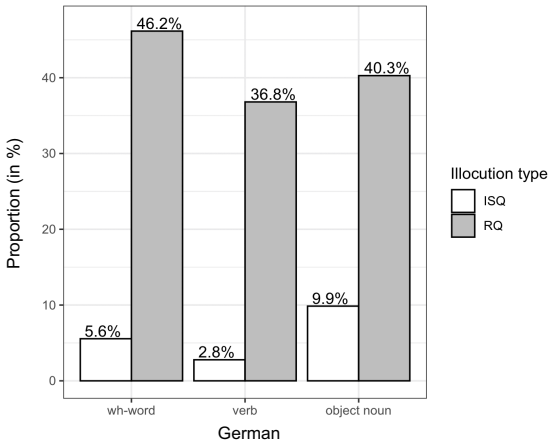

Figure 1: *Proportions (in %) of vowels classified as breathy for G in all three positions.*

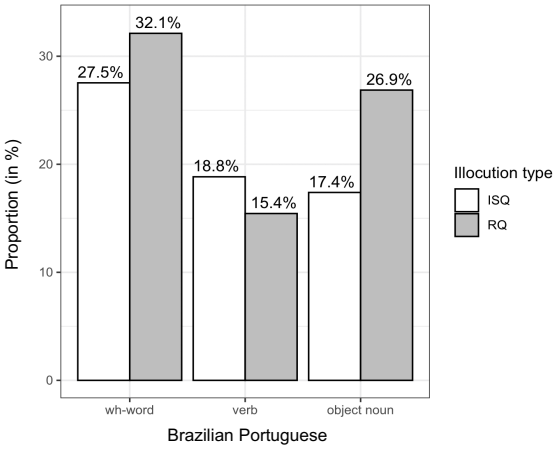

Figure 2: *Proportions (in %) of vowels classified as breathy for BP in all three positions.*

Concerning the overall sentence duration, our analysis showed an interaction between *illocution type* and *language* (β=0.08, *SE*=0.03, p=0.02) indicating a longer overall sentence duration for RQs than for ISQs in both G (β=0.2, *SE*=0.04, p<0.001) and BP (β=0.1, *SE*=0.1, p=0.04), but overall with a greater effect for G (see Figure 3).

Analyzing the absolute constituent duration of the initial *wh*-word showed a main effect of *illocution type* (β=0.04,

SE=0.01, *p*<0.001) with a significantly longer mean duration of the initial question word in both languages (G: β=0.04, *SE*=0.01, *p*<0.001; BP: β=0.03, *SE*=0.01, *p*=0.01). In both languages, we found no difference concerning the duration of the verb (*p*-values > 0.6). With respect to the absolute constituent duration of the sentence-final object noun, our findings revealed an interaction between *illocution type* and *language* (β=0.1, *SE*=0.03, *p*<0.001) indicating a longer mean duration for the object noun in *wh*-RQs in G (β=0.1, *SE*=0.03, *p*<0.001). In contrast, there was no difference in the data subset of BP (*p*=0.5).

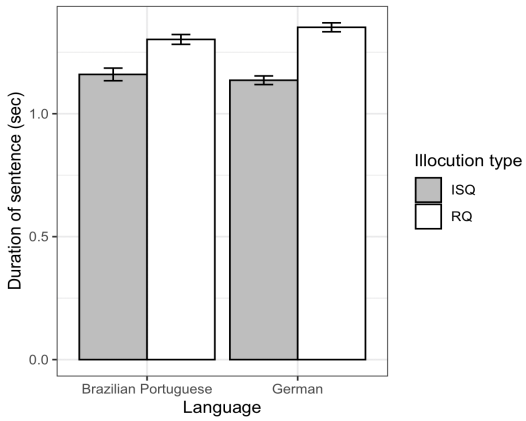

Figure 3: *Overall sentence duration (in sec) for both illocution types in both languages.*

4. Discussion

Taken together, our results concerning voice quality and duration in the realization of RQs in G are similar to previous studies (see [4,9]). Additionally, concerning the phonetic characteristics investigated here, our results indicate similarities between the realization of RQs compared to ISQs in both languages: Our findings suggest that longer durations as well as a breathier voice quality seem to be crucial phonetic features in the realization of RQs compared to string identical ISQs not only in G but also in BP. However, both languages differ from one another with respect to the placement of the realized breathiness: while speakers of G realize RQs more often with a breathy voice than ISQs in the three positions – mainly in the beginning – speakers of BP seem to have a breathier voice quality especially in the end of the target utterance (i.e., object noun).

