



# Prosodic and Phonetic Aspects of Paralinguistic Utterances with the German Modal Particle *schon* in L1 and L2

Miki Ikoma

Waseda University, Tokyo, Japan

ikoma@waseda.jp

## Abstract

In this study, prosodic aspects of the German modal particle *schon* are investigated, expressing three kinds of paralinguistic information, “conviction,” “reservation,” and “rebuttal.” To clarify if there are common vs. language-dependent prosodic aspects of utterances with MP *schon* conveying paralinguistic meaning in L1 and L2, both production and perception experiments for L1 native speakers and L2 Japanese learners of German at the A1 level were conducted. Japanese nonlearners of German (NL) also attended the perception experiment, to compare the results with L1 speakers and L2 learners of German.

The results in production show common prosodic properties in L1 and L2 on the one hand, such as longer duration of utterance, larger pitch range, and higher F1 in the accented vowel in the “rebuttal” utterance; but we also found language-dependent properties, such as lower pitch and higher intensity in the whole utterance in “rebuttal” by L1 speakers, while higher mean pitch and lower intensity in “rebuttal” utterances by L2 speakers were observed. In perception, the results show common tendencies in L1 and L2, such as perceptual confusion between “conviction” and “rebuttal,” although the overall percentage of correct answers by L1 speakers was significantly higher than by L2 speakers and Japanese nonlearners.

**Index Terms:** paralinguistic information, German speech, production and perception of prosody, vowel quality, modal particles

## 1. Introduction

In spoken German conversation, modal particles (MPs) ([1], [2], [3]) are widely used to express various kinds of paralinguistic information, that is, speakers’ intentions and attitudes, depending on different contexts. In this study, we investigate the prosodic aspects of MP *schon*, which can express three main intentions – “conviction,” “reservation,” and “rebuttal,” as in the following examples ([3]):

- A) *Peter kommt schon.* ‘Peter will certainly come.’ “conviction”
- B) *Peter kommt schon.* ‘Peter will come, but ...’ “reservation”
- C) *Peter kommt schon.* ‘Peter WILL come.’ “rebuttal”

As the utterances above show, the single modal particle *schon*, like most other MPs, such as *doch* and *ja*, may convey several different speakers’ intentions, depending on the context.

As far as concerns the prosodic properties of MP *schon* by L1 speakers, it has gradually become clear that in the case of “rebuttal” *schon*, longer sentence duration, a lower F0

minimum of the whole sentence, a larger F0 range, and later F0 peak timing of the accented verb compared with the case of “conviction” *schon* all play an important role in identifying the exact intention of the speaker ([3], [4], [5]).

In the previous studies on prosodic aspects of paralinguistic information, few studies focused on utterances conveying speakers’ attitudes and intentions ([6], [7]), compared to the studies on emotional speech such as “anger” and “happiness” ([8], [9], [10]). [6] analyzed utterances of *Excuse me* in hyperarticulated speech in German and Japanese L1 and L2 production and found there are “both L1-dependent and -independent ways to convey paralinguistic information by phonologically and phonetically modifying F0 contours and segmental durations” ([6, p.192]). [7] analyzed production and perception of utterances in Japanese conveying several paralinguistic messages, such as “admiration” and “suspicion,” and found specific prosodic properties such as longer duration, larger pitch range, later timing and larger pitch range of phrase-initial rise, and higher F1 and F2 in the utterance of “suspicion” than in “neutral.” Based on the results from perceptual experiments by L1 native speakers, L2 learners, and nonlearners of Japanese, they argued that there are language-dependent aspects in the perception of paralinguistic information.

The question then arises: Are there any common vs. language-dependent properties in the production and perception of paralinguistic information expressed by MPs between native and nonnative speakers of German?

