



# Attractiveness of French voices for German listeners – results from native and non-native read speech

*Jürgen Trouvain and Frank Zimmerer*

Language Science and Technology, Saarland University, Saarbrücken, Germany

trouvain|zimmerer@coli.uni-saarland.de

## Abstract

This study investigated how the perceived attractiveness of voices was influenced by a foreign language, a foreign accent, and the level of fluency in the foreign language. Stimuli were taken from a French-German corpus of read speech with German native speakers as raters. Additional factors were stimulus length (syllable or entire sentence) and sex (of the raters and speakers). Results with German native raters reveal that stimuli spanning just a syllable were judged significantly less attractive than those containing a sentence, and that stimuli from French speakers were assessed as more attractive than those of German speakers. This backs the cliché that French has an attractive image for German listeners. An analysis of the best vs. the worst rated sentences suggests that an individual mix of voice quality, disfluency management, prosodic behaviour and pronunciation precision is responsible for the results.

**Index Terms:** attractiveness, French, German, accent, foreign language, fluency

## 1. Introduction

Attractiveness of voice unfolds in various ways of physical and also social attractiveness. It plays a role in sexual encounter for mating and seduction but also in the evaluation of beauty and pleasantness, sympathy, politeness, competence, persuasion and charisma, see e.g. [1-7]. The most frequently investigated phonetic parameters are related to fundamental frequency (e.g. mean pitch and pitch range) and there are many more, related to timing (e.g. vowel duration), vowel quality (e.g. formants and harmonics-to-noise ratio) and voice quality (e.g. jitter and shimmer).

In this study we investigate the effect of i) foreign accent, ii) native (or first) language (L1) and non-native (or second) language (L2), and iii) fluency (which usually differs when using L1 or L2), see e.g. [8]. In our study we use read speech of French and German speakers in both languages rated by German listeners. In addition to the language-dependent questions factors such as stimulus length and sex of the speakers and listeners are taken into account.

### 1.1. Foreign accent

Prior studies show that accents influence attractiveness ratings, e.g. [9] found that foreign accented speakers were rated as less trustworthy which can be seen as a special aspect of attractiveness. In Germany there are various nationally based questionnaires in which people indicate how attractive they consider various regional and foreign accents. Most of these 'studies' (with a low level of linguistic, dialectological or other scientific quality) reveal that for example, the accents spoken in Bavaria are ranked high regarding popularity and sympathy, and those spoken in Saxony are ranked low. Likewise, for non-

native accents, German listeners rank the attractiveness of French high, and that of Russian low. Although these results are backed by serious studies such as [10, 11] it remains unclear how pre-fabricated clichés are still at work when assessing a concrete voice with, say, a French accent.

### 1.2. Language

By far the most widely spoken foreign language in the German-speaking countries is English. In terms of languages taught in schools in Germany, French as L2 ranks on number 2 (with far fewer learners), followed by Spanish as L2 [12]. This picture is also transported by surveys reported in non-scientific media (with Spanish apparently becoming more popular than French). Here again, the cliché of French as a popular language for German speakers hints to an expected high level of attractiveness of French speech. It remains unclear whether this expected effect is a graded one (e.g. French native speech more attractive than French accented German which in turn is more attractive than German accented French) or whether this expected attractiveness is found in general to the native language (i.e., German).

### 1.3. Fluency

Zuta [13] shows that less fluent (male) speakers are rated as less attractive than more fluent speakers in their L1 (German). This suggests that the level of fluency influences the perceived attractiveness of the speaker.

A widely observed phenomenon is that L2 speakers are usually less fluent than L1 speakers, and that L2 beginners are less fluent than L2 advanced learners, e.g. [8]. In speech production this is reflected by longer pauses, pauses at unacceptable locations, and false starts with or without repairs. A repair after a possible re-start can consist of a repetition, a substitution or an insertion, or a combination of those. In spontaneous speech, a filled pause often occurs after the interruption. Usually, a slower articulation rate also contributes to a lower level of fluency.