Regarding duration, our findings indicate that in both languages the overall sentence duration is longer in RQs than in ISQs, which is in line with previous studies [4,5,6]. More specifically, durational differences for RQs in G and BP become most obvious on the basis of different constituents: While a longer duration becomes most obvious in the initial position of RQ realizations in BP, it is the sentence-final object noun in RQs of G that is realized with a longer duration than in ISQ. Hence, the present study does affiliate to previous results in other languages reporting longer constituent durations in RQs than in ISQs, but the initial part of the target utterance seems to be more crucial for the realization of RQs in BP than the final part of the utterance, vice versa for G.

Taken together, the two phonetic characteristics of voice quality and duration differ significantly between both illocution types in G and BP with effects going in the same direction. A closer look, however, shows that these two phonetic cues are realized differently. That is, first, while RQs in G show a significant lengthening of the last constituent, RQs in BP are usually realized with a significantly lengthened initial constituent. Second, while RQs in G are realized with more breathy vowels in the beginning of the target interrogative, RQs in BP are realized with more breathy vowels at the end of the target utterance.

With respect to the breathing-related parameters, there were no clear links between the RQ/ISQ breathiness differences and the surrounding speech-breathing patterns. However, note that the speakers of G who breathed in more deeply and dynamically also produced RQs with more breathy vowels than the speakers of BP. That is, perhaps planning a modally voiced utterance does not necessarily cause a lower amplitude and dynamics of breathing; rather, conversely, there could be something like a fixed amplitude and dynamics of breathing that determines how pronounced a meaningful difference between modal and breathy voice can be implemented by the speakers of a language. Future experiments will shed more light on this assumption.

We will also vary the overall length of the utterances in these future experiments, because perhaps the target interrogatives used here were too short for voice quality related differences to affect inhalation and exhalation behavior. Instead, participants could read the whole context out loud followed by the target interrogative to allow for a monitoring over a longer time frame in the respective conditions.

Finally, it is also interesting that differences in breathing behavior when expressing irony and emotions were found in previous studies, but not for the present difference between matter-of-fact ISQs and emotion-loaded RQs. This could indicate (assuming an authentic implementation of our target interrogatives) that speakers do not express real emotions through RQs, but rather reduced and conventionalized images of these emotions. We will continue to investigate this question in future experiments as well.

5. Conclusions

Beyond prosodic similarities between Germanic languages, the present results indicate similarities for the realization of RQs in BP as a Romance language. More precisely, the two phonetic characteristics duration and voice quality have been shown to play a role in the realization of RQs in both G and BP. Furthermore, the present study stresses the importance of having a closer look at cross-linguistic phonetic characteristics that are distinctive of the realization of the two illocution types. That is, besides the similarities concerning duration and voice quality in both languages, it is also important *where* exactly these characteristics are most pronounced within the question utterance. To investigate if the presented phonetic distinctions are also crucial for the disambiguation between the two illocution types in BP, subsequent perception studies are necessary. Additionally, the role of a glottalized voice needs to be further investigated with respect to BP.

In order to examine our assumption mentioned above, a combined investigation of both the signal that is measured at the glottis and by the RIP is conceivable. More specifically, an investigation of two respective time-aligned systems could shed light on the interplay between breathing and voice quality by making use of the RIP system and an Electroglottography system (EGG). An EGG system could also detect more subtle differences concerning a speaker's voice quality during the realization of real emotions and conventionalized images of these emotions.

6. References

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

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

[3] I. Caponigro and J. Sprouse, "Rhetorical Questions as Questions", In: Puig-Waldmüller, E. (ed.), *Proc. of Sinn und Bedeutung* 11, Barcelona, Spain: Universitat Pompeu Fabra, 121-133, 2007.

[4] B. Braun, N. Dehé, J. Neitsch, D. Wochner, and K. Zahner, "The prosody of rhetorical and information-seeking questions in German," *Language and Speech*, vol. 62, no. 4, pp. 779-807, 2019.

[5] N. Dehé and B. Braun, "The prosody of rhetorical questions in English," *English Language & Linguistics*, pp. 1-29, 2019.

[6] N. Dehé and B. Braun, "The Intonation of Information-Seeking and Rhetorical Questions in Icelandic," *Journal of Germanic Linguistics*, vol. 32, no. 1, pp. 1-42, 2020.

[7] V. J. v. Heuven and E. v. Zanten, "Speech rate as a secondary prosodic characteristic of polarity questions in three languages," *Speech Communication*, vol. 47, pp. 87-99.