In [5], we investigated L1 and L2 German utterances including MP *schon*. The results showed a longer utterance duration in “rebuttal” than “conviction” as a common prosodic property by both L1 and L2 speakers with the experience of study abroad. On the other hand, a higher utterance F0 in “rebuttal” than in “conviction” was observed by L2 speakers without any experience of study abroad. The result from the perception experiment showed similar tendencies by L2 with the experience of study abroad to L1 native speakers of German. However, it has not yet been clarified what the production and perception of paralinguistic utterances including MP *schon* by L2 speakers at the beginner (A1) level look like. Therefore, we will ask the following research questions for this study to clarify the prosodic aspects of paralinguistic utterances with MP *schon* in L1 and L2:

1. What are the common vs. language-dependent prosodic properties in paralinguistic utterances using MP *schon* by L1 speakers of German (L1) and L2 learners of German (L2) at the A1 level?
2. Are there any common vs. language-dependent properties in the perception of utterances with

paralinguistic information using MP *schon* by L1, L2 speakers, and nonlearners?

To answer the questions above, we conducted production and perception experiments for both L1 speakers and Japanese L2 learners of German, using utterances including MP *schon*. The perception experiment was also conducted for Japanese nonlearners (NL) in order to compare the results with L1 and L2 speakers.

## 2. Experiments in production and perception

In this study, a perception experiment (see 2.1.) was conducted for three different groups: for L1, L2, and NL, using the recorded utterances of the production experiment spoken by L1 speakers. A production experiment (see 2.2.) was performed in parallel, using utterances including MP *schon*.

### 2.1. Experiments in perception

#### 2.1.1. Subjects

As L1 speakers, 22 native speakers of German (11 females and 11 males, aged between 20 and 35) participated in the perception test. As L2 learners of German, 13 Japanese students (8 females and 5 males, aged between 18 and 20) participated in the perception test. Seven of these students took part in the production experiment, which was conducted one month after this perception experiment. As the perception test was conducted for L2 speakers in June 2019, the L2 learners had been learning German for only two months, at the total beginner level, as they had all just begun to learn German in April 2019. They have been learning with the same textbook, “Menschen A1.1” [11], where the word *schon* has been introduced not as a MP but only as a temporal adverb meaning “already.” As a further control group, 18 Japanese nonlearners of German (9 females and 9 males, aged between 18 and 22) also took part in the perception experiment.

#### 2.1.2. Materials

As listening stimuli, the recorded utterances of the two short sentences *Peter kommt schon* ‘Peter will come’ and *Peter kommt schon zum Unterricht* ‘Peter will come to the lesson’ by L1 speakers of German were used for the perception experiment. For each sentence, we had 90 utterances in the three different paralinguistic contexts. By using these stimuli, a perception program was developed by means of Praat ([12]), so that the subjects could listen to the randomized stimuli three times in total.

#### 2.1.3. Procedure

After explaining the experiment, all the subjects were individually asked to listen to the stimuli and to decide which of the three given intentions, “conviction,” “reservation,” or “rebuttal,” the utterance sounded most appropriate for. The subjects listened to the stimuli one by one through headphones connected to a PC. There were two sections, with each section corresponding to each sentence. In each section, the subjects listened to 270 stimuli in total. Having taken a short break after the first section, the subjects then started on the second section. The subjects’ answers were automatically saved to a text file.

### 2.2. Experiments in production

#### 2.2.1. Subjects

As L1 speakers, 10 native speakers of German (5 females and 5 males, aged between 20 and 40), were asked to participate in the experiment. As Japanese L2 speakers, 10 Japanese learners of German (5 females and 5 males, aged between 18 and 20) participated in the experiment. They were all first-year university students. They began to learn German as a foreign language in an intensive course three months before the experiment was conducted (April 2019). They had no experience of living in a German-speaking country.

#### 2.2.2. Materials

Two short sentences containing MP *schon*, (1) *Peter kommt schon* and (2) *Peter kommt schon zum Unterricht* were used in the experiment. In (1), the position of MP *schon* is at the end of the sentence, while in (2), MP *schon* is located in the middle of the sentence. According to our previous studies ([5]), three different contexts corresponding to the three major kinds of paralinguistic information “conviction,” “reservation,” and “rebuttal” are created for the two sentences. Figure 1 shows the short sentence “Peter kommt schon” embedded in “rebuttal.”