We thus can expect that the effect of an L2 accent is reinforced by disfluencies. However, disfluencies in L1 are processed and interpreted differently than those in L2, see e.g. [8]. In addition, we can distinguish between utterance fluency and perception fluency, i.e. a sample of speech with many observed disfluencies is not necessarily judged by listeners as very disfluent and thus less attractive.

### 1.4. Stimulus length

Former studies on vocal attractiveness used stimuli with differing length. This ranges from longer narratives, like fairy tales [13], opening statements in job interviews [1], single sentences of diverse length to syllable-sized units, mono-

syllabic words [5, 6], or isolated vowels [2, 4]. Also, a combination of several variable sized stimuli was investigated [5]. Ferdenzi et al. [14] examined stimuli of variable length (single vowel, series of three vowels, two-syllabic word). They found that those stimuli were rated as most attractive which were the longest. Furthermore, after manipulating the duration of stimuli, unmanipulated versions were found as most attractive.

We decided to use very short stimuli with the length of a syllable and longer stimuli with the length of a complex sentence to see whether stimulus length makes a difference for our data. The syllable condition, of course, cannot be used for the investigation of the effect of hesitation on attractiveness.

### 1.5. Sex

Adult hetero-sexual listeners find speakers of the opposite sex more attractive than speakers of the same sex [1-7] which can be explained with a typical mating behaviour. In this setting, female listeners prefer same-sex speakers more than male listeners do, see e.g. [6]. We would expect similar results in this study.

### 1.6. Research questions

From the presented factors our main research questions can be summarised as follows:

- Do German native speakers rate French accented read speech as attractive as German L1 read speech?
- Are the voices of French native speakers reading French as attractive as the German voices in their L1 – or even more attractive?

In addition, we are interested in the following questions:

- Is L2 read speech with disfluencies (hesitations) less attractive than L2 fluent speech?
- Does language proficiency of French of the German raters influence the attractiveness ratings?
- Are the results for short stimuli comparable to those for long stimuli?

## 2. Material

### 2.1. Material and stimuli

A subset of the French-German IFCASL corpus [15] served as material. This corpus contains read speech produced by German and French speakers. Both groups read the entire material in both languages. The level of proficiency in the L2 ranged from beginners to advanced speakers. In total, samples from 40 speakers were used: 20 German and 20 French native speakers. Each of these groups consisted of 10 beginners and 10 advanced learners in the respective L2. In each sub-group were five females and five males. The speakers in each sub-group were randomly selected.

The reading material of the corpus consists mainly of single sentences, but also one text (a translation of "The three little pigs" with 13 sentences in both language versions) of which we selected one specific sentence where pronunciation errors and disfluencies are quite frequent (in German: "Da beschließt der Wolf durch den Schornstein in das Haus zu steigen, aber die kleinen Schweinchen haben einen Kessel mit kochendem Wasser vorbereitet."; in French: "Le loup décide alors de passer par la cheminée mais les petits cochons ont préparé un chaudron d'eau bouillante."). In addition to this longer sentence (SENT

condition), we selected a short section with a size of a syllable (SYLL) from this text. For the German version it was [did] from "Die drei kleinen Schweinchen ..." and for the French version it was [let] from "Les trois petits cochons ..." (English translation: "The three little pigs") from the first sentence of the text each, i.e. a consonant-vowel-consonant sequence (with the last consonant being the onset of the second word) from words in non-prominent position. The duration of the SENT stimuli in the L1 was 8,213 ms on average (standard deviation (sd): 1,437 ms), in the L2 the mean duration was 10,135 ms (sd: 2,228 ms), the mean duration of the SYLL stimuli was 170 ms (sd: 38 ms) in the L1, and 206 ms (sd: 56 ms) in the L2. The total number of stimuli was 160 (4 x 40 speakers).