[8] O. Niebuhr, J. Bergherr, S. Huth, C. Lill, and J. Neuschulz, "Intonationsfragen hinterfragt! – Die Vielschichtigkeit der prosodischen Unterschiede zwischen Aussage- und Fragestäzen mit deklarative Syntax - On the complexity of prosodic differences between declaratives and declarative questions," *Zeitschrift für Dialektologie und Linguistik*, vol. 77, pp. 304-346, 2010.

[9] J. Neitsch, *Who cares about context and attitude? Prosodic variation in the production and perception of rhetorical questions in German*, Dissertation, Germany: University of Konstanz, 2019.

[10] T. Rathcke and J. Harrington, "Is there a distinction between H+!H* and H+L* in standard German? Evidence from an acoustic and auditory analysis," in *Speech Prosody 2006 - 3rd International Conference on Speech Prosody, May 2-5, Dresden, Germany, Proceedings,* pp. 783-786, 2006.

[11] G. de Krom, "Some spectral correlates of pathological breathy and rough voice quality for different types of vowel fragments," *Journal of Speech, Language, and Hearing Research*, vol. 38, no. 4, pp. 794–811, 1995.

[12] A. F. Freed, "The form and function of questions in informal dyadic conversation," *Journal of Pragmatics*, vol. 2, no. 6, pp. 621-644, 1994.

[13] R. A., Hudson, "The meaning of questions," *Language*, vol. 51, no. 1, pp. 1-31, 1975.

[14] J. Frank, "You call that a rhetorical question?: Forms and functions of rhetorical questions in conversation," *Journal of Pragmatics*, vol. 14, no. 5, pp. 723-738, 1990.

[15] J. Neitsch and T. Marinis, "The prosodic Mapping of Irony in German Rhetorical Questions," In M. Elmentaler & O. Niebuhr (eds.): *An den Rändern der Sprache*. Kiel: CAU. Accepted.

[16] I. Homma and Y. Masaoka, "Breathing rhythms and emotions," *Experimental physiology*, vol. 93, no. 9, pp. 1011-1021, 2008.

[17] F. A., Boiten, N. H. Frijda, and C. J. Wientjes, "Emotions and respiratory patterns: review and critical analysis," *International journal of psychophysiology*, vol. 17, no. 2, pp. 103-128, 1994.

[18] H. S. Cheang, and M. D. Pell, "The sound of sarcasm," *Speech communication*, vol. 50, no. 5, pp. 366-381, 2008.

[19] H. S. Cheang, and M. D. Pell, "Acoustic markers of sarcasm in Cantonese and English," *The Journal of the Acoustical Society of America*, vol. 126, no. 3, pp. 1394-1405, 2009.

[20] W. Silva, and P. A. Barbosa, "Prosodic correlates of the expression of pure sarcasm and sarcastic irony in Brazilian Portuguese," In: *Anais do Congresso Brasileiro de Prosódia*, vol. 1, no. 4, pp. 111 – 114, 2014.

[21] M. Thurmair, *Zum Gebrauch der Modalpartikel 'denn' in Fragesätzen. Eine korpusbasierte Untersuchung.* Tübingen: Niemeyer, 1991.

[22] E. Hentschel, "Funktion und Geschichte deutscher Partikeln: Ja, doch, halt und eben," Berlin: Walter de Gruyter, 2011.

[23] M. Włodarczak, and M. Heldner, "Respiratory belts and whistles: A preliminary study of breathing acoustics for turn-taking," in *INTERSPEECH 2016 – 17th Annual Conference of the International Speech Communication Association, September 8-12, San Francisco, USA, Proceedings*, 2016, pp. 510–514

[24] P. Boersma, and D. Weenink, "Praat - doing phonetics by computer (Version Version 5.3.64)," Amsterdam: Institute of Phonetics, retrieved from http://www.praat.org, (last access 2/12/2020).

[25] R Development Core Team. "R: A language and environment for statistical computing (Version 3.2.2)," Vienna: R Foundation for Statistical Computing, retrieved from http://www.r-project.org./, last access 3/4/2020).

[26] H. R. Baayen, D. J. Davidson, and D. M. Bates, "Mixed-effects modeling with crossed random effects for subjects and items," Journal of Memory and Language, vol. 59, no. 4, pp. 390-412, 2008.

[27] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "lmerTest Package: Tests in linear mixed effects models," *Journal of Statistical Software*, vol. 82, no. 13, pp. 1-26, 2017.