

# GENDER DIFFERENCES IN THE PROSODY OF GERMAN QUESTIONS

*Oliver Niebuhr*

Dept. of Design and Communication, IRCA, University of Southern Denmark, Sønderborg, Denmark  
olni@sdu.dk

## ABSTRACT

Based on enacted dialogues of 60 native speakers, the present study shows that males and females differ in their prosodic realization of lexically and syntactically marked questions in German. Frequency counts of final rises and falls, as well as measurements of mean F0, speaking rate, and intensity levels suggest that males rather than females are more polite, open, and amenable – i.e. overall more hearer oriented – when asking questions. This applies most clearly to syntactically marked questions, i.e. when male speakers ask their dialogue partner to make decisions or confirm pieces of information. Female speakers show signs of hearer orientation in *wh*-questions, i.e. when it comes to requesting explanations or justifications from the dialogue partner.

**Keywords:** gender, question, intonation, prosody.

## 1. INTRODUCTION

It is well known that German questions can occur with rising and falling phrase-final intonations. Early descriptive studies like that of [1] claimed that morphologically marked questions (*wh*-questions) end in a falling intonation, whereas syntactically marked questions (*vi*-questions, with the verb in initial position) end in a rising intonation. This claim was generally accepted for a long time and disseminated in linguistic papers and foreign-language courses on German (e.g., [2,3,4]), until its absoluteness was qualified by analyses of larger dialogue corpora. On the one hand, these analyses confirmed that *wh*-questions and *vi*-questions are in fact predominantly falling or rising, respectively. But, on the other hand, they also yielded more than 30 % counterexamples for both question types, see [5,6,7].

Closer semantic-pragmatic examinations of these counterexamples finally decoupled the phrase-final intonation movements in questions from their supposed morphosyntactic triggers, redefined these movements as communicative elements in their own right, and associated them with an attitudinal meaning difference: Final rises are, across question types, hearer-oriented, convey politeness, interest, and invite the dialogue partner to determine content and length of the answer independently of the questioner. The opposite is true for final falls, cf. [6]. Thus,

“Would that suit you?” produced with a final rise allows for a negative answer and invites the dialogue partner to add additional explanations. In contrast, the same question wording, but with a final fall, only allows the dialogue partner to answer with a short “yes” or “no”, and clearly calls for the positive answer the questioner would like to hear.

Recent findings of [8,9] suggest that rising and falling question intonations in Greek have – roughly speaking – the same communicative meaning as in German. Moreover, the findings on Greek revealed another variable in the utilization of rising and falling questions: Gender. Female speakers of Greek used final rises more frequently than male speakers. In addition, final rises sounded more polite than final falls when used by female speakers.

This finding is remarkable insofar as studies on gender-related differences in speech have, as yet, mainly focused on mode of expression and discourse behavior or – in the phonetic domain – on the production of sound segments [10,11]. If studies did deal with gender-specific prosodies, then typically just at the very general level of anatomical differences or learned social roles that shape, for example, speaking rate, voice quality, and F0 level/range [12].

Among the rare exceptions are the studies of [13, 14]. They found two variants of final falls in German statements. One variant falls steeply to the lower limit of the speaker's individual F0 range. The other variant decelerates towards the end, and then changes into a short and tiny rise at the phrase boundary. This second variant was called “pseudo-terminal fall”, as opposed to the regular terminal fall. Compared with the terminal fall, the pseudo-terminal fall is less categorical, dominant, and impolite. Rather, it conveys the speaker's willingness to compromise, discuss, and accommodate. Pseudo-terminal falls are more often used by female than by male German speakers [13].

Inspired by the gender-specific use of final rises in Greek questions, and taking up the line of research of [13,14] on the gender-specific use of final falls in German statements, the present production study investigates whether male and female speakers also differ in the way they realize German questions. The conducted prosodic analysis pays special attention to the direction of the phrase-final intonation movement. However, other long-term para-

meters like speaking rate, intensity level, and F0 level are taken into account as well. The analysis is based on a sub-set of dialogues from the SPID corpus that was recorded for a large project on In-Car Communication at Kiel University, see [15].