### 2.2. Rating experiment

The rating experiment was performed with 18 participants using the web tool "SoSci Survey" [16]. Questions before starting the test concerned the L1(s), reported hearing deficiency (none), sexual orientation, and self-reported proficiency of French (7 steps, from no skills to native-like). One bi-sexual and one homo-sexual participant were excluded from further analysis in order to have enough subjects in one condition of sexual orientation, here hetero-sexual. At the end, ratings of 16 German listeners were analysed: 7 males and 9 females, mean age: 33 years (range from 20 to 62 years); language proficiency ranged from advanced (n=4) over beginners (n=7) to no skills (n=5). The task itself was presented in German as follows: "You now hear samples of many different voices who speak either German or French. Some of the samples are very short. Click on the play button to listen to a sample. After listening you have to select a point on a scale expressing how attractive you found the voice in the sample. Then click 'continue' and the next item will be displayed."

The stimuli were presented in a fixed randomised order. The test took about 30 minutes on average.

### 2.3. Analysis

The scale ranged between 1 and 101 and was treated as continuous variable for the statistical analysis. This analysis was performed with JMP [17].

In addition, the 5 worst rated and 5 best rated sentences were auditorily assessed by both authors with regard to peculiarities such as extreme pitch range, disfluencies, extreme articulation rate, prominent voice quality features, increased nasality or pronunciation errors. This rather informal procedure was chosen to shed light on possible explanations for the results of the rating study.

## 3. Results

A linear mixed model was calculated with rating as dependent variable, speaker and item as random variable, listener sex, speaker sex, item language, native language, stimulus length (SENT or SYLL) and all two-way interactions as well as French proficiency (no skills, beginners, advanced) and hesitation (with disfluency or not) without interactions as fixed factors.

### 3.1. Stimulus length

The length of the stimuli has a significant impact on the attractiveness ratings. Short stimuli of the size of a syllable are judged as less attractive than those of a size of a longer sentence. Figure 1 shows the significant difference between the SENT (mean: 53.6) and the SYLL condition (mean: 47.0).

To investigate a possible effect of hesitation on the attractiveness ratings the SYLL condition could not be used anymore because there was no hesitation in any of the items. Furthermore, some participants reported insecurity about their response to SYLL condition in which it can also be questioned that listeners recognised the L1 of the speaker. Moreover, a longer stimulus such as in the SENT condition allows a more global rating and therefore a better consideration of prosodic aspects.

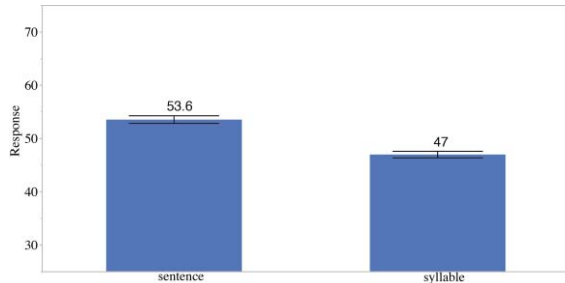


Figure 1: Mean of ratings in the condition SENT (left) and SYLL (right).

### 3.2. Sex

After removing the SYLL stimuli, the analysis was recalculated without stimulus length. The results can be found in Table 1, the effects will be reported in subsequent sections.

Table 1: The results of the linear mixed model with rating as dependent variable.

Variable	F value	p-value
Listener Sex*Speaker Sex	123.4061	<.0001*
Speaker L1	42.9005	<.0001*
Speaker L1*Speaker Sex	7.3117	0.0085*
Speaker Sex	5.0285	0.0280*
Item Language	3.3704	0.0705
Hesitation	1.8212	0.1814
L2 Proficiency	0.2279	0.7995
Listener Sex	0.0298	0.8658
Speaker L1*Item Language	0.0267	0.8707

As expected, listeners rate the voices of the opposite sex as more attractive than voices of the own sex (see Figure 2). In addition, male listeners judge male voices as less attractive than do female listeners with female voices.

