

# Pitch synchrony as an effect of perceived attractiveness and likability

Jan Michalsky

*Institute of German Studies, University of Oldenburg, 26129 Oldenburg, E-Mail: j.michalsky@uol.de*

## Abstract

While there exists a growing body of research on acoustic correlates of attractiveness and likability, there is little to no research on how the perception of these impressions affects a listener's acoustic features. In this study we investigate the impact of perceived attractiveness and likability as aspects of the overall impression of an interlocutor on a speaker's pitch features. In addition, we examine pitch synchrony, i.e. whether perceived attractiveness and likability affect the relationship between the pitch features of the speaker and the interlocutor, both on a global and a local level. We investigated 10 female and 10 male speakers in 98 spontaneous dialogues of 15 minutes each in a speed dating setting. The results of the study suggest that 1) perceived attractiveness and likability systematically affect a speaker's pitch, 2) attractiveness and likability constitute two separate impressions that can and need to be distinguished, 3) the presence and degree of pitch synchrony is affected by perceived attractiveness and likability, and 4) the effects differ greatly between the global and the local level.

## 1. Introduction

Attractiveness and likability play a vital role in social interaction. Accordingly, previous studies examined which pitch features render a speaker's voice attractive or likable (for an overview see [1] and [2]). The study at hand alters the perspective by examining the effects of a speaker's impression of an interlocutor regarding attractiveness