*A und B haben sich mit Peter verabredet und warten lange auf ihn. A glaubt, dass Peter heute nicht kommt. B kennt Peter besser und glaubt, dass Peter kommt. ‘A and B have an appointment with Peter and wait for him for a long time. A believes that Peter will not come today. B knows Peter better and believes that Peter will come.’*  
*A: Peter ist unzuverlässig. Er kommt heute nicht. ‘Peter is unreliable. He won’t come today.’*  
*B: Peter kommt schon. Ich kenne ihn viel besser! ‘Peter WILL come. I know him much better!’*

Figure 1: Example of the short sentence “Peter kommt schon” embedded in the context of “rebuttal.”

#### 2.2.3. Procedure

The recording of L1 speakers was made in the phonetics laboratory at the Martin-Luther-Universität Halle-Wittenberg, while the recording of Japanese L2 learners of German was made in the phonetics laboratory at Waseda University. In both recordings, we adopted the same procedure as in our previous studies ([3], [4], [5]). For each subject, nine utterances including the three different intentions were recorded, giving 18 utterances in total for each subject for the two short sentences.

The recorded data were analyzed acoustically using Praat ([12]) in the form of annotation files. On the basis of [5], prosodic properties such as duration, F0 and intensity of the whole utterances, and the accented syllable were considered in the acoustic measurements. In measuring pitch and intensity, the overall mean utterance values of pitch (in semitones) and intensity (in dB) of each subject were used as the standard value. The difference (+/-) between the standard and the measured values within each subject was considered as the relative value of each measured variable. In addition to the measured variables of prosody adopted in [5], the F0 range from F0 peak of the accented syllable to the end of the utterance (F0 slope after the pitch accent) was considered as a new variable in this study. Furthermore, the F1 and F2 of the accented vowels in the utterances as a vowel quality regarding paralinguistic information were also taken into consideration.

### 2.3. Results

#### 2.3.1. Perception of different paralinguistic utterances containing the German MP *schon* by L1, L2, and NL.

Table 1 shows the percentage of answers in the perception experiment.

Table 1: Percentage of answers in the perception experiment. (Percentage of correct answers is marked in yellow)

Sentence	Paralinguistic meaning	Group	Answer		
			Conviction	Reservation	Rebuttal
1	Conviction	L1	73.1%	8.7%	18.2%
		L2	58.1%	11.9%	30.0%
		NL	59.3%	12.9%	27.8%
	Reservation	L1	21.5%	58.5%	20.0%
		L2	31.6%	46.3%	22.1%
		NL	35.0%	44.7%	20.3%
	Rebuttal	L1	36.6%	16.0%	47.4%
		L2	37.2%	14.7%	48.1%
		JP	44.2%	20.6%	35.2%
2	Conviction	L1	73.2%	2.5%	24.3%
		L2	59.7%	10.5%	29.7%
		NL	63.4%	12.7%	24.0%
	Reservation	L1	18.3%	64.0%	17.7%
		L2	39.5%	33.1%	27.5%
		NL	39.0%	37.9%	23.1%
	Rebuttal	L1	27.4%	29.7%	42.9%
		L2	47.5%	16.1%	36.4%
		NL	48.6%	21.9%	29.6%

The percentage of correct answers regarding the three paralinguistic meanings “conviction,” “reservation,” and “rebuttal” in the two sentences among the three subject groups L1, L2, and NL was analyzed statistically by ANOVA and Tukey's multiple comparison test. The results from the statistical analysis show the percentage of correct answers in “conviction” and “reservation” was significantly higher in L1 than in L2 and NL ( $p < .05$ ). In the case of “rebuttal,” on the other hand, significant differences in the percentage of correct answers were not found between L1 and L2 but between L1/L2 and NL ( $p < .05$ ).

#### 2.3.2. Prosodic and phonetic properties of paralinguistic utterances containing MP *schon* in L1 and L2

On the basis of the results from the perception experiment (see 2.3.1.) in L1, a limited number of utterances were picked up, of which the percentage of correct answers showed more than 70% by L1 subjects, in order to look at the crucial prosodic and phonetic properties in the utterances with MP *schon*. Table 2 shows the well-identified utterances by L1 in the perception experiment.

Table 2: Utterances, of which the percentage of correct answers by L1 showed more than 70%.