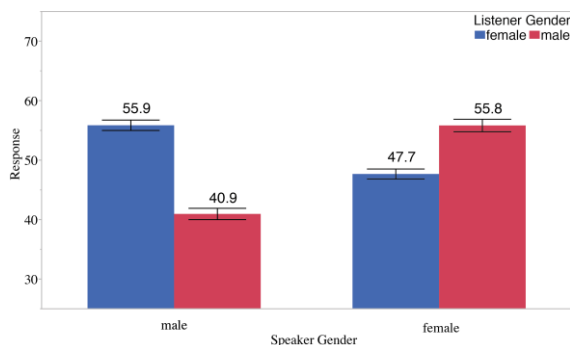


Figure 2: Ratings of male (M, left) and female participants (W, right). Speakers to be rated are in blue (female) and red (male).

### 3.3. French accent

The German native subjects rated French accented speech as significantly *more* attractive than German L1 speech.

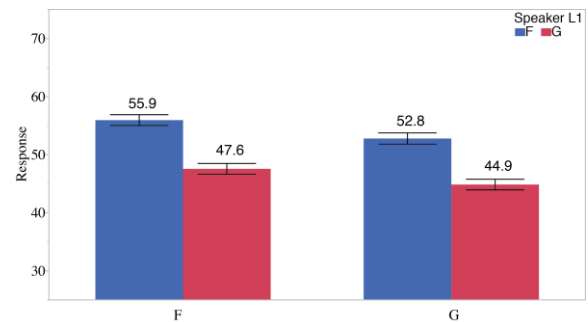


Figure 3: Ratings of the item language in which the sentence was read: French (F, left) and German (G, right). L1 of the speakers in blue (French) and in red (German).

### 3.4. Language

It is visible in Figure 3 that our French speakers were rated as *more* attractive than our German speakers. This finding is valid for both language versions. As can be seen in Table 1, there is no significant interaction between speaker L1 and item language despite a small numerical difference towards German accented French.

Figure 4 shows the tendency of the German listeners to rate the French speakers (black dots) as more attractive than the German speakers (light grey dots) in their L1 speech *and* their L2 speech.

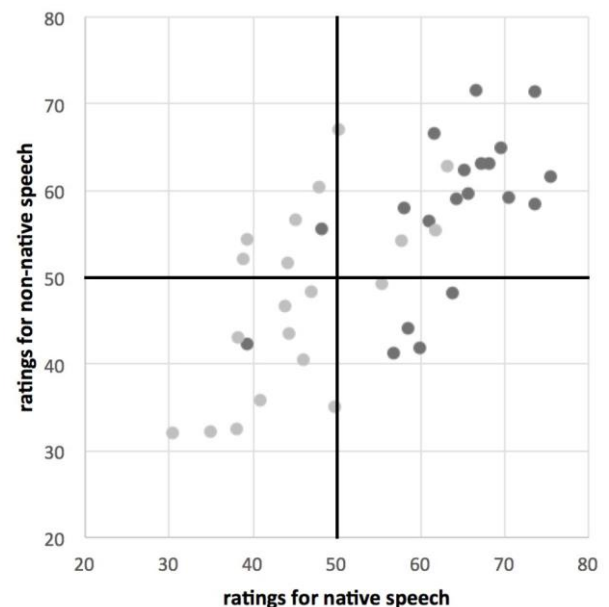


Figure 4: Ratings (the higher the value the more attractive) of all 40 speakers. Native speech on the x-axis and non-native speech on the y-axis. French speakers in black, German speakers in light grey.

### 3.5. Disfluencies

There was no significant effect of disfluency/hesitation.

### 3.6. L2 proficiency of the raters

The French proficiency levels of the subjects had no significant effect on the investigated variables (see Table 1).

### 3.7. Auditory analysis of the extreme samples

The five worst rated sentences (W1-5) were all by German speakers who produced French, whereas the five most attractive rated samples (A1-5) stem from French female speakers only (compare also Figure 4). The auditory analysis performed by two trained phoneticians shows an extremely diverse picture regarding the examined parameters. The speaker of W1 performs in his L2 speech a word-by-word reading style with a strong staccato rhythm, leaving an auditory impression of being very strained with a high tension, plus some pronunciation errors. In contrast, the L1 speech of the same speaker (W2) reflects a reduced jaw opening. This feature can also be observed for the female speaker of W5, accompanied with a reduced pitch range and a strong degree of nasality which leaves the impression that the speaker was bored. A typical feature of W3 (male) is a very rigid German segmental and rhythmic accent, and W4 (female) shows markers of a non-professional staging, pressed voice quality, nasality and regional German accent.