The preference for pseudo-terminal falls in statements suggests that female speakers in German still follow traditional gender roles. Based on this assumption, female speakers should aim to be more polite, open, and hearer-oriented than male speakers when asking questions. Translated into the attitudinal meanings of final rises and falls in German questions, this means that the present results should replicate those of Greek [8,9]: Rises at the ends of questions should occur more frequently for female than for male speakers. Besides, female speakers should also show slower speaking rates, higher F0 levels, and lower intensity levels than those of male speakers, as these phonetic characteristics are also known to be associated with politeness, openness, interest, and hearer orientation, cf. [16,17,18].

## 2. METHOD

### 2.1. Corpus, speakers, and speech material

The SPID corpus (SPontaneous In-car Dialogues, [19]) includes two data sets. A larger set of spontaneous dialogues was elicited on the basis of map-tasks and the less formal daily-soap scenario [20].

A smaller data set was elicited by means of written dialogue texts. These texts were trained by pairs of speakers prior to the recording, and then performed at the recording as if they were developing spontaneously in everyday conversations about traffic situations or travel destinations. In order to enhance the naturalness of this task, the dialogues were performed by good friends inside a real car, which stood in a large sound-treated booth at Kiel University. Moreover, common speech reduction patterns were already implemented in the orthographic representation of the dialogue texts; and speakers were explicitly allowed to adjust the texts to their own mode of expression by replacing, omitting, or adding words. The present analysis is based on these performed dialogue texts, as they offered a greater control over semantic-pragmatic and phonetic variables than actual spontaneous speech.

The dialogue texts were performed by 30 pairs of Standard German native speakers, i.e. by a total of 60 speakers. They were between 20-30 years old and undergraduate students of linguistics at Kiel University. All students had basic experience with speech recordings. Thus, it was easy for them to relax and let themselves go in the recording situation. Speaker pairs were gender balanced. That is, the re-

cordings were conducted with equal numbers of male-male, female-female, and mixed-gender pairs.

The dialogue texts contained 6 different wh-questions and 8 different vi-questions. Thus,  $6 \times 30 = 180$  wh-questions and  $8 \times 30 = 240$  vi-questions were available for analysis. All questions were at least 5 syllables long, contained at least 2 pitch accents, and were consistently located in turn-final position.

### 2.2. Analysis

The analysis of the two question samples was conducted in three steps. In the first step, we simply counted on an auditory basis how many of the 180 wh-questions and 240 vi-questions occurred with final rising and falling intonations.

The second step was the extraction and annotation of all 420 questions. The annotation was conducted on the basis of TextGrid intervals in Praat [21]. The TextGrids specified speaker gender and question type (wh vs. vi) for each token.

In the third step, the annotated tokens were processed by ProsodyPro [22] and automatically analyzed with respect to question duration, mean F0, and mean intensity. Question duration was converted into the established measure of phonological syllables per second. Intensity measurements were analyzable, as the speakers wore head-mounted microphones. In this way, the microphone-to-mouth distance was kept constant, independently of the speakers' head movements.

## 3. RESULTS

### 3.1. Frequencies of rising and falling intonations

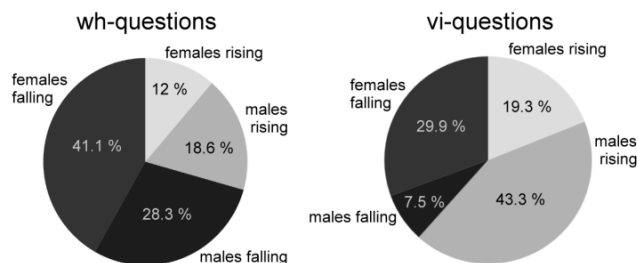
Figure 1 displays the relative frequencies of rising and falling intonations produced by male and female speakers at the ends of wh-questions and vi-questions. The most obvious result is that the majority of wh-question intonations fell at the end (69.4 %), whereas vi-questions were predominantly realized with final rises (62.6 %).