Sentence	Paralinguistic meaning	Number of utterances	%
1	Conviction	22	73.3
	Reservation	15	50
	Rebuttal	8	26.7
2	Conviction	20	66.7
	Reservation	20	66.7
	Rebuttal	5	16.7

In the following results from the production experiment, the prosodic properties of the well-identified utterances by L1

shown above are introduced and compared with those of the utterances by L2 Japanese learners of German. In the utterances spoken by L1, most of the well-identified utterances in the context of “conviction” (100% for sentence 1, 90% for sentence 2) and “rebuttal” (100% for sentence 1 and sentence 2) carried a pitch accent on the verb, not on MP *schon*, while a pitch accent was located on MP *schon* in most of the utterances in the case of “reservation” (93% for sentence 1, 95% for sentence 2). In the utterances spoken by L2, MP *schon* carried pitch accents in most cases in all three contexts (83% in “conviction” and “reservation,” and 97% in “rebuttal” for sentence 1, 77% in “conviction,” 87% in “reservation,” and 83% in “rebuttal” for sentence 2). In the further analysis, we focused primarily on the two paralinguistic meanings “conviction” and “rebuttal,” in both of which, when spoken by L1, a pitch accent was located on the verb *kommt*.

Table 3 shows the results from the statistical analysis (ANOVA and Tukey's multiple comparison test) of the measured variables among three paralinguistic meanings.

Table 3: Measured variables in L1 and L2 between “conviction” (a) and “rebuttal” (c) (\*\*\*) means  $p < .001$ ; \*\* means  $p < .01$ ; \* means  $p < .05$ .

Measured Variables	L1	L2
Utterance duration	c>a ***	c>a ***
Duration of each syllable	c>a ***	c>a *
Utterance pitch maximum	a>c ***	n.s.
Utterance pitch mean	n.s.	c>a ***
Utterance pitch range	a>c ***	n.s.
Onset F0 of accented syllable	c>a ***	c>a ***
Pitch maximum of accented syllable	c>a ***	c>a ***
Pitch range from onset to peak	c>a ***	c>a ***
Pitch range from F0 peak to the end of utterance	c>a ***	c>a ***
Intensity maximum	c>a ***	a>c ***
Vowel F1 of accented syllable	c>a ***	c>a ***
Vowel F2 of accented syllable	c>a *	n.s.

As Table 3 shows, similar prosodic properties in the utterance duration, the duration of each syllable, and the pitch values of the accented syllable *kommt* could be found in the utterances spoken by L1 and L2, while pitch in the whole utterances and intensity showed language-dependent properties as in [5]: In “rebuttal” spoken by L1, pitch maximum ( $p < .001$ ) and range ( $p < .001$ ) were significantly lower, but pitch mean was significantly higher ( $p < .001$ ) in “rebuttal” spoken by L2. The intensity maximum of the whole utterance was significantly higher ( $p < .001$ ) in “rebuttal” than in “conviction” in L1, and vice versa in L2 ( $p < .001$ ). Pitch range from F0 peak to the end of utterance (F0 slope after the F0 peak of the pitch accent) was in both L1 and L2 utterances significantly larger ( $p < .001$ ) in “rebuttal” than in “conviction” in sentence 1. For sentence 2, however, a significant difference could be observed only in L1 ( $p < .001$ ) utterances.

Regarding vowel quality, vowel F1 ( $p < .001$ ) and F2 ( $p < .05$ ) of the accented vowel [ɔ] in the verb “kommt” were significantly higher in “rebuttal” than in “conviction” spoken by L1, as Figure 2 shows. In L2, vowel F1 of [o:] in MP *schon* was significantly higher ( $p < .05$ ) in “rebuttal” than in “conviction,” but no significant difference in vowel F2 of [o:] was found (Figure 3).

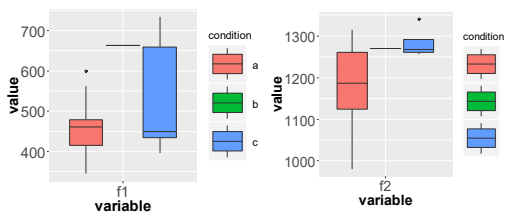


Figure 2: F1 (left) and F2 (right) (in Hz) of the accented vowel [ɔ] in verb “kommt” in the case of “conviction” (a) and “rebuttal” (c) spoken by L1 in sentence 2.