A3 and A5 are produced by the same female speaker. In her L2 sample she also shows clear disfluencies but she manages easily to repair this difficulty by producing only a short silent pause (300 ms) after the interruption point. Her articulation rate as well as her segmental skills hint to a high perceived level of spoken language proficiency. Interestingly, also A4 is not fluent in her L2 speech with a repair (marked by a short silent pause) and a clear elongation but with a stable articulation rate. The particular feature of the speaker of A2 is a breathy voice quality whereas A1 shows no prominent phonetic feature.

## 4. Discussion

Results on sex are as expected: male listeners find female voices more attractive than male voices, and vice versa, i.e. female listeners prefer male voices over female voices.

German native speakers rate French speakers producing foreign accented read speech as more attractive than German L1 read speech. This result is contrary to the findings in [9]. It might be that the L1 of the speakers with the foreign accent plays a substantial role and that L2 speakers differ in their perceived attractiveness depending on their L1. Overall, we do not see any effect of French language proficiency on the German participants' attractiveness ratings.

The voices of French native speakers reading French are not only equally attractive as the German voices in their L1, they are even more attractive. Apparently, the stereotypical picture of French as a popular and sympathetic language for German speakers is fully reflected by these results.

Our expectations regarding the attractiveness of L2 read speech containing disfluencies were not, or only partially, valid. We found that disfluent or hesitant speech is not unattractive per se. Two of the top five attractive samples contained disfluencies whereas not each of the least attractive samples was hesitant. The auditory analysis suggests that a good control of how to handle disfluencies is more important than the pure

existence of disfluencies themselves. Thus, the consideration of disfluencies in L2 teaching and L2 assessment should be refined into the direction of disfluency management rather than a strict avoidance of disfluencies.

Results on stimulus length show that short stimuli are *not* comparable to those for long stimuli, at least not in our data. Therefore we recommend to use stretches longer than syllables for rating experiments as ours.

As in other studies on voice attractiveness the voice quality can play an important role [6]. In an investigation of the laryngeal phonatory settings of the corpus data also used in this study [18] it could be shown that the behaviour of German speakers is characterised by more strained adduction of the vocal folds whereas French speakers show more incomplete glottal closure.

## 5. Conclusions

In a recent study by Babel et al. [6] the authors showed that voice attractiveness is more complex than previously assumed. This study adds a further bit to the complexity by stating that

- i) foreign accented speech can be *more* attractive than native accented speech,
- ii) speakers of a foreign language can sound *more* attractive than speakers of the native language of the listeners, and
- iii) that disfluencies do *not per se* lead to less attractiveness.

Foreign language, obviously, has an effect on the perceived attractiveness, be it as foreign accent or as native speech. It would be interesting to investigate other language pairs with speech recorded in both languages by the same speaker. This could include on the one hand languages with a stereotypical lower prestige (e.g. Russian for German listeners) but also with languages that presumably are unknown to the raters, i.e. where a cliché is arguably not existent.

This study investigated read speech and is therefore very limited in the number of factors which we analysed. In spontaneous speech listeners probably use many other important factors to assess the attractiveness of a speaker. This could involve other types of disfluencies. Furthermore, discourse structure, lexical choice and syntactic complexity will play a role as well as behaviour strategies in talk-in-interaction, see also [19].