These numbers are consistent with those found in previous studies on spontaneous dialogues in German. They are even in the same order of magnitude [5,6,7]. That is, more than 30 % of our final intonation movements run counter to the old but persistent claim that syntactically marked questions rise and lexically marked questions fall in German.

This large number of counterexamples supports the revised claim that the direction of the final intonation movement makes a separate contribution to the communicative meaning of German questions: Final rises are hearer-oriented, convey politeness, interest, and invite the dialogue partner to determine content and length of the answer independently of

the questioner. Final falls have the opposite attitudinal meaning.

**Figure 1:** Percentages of rising (light gray) and falling (dark gray) intonations at the ends of wh-questions (n=180) and vi-questions (n=240).



The relative frequencies in Figure 1 also support the ecological validity of our dialogues. They were performed on the basis of trained text templates, in order to achieve a high degree of contextual control. Yet, the frequencies we got in this way are qualitatively and quantitatively comparable with those of previous studies on actual spontaneous speech.

Based on this ecological validity, it is particularly worth noting that Figure 1 also shows a clear gender difference: Females produced more final falling intonations than males. Almost two-thirds of all falling wh-questions were produced by the female speaker group. In the case of vi-questions, the female speaker group is even responsible for more than three-fourths of all falling intonations. In contrast, male speakers asked questions more frequently with final rising intonations, particularly in the vi-condition. Chi-squared tests showed that these gender-related frequency differences are significant for both wh-questions ( $\chi^2[1]=6.342$ ,  $p<0.05$ ) and vi-questions ( $\chi^2[1]=52.702$ ,  $p<0.001$ ).

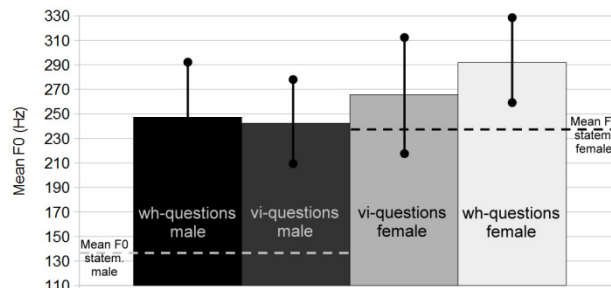
### 3.2. Acoustic parameters

The measurement data generated by ProsodyPro was statistically analyzed in a MANOVA with Gender (male vs. female) and Question Type (wh vs. vi) as fixed factors. Speaker Pairing (m/m, f/f, m/f) was included as a covariate, but had no significant effects.

Regarding F0 level, the MANOVA yielded two significant main effects: a strong effect of Gender ( $F[1,836]=635.669$ ,  $p<.001$ ,  $\eta^2_p=.432$ ), and a weaker effect of Question Type ( $F[1,836]=46.023$ ,  $p<.001$ ,  $\eta^2_p=.052$ ). There was no interaction between the two. The weak effect of Question Type reflects that both male and female speakers produced wh-questions at a slightly higher mean F0 level than vi-questions, see Figure 2. Since we took care that the two types of questions had overall about the same number of syllables, the difference in mean F0 level cannot be an artefact of question length. We can also rule out that the F0 difference is an artefact of a different

situational embedding, as this variable was as well controlled in the dialogue recordings.

**Figure 2:** Mean F0 levels of male (dark gray) and female (light gray) speakers in vi-questions (n=240) and wh-questions (n=180). Dotted lines refer to mean F0 levels in statements.