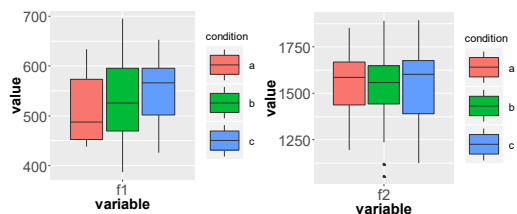


Figure 3: F1 (left) and F2 (right) (in Hz) of the accented vowel [o] in MP *schon* in the case of “conviction” (a), “reservation” (b), and “rebuttal” (c) spoken by L2 in sentence 2.

### 3. Discussion

#### 3.1. Common and language-dependent prosodic and phonetic properties in paralinguistic utterances using the German MP *schon* by L1 speakers and L2 learners of German at the A1 level

According to the results described above, the duration of the whole utterance and the pitch properties of the accented syllable spoken by both L1 and L2 seem to have common functions in distinguishing the paralinguistic utterances between “rebuttal” and “conviction.” However, the results in the syllable duration seem to be different between L1 and L2. More precise analysis about the syllable duration is needed to conclude a similar function of duration regarding paralinguistic utterances by L1 and L2 speakers.

In the vowel F1, we found similar results both in L1 and L2: The vowel F1 of the accented syllable was in both L1 and L2 significantly higher in the case of “rebuttal” than in “conviction.” This result suggests that the vowel F1 is one of the common phonetic properties regarding the speaker’s attitude of “rebuttal.” [7] also report a higher vowel F1 and F2 in the case of “suspicion,” which corresponds to the results in this study. However, we could not find any specific properties in the vowel F2 in the utterances spoken by L2, while in L1 the vowel F2 of the accented syllable was significantly higher in “rebuttal” than in “conviction.” The L2 speakers pronounced the accented vowel of MP *schon* in most cases in a way that was wrong, using [ø:] instead of [o:]. This is perhaps because the L2 learners were not yet used to pronouncing the word *schon* and therefore mistook it for the word *schön*, ‘beautiful,’ although the word *schon* as a temporal adverb has already been introduced in lessons. Further production experiments should be performed with L2 learners, who continue to learn German and can proceed to a higher level (from A1 to A2) in German as a foreign language.

Considering the pitch and intensity properties in the whole utterances, we found language-dependent characteristics of the

paralinguistic utterances by L1 and L2 speakers: In L1 utterances, the maximum pitch was significantly lower, while the maximum intensity was significantly higher in “rebuttal” than in “conviction.” In L2 utterances, on the other hand, the mean pitch was significantly higher in “rebuttal” than in “conviction,” while the intensity maximum was lower in “rebuttal.”

#### 3.2. Common and language-dependent properties in the perception of paralinguistic utterances using MP *schon* by L1 and L2 speakers

As the results from the perception experiment show, the overall tendencies of answers appeared to be similar among the three subject groups L1, L2, and NL, although the percentage of correct answers in “conviction” and “reservation” was in L1 significantly higher than in the L2 groups. It should also be noted that the utterances in the case of “rebuttal” were answered more often correctly by L1 and L2 than by NL, while there were no significant differences in the perception of “conviction” and “reservation” utterances between L2 and NL. These results suggest that some kinds of paralinguistic information such as “rebuttal” show language-dependent properties, which can be gradually acquired in learning German as a foreign language. However, more analysis should be done in order to clarify common vs. language-dependent aspects in the perception of paralinguistic utterances using MP *schon*.

### 4. Conclusions

In this study, we investigated prosodic aspects of utterances conveying paralinguistic information using MP *schon* by L1 speakers and L2 Japanese learners of German. On the basis of the results from the production and perception experiments, both common (duration of the whole utterance, pitch of the accented syllable, and vowel F1) and language-dependent (duration of each syllable, pitch and intensity of the whole utterance and vowel F2) prosodic properties could be observed.