## 6. References

- [1] M. Zuckerman and R.E. Driver, "What sounds beautiful is good: The vocal attractiveness stereotype," *Journal of Nonverbal Behavior*, vol. 13, pp. 67–82, 1989.
- [2] D.R. Feinberg, B.C. Jones, L.M. DeBruine, F.R. Moore, M.J. Law Smith, R.E. Cornwell, B.P. Tiddeman, L.G. Boothroyd, and D.I. Perrett, "The voice and face of woman: One ornament that signals quality?" *Evolution and Human Behavior*, vol. 26, pp. 398–408, 2005.
- [3] B. Weiss and F. Burkhardt, "Voice attributes affecting likability perception," *Proc. Interspeech*, Makuhari, pp. 1934–1937, 2010.
- [4] D.E. Re, J.J.M. O'Connor, P.J. Bennett, and D.R. Feinberg, "Preferences for very low and very high voice pitch in humans," *PLoS one*, vol. 7, no. 3, paper e32719, 2012.
- [5] P. Abend, L.S. Pflüger, M. Koppensteiner, M. Coquerelle, and K. Grammer, "The sound of female shape: a redundant signal of vocal and facial attractiveness," *Evolution and Human Behavior* vol., 36, no. 3, pp. 174–181, 2014.

- [6] M. Babel, G. McGuire, and J. King, "Towards a more nuanced view of vocal attractiveness," *PLoS one*, vol. 9, no. 2, paper e88616, 2014.
- [7] H. Quené, G. Boomsma, and R. van Erning, "Attractiveness of male speakers: effects of voice pitch and of speech tempo," *Proc. Speech Prosody*, Boston, pp. 1086-1089, 2016.
- [8] H.R. Bosker, H. Quené, T.J.M. Sanders, and N. de Jong, "The perception of fluency in native and non-native speech," *Language Learning*, vol. 64, pp. 579-614, 2014.
- [9] S. Lev-Ari and B. Keysar, "Why don't we believe non-native speakers? The influence of accent on credibility," *Journal of Experimental Social Psychology*, vol. 46, pp. 1093-1096, 2010.
- [10] L.M. Eichinger, A.-K. Gärtig, A. Plewnia, J. Roessel, A. Rothe, S. Rudert, C. Schoel, D. Stahlberg, and G. Stickel, "Aktuelle Spracheinstellungen in Deutschland. Erste Ergebnisse einer bundesweiten Repräsentativumfrage," Mannheim: Institut für Deutsche Sprache and Universität Mannheim, 2009.
- [11] J. Setinieri, "Soziale Akzeptanz unterschiedlicher Normabweichungen in der L2-Aussprache Deutsch," *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, vol. 16, no. 2, 66-80, 2011.
- [12] A. Malecki, "Schulen auf einen Blick – Ausgabe 2016," Statistisches Bundesamt, Wiesbaden, 2016.
- [13] V. Zuta, "Phonetic criteria of attractive male voices," *Proc. International Congress of the Phonetic Sciences (ICPhS)*, Saarbrücken, pp. 1837-1840, 2007.
- [14] C. Ferdenzi, S. Patel, I. Mehu-Blantar, M. Khidasheli, D. Sander, and S. Delplanque, "Voice attractiveness: Influence of stimulus duration and type," *Behavior Research Methods*, vol. 45, pp. 405-413, 2013.
- [15] J. Trouvain, A. Bonneau, V. Colotte, C. Fauth, D. Fohr, D. Juvet, J. Jügler, Y. Laprie, O. Mella, B. Möbius, and F. Zimmerer, "The IFCASL corpus of French and German non-native and native read speech," *Proc. 9th Language Resources and Evaluation Conference (LREC)*, Portorož, pp. 1333-1338, 2016.
- [16] SoSci Survey <https://www.soscisurvey.de/>
- [17] SAS, "JMP," 13 ed. Cary (NC): SAS Institute, 2016
- [18] M. Pützer, F. Zimmerer, W. Wokurek, and J. Jügler, "Evaluation of phonatory behavior of German and French speakers in native and non-native speech," *Proc. Interspeech*, San Francisco, 520-524, 2016.
- [19] B. Weiss and K. Schoenenberg, "Conversational structures affecting auditory likeability," *Proc. Interspeech*, Singapore, pp. 1791-1795, 2014.