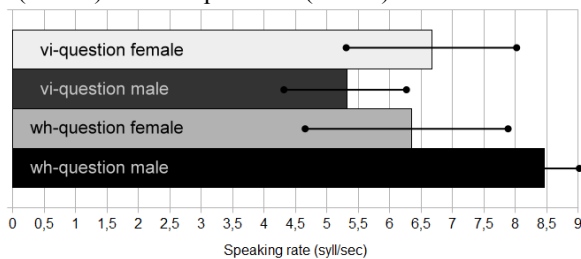


The strong effect of Gender on mean F0 level is obviously related to the anatomical differences between male and female larynges, due to which females phonate at a higher F0 level than males. However, if the effect of Gender was only anatomical, then the mean F0 level of females should have been almost one octave above that of males, i.e. 210-230 Hz vs. 120-130 Hz according to reference values in the literature [12,23,24]. In contrast, in our data, the mean F0 level of females was only about one quarter of an octave above that of males. In order to check whether or not we analyzed representative speakers, we additionally measured the mean F0 level in a randomly compiled set of 300 declarative statement utterances from the dialogue texts (five per speaker). The mean F0 levels across all male and female statement utterances fit in very well with the reference values from the literature, see the dotted lines in Figure 2. This left us with only one possible conclusion: Our results on mean F0 include a *behavioral gender difference*. That is, both male and female speakers raised their F0 level when asking questions, but the magnitude of this F0 raising was much larger for male than for female speakers.

The results for speaking rate are depicted in Figure 3. The MANOVA yielded a highly significant main effect of Question Type ( $F[1,836]=467.512$ ,  $p<.001$ ,  $\eta^2_p=.359$ ), but no main effect of Gender. However, there was a significant interaction of Gender and Question Type ( $F[1,836]=96.443$ ,  $p<.001$ ,  $\eta^2_p=.103$ ). The main effect of Question Type is due to the fact that wh-questions were produced at a faster speaking rate than vi-questions. The average difference was about 1.2 syll./sec. The interaction of Question Type and Gender was disordinal, i.e. due to an inversion effect, which made additional t-tests necessary. The t-tests showed ( $p<0.01$ ) that males were faster than females in producing wh-questions. By contrast, in vi-questions, it were the females who had a higher speaking rate than the males.

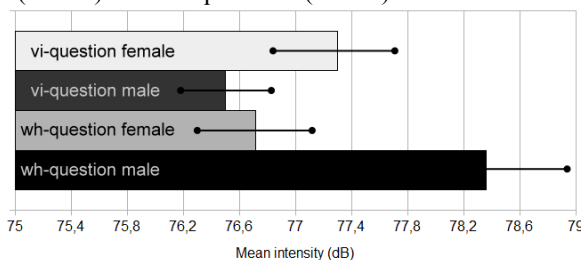
The results pattern of mean intensity resembles that of speaking rate and is summarized in Figure 4. There was a significant effect of Question Type ( $F[1,836]=54.953$ ,  $p<.001$ ,  $\eta_p^2=.062$ ) and, unlike for speaking rate, also a significant main effect of Gender ( $F[1,836]=52.280$ ,  $p<.001$ ,  $\eta_p^2=.059$ ). Wh-questions had on average a higher mean intensity level than vi-questions. Likewise, male speakers had on average a higher mean intensity level than female speakers. This main effect could again be due to anatomical differences that underlie male and female phonation. Everything else being equal, females are known to speak with a breathier voice quality than males. Breathiness causes a higher spectral tilt and hence a lower mean intensity level, cf. [12,25].

**Figure 3:** Mean speaking rates of male (dark gray) and female (light gray) speakers in vi-questions ( $n=240$ ) and wh-questions ( $n=180$ ).



Over and above these main effects, there was also a significant interaction between Gender and Question Type ( $F[1,836]=117.708$ ,  $p<.001$ ,  $\eta_p^2=.123$ ). Once again, this interaction was disordinal. But, t-tests showed ( $p<0.01$ ) that female speakers had a lower intensity level than male speakers in wh-questions. In contrast, in vi-questions, it were the males whose intensity level was lower than that of the females.

**Figure 4:** Mean intensity of male (dark gray) and female (light gray) speakers in vi-questions ( $n=240$ ) and wh-questions ( $n=180$ ).



#### 4. DISCUSSION AND CONCLUSIONS

The present study searched for gender differences in the realization of German wh- and vi-questions, based 60 male and female native speakers. If, in accordance with the traditional gender roles, female speakers are more polite, open, interested and amenable than male speakers when asking questions, then female questions should more often end in a rising intonation, show higher F0 levels, lower intensity levels, and lower speaking rates than those of males.