The results from the perception experiment in the three subject groups L1, L2, and NL show similar tendencies, for example, in distinguishing the two attitudes “rebuttal” and “conviction.” “Rebuttal” was identified more correctly by L1 and L2 than by NL, which suggests there are language-specific aspects in “rebuttal” that can only be perceived through competence in German. On the other hand, the results suggest that there is a language-independent aspect in perceiving the paralinguistic utterance of “conviction.” Further studies are needed on the basis of the evaluation test for the recorded utterances spoken by L2. More production and perception experiments with the same L2 learners at a higher level should also be conducted in the future.

### 5. Acknowledgements

I would like to thank Takayuki Konishi (Waseda University, Tokyo) for developing the Praat script for the acoustic analysis. I also thank Prof. Dr. phil. habil. Dr. h.c. Ursula Hirschfeld for permission to use the phonetics laboratory at the Martin-Luther-Universität Halle-Wittenberg.

This research was supported by JSPS KAKENHI Grant Number JP19K00893, iLaSS (Waseda Univ. Institute of Language and Speech Science), Waseda University Grants for Special Research Projects (“Tokutei Kadai”) 2019C-005 and Waseda Institute of Political Economy (WINPEC) Special Research Project “Social Behaviour and Language Choice.”

## 6. References

- [1] H. Weydt, *Abtönungspartikel: die deutschen Modalwörter und ihre französischen Entsprechungen*. Bad Homburg: Gehlen, 1969.
- [2] M. Ikoma, *Prosodische Eigenschaften der deutschen Modalpartikeln*. (Schriftenreihe PHONOLOGIA, Band 103). Hamburg: Dr. Kovač, 2007.
- [3] M. Ikoma, “Prosody of German Modal Particles and Japanese Sentence Final Particles,” in *Proceedings of the 31<sup>st</sup> General Meeting of the Phonetic Society of Japan (PSJ)*, 2017, pp. 216-221.
- [4] M. Ikoma, „Prosodie und Bedeutung der unbetonten und betonten Modalpartikel schon“, in *Akten des 44. Linguisten-Seminars, Tokyo 2016*. München: Iudicium, pp. 53-68, 2018.
- [5] M. Ikoma and T. Konishi, “Prosody of modal particles – production and perception of L1 and L2 German speech,” in *Proceedings of the 32<sup>nd</sup> General Meeting of the Phonetic Society of Japan (PSJ)*, 2018, pp. 285-290.
- [6] Y. Asano and M. Gubian, ““Excuse meeee!”: (Mis)coordination of lexical and paralinguistic prosody in L2 hyperarticulation,” in *Speech Communication* 99, 2018, pp.183-200.
- [7] H. Mori, K. Maekawa and H. Kasuya, *What does speech convey? – Speech Science of Emotion, Paralinguistic Information, and Speaker Individuality* (eds. The Acoustical Society of Japan) Tokyo: Corona, 2014.
- [8] F. Burkhardt, A. Paeschke, M. Rolfes, W. Sendlmeier and B. Weiss, “A Database of German Emotional Speech,” in *Proceedings of Interspeech*, 2005, pp. 1517-1520.
- [9] M. Kienast and W. F. Sendlmeier, “Acoustical Analysis of Spectral and Temporal Changes in Emotional Speech,” in *Proceedings of the ISCA-Workshop on “On Speech and Emotion” in Belfast, Newcastle, Northern Ireland, UK*, 2000, pp. 92-97.
- [10] A. Paeschke and W. F. Sendlmeier, „Prosodic Characteristics of Emotional Speech: Measurements of Fundamental Frequency Movements,“ in *Proceedings of the ISCA-Workshop on “On Speech and Emotion” in Belfast, Newcastle, Northern Ireland, UK*, 2000, pp. 75-78.
- [11] S. Evans, A. Pude and F. Specht, *Menschen A1.1 Kursbuch Deutsch als Fremdsprache*, Hueber, 2018.
- [12] P. Boersma and D. Weenink, *Praat: doing phonetics by computer* [Computer program]. Version 6.0.43, retrieved 13 September 2018 from <http://www.praat.org/>, 2018.