The results are not consistent with this assumption. The analysis did reveal clear gender differences. However, it were the *male* rather than the female speakers who produced more final rising intonations in both types of questions. Likewise, the *male* speakers raised their mean F0 level in questions to such an extent that they almost evened out their intrinsic F0 distance (of about one octave) to female speakers. The *male* speakers also had lower intensity levels and speaking rates than the female speakers, but only in vi-questions. In wh-questions, it were the females who had lower intensity levels and speaking rates than the males; and if female speakers showed signs of F0 raising, then mainly in wh-questions, see Figure 2.

So, overall, our behavioral "snapshot" of young German speakers suggests that males rather than females take the role of "the weaker sex" when asking questions. The prosodic profile of males can be considered more strongly hearer oriented, particularly in vi-questions, i.e. when male speakers ask the interlocutor to make decisions or confirm pieces of information. Females show stronger signs of hearer orientation than males only beyond F0, and only in wh-questions, i.e. when it comes to requesting explanations or justifications from their interlocutor.

In summary, the empirical picture drawn by the present study is much more complex than the simple stereotypes of men and women and their manifestation in speech. It is possible that traditional gender roles are changing. But, what is more likely with respect to previous studies [26,27] is that simple stereotypes just never bear close linguistic examination.

In any case, the results are clearly consistent with the basic cross-linguistic conclusion of [9,13]: There *basically are* gender differences in the use of prosodic patterns, and future studies should take this fact into account (a) by controlling gender more carefully in experiments, and (b) by putting more emphasis on investigating gender-specific prosodies within and across languages, also with respect to their pragmatic interpretation and interaction with other linguistic and contextual factors. For example, it would be interesting to find out if the present conclusions only hold for dialogues inside a car, which is a traditionally male-connoted environment, or if the same gender differences in the realization of wh- and vi-questions also occur in other everyday situations.

Finally, we found general prosodic differences between wh- and vi-questions. Wh-questions were produced faster, louder, and at a higher F0 level than vi-questions. This new insight should also be examined in more detail in follow-up studies, particularly on the basis of actual spontaneous speech, even though we can assume (see 3.1) that our performed dialogue texts did generate reliable speech behavior.

## 5. REFERENCES

- [1] von Essen, O. 1964. *Grundzüge der hochdeutschen Satzintonation*. Ratingen: Hehn
- [2] Griesbach, H. 2000. *Bauplan Deutsch: Übungsgrammatik und Satzbaufehler*. Frankfurt: Libri.
- [3] Batliner, A. 1991. Ein einfaches Modell der Frageintonation und seine Folgen. In: Klein, E., Pouradier Duteil, F., Wagner, K.H. (eds), *Betriebslinguistik und Linguistikbetrieb*. Tübingen: Niemeyer, 147-160.
- [4] Brinckmann, C., Benz Müller, R. 1999. The Relationship between Utterance Type and F0 Contour in German. *Proc. 6th European Conference on Speech Communication and Technology (EUROSPEECH), Budapest, Hungary*, 21-24.
- [5] Selting, M. 1995. *Prosodie im Gespräch. Aspekte einer interaktionalen Phonologie der Konversation*. Tübingen: Niemeyer.
- [6] Kohler, K.J. 2004. Pragmatic and attitudinal meanings of pitch patterns in German syntactically marked questions. In: Fant, G., Fujisaki, H., Cao, J., Yi, X. (eds), *From traditional phonology to modern speech processing – Festschrift for Professor Wu Zongji's 95th Birthday*. Beijing: Foreign Language Teaching and Research Press, 205-215.
- [7] Kügler, F. 2004. Do we know the answer? Variation in yes-no-question intonation. In: Fischer, S., van de Vijver, R., Vogel, R. (eds), *Experimental studies in linguistics 1*. Potsdam: Potsdam University Press, 9-29.
- [8] Arvaniti, A., Ladd, D.R. 2009. Greek wh-questions and the phonology of intonation. *Phonology* 26, 43-74.
- [9] Arvaniti, A., Baltazani, M., Gryllia, S. 2014. The pragmatic interpretation of intonation in Greek wh-questions. *Proc. 7th International Conference of Speech Prosody, Dublin, Ireland*.
- [10] Whiteside, S.P., Henry, L., Dobbin, R. 2004. Sex differences in voice onset time: A developmental study of phonetic context effects in British English. *JASA* 116, 1179-1183.
- [11] Coates, J. 2004. *Women, Men and Language: A Sociolinguistic Account of Gender Differences in Language*. Harlow: Pearson Education
- [12] Simpson, A. 2009. Phonetic differences between male and female speech. *Language and Linguistics Compass* 3, 621-640.
- [13] Peters, B. 2000. Individuelle und geschlechtsspezifische Unterschiede in der prosodischen Gestaltung deutscher Lese- und Spontansprache. In: Schlosser, H.D. (ed.), *Sprache und Kultur*. Frankfurt: Peter Lang, 153-162.
- [14] Ambraszaitis, G.I. 2005. Between fall and fall-rise: substance-function relations in German phrase-final intonation contours. *Phonetica* 62, 196-214.
- [15] Landgraf, R. in press. Simulating complex speech production environments. In: Niebuhr, O., Skarnitzl, R. (eds), *Tackling the complexity of speech*. Prague: Nakladatelství Epocha.
- [16] Culpeper, J., Bousfield, D., Wichmann, A. 2003. Impoliteness revisited: With special reference to dynamic and prosodic aspects. *Journal of Pragmatics* 35, 1545-1579.
- [17] Apple, W., Streeter, L.A., Krauss, R.M. 1979. Effects of Pitch and Speech Rate on Personal Attributions. *Journal of Personality and Social Psychology* 37, 715-727.
- [18] Liscombe, J.J. 2007. *Prosody and Speaker State: Paralinguistics, Pragmatics, and Proficiency*. PhD thesis, Columbia University.
- [19] Niebuhr, O., Peters, B., Landgraf, R., Schmidt G. 2015. The Kiel Corpora of "Speech and Emotion" - A summary. *Proc. 41st Jahrestagung für Akustik, Nuremberg, Germany*, 1-4.
- [20] Peters, B. (2001). 'VideoTask' or 'Daily Soap Scenario' - A new method for the controlled elicitation of spontaneous speech. Manuscript. URL: [http://www.ipds.uni-kiel.de/pub\\_exx/bp2001\\_1/Linda21.html](http://www.ipds.uni-kiel.de/pub_exx/bp2001_1/Linda21.html).
- [21] Boersma, P. 2001. Praat, a system for doing phonetics by computer. *Glott International* 5, 341-345.
- [22] Xu, Y. 2013. ProsodyPro — A Tool for Large-scale Systematic Prosody Analysis. *Proc. Tools and Resources for the Analysis of Speech Prosody (TRASP), Aix-en-Provence, France*, 7-10.
- [23] Traunmüller, H., Eriksson, A. 1995. *The frequency range of the voice fundamental in the speech of male and female adults*. Manuscript, University of Stockholm, [http://www.ling.su.se/staff/hartmut/f0\\_m&f.pdf](http://www.ling.su.se/staff/hartmut/f0_m&f.pdf)
- [24] Henton, C.G. 1989. Fact and fiction in the description of female and male pitch. *Language and Communication* 9, 299-311.
- [25] Klatt, D. H., and L. C. Klatt. 1990. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *JASA* 87, 820-857.
- [26] Hyde, J.S. 2005. The gender similarity hypothesis. *American Psychologist* 60, 581-592.
- [27] Anderwald, L. 2014. You just don't understand! Nichtverstehen zwischen Männern und Frauen. In: Niebuhr, O. (ed.), *Formen des Nicht-Verstehens*. Frankfurt: Peter Lang, 113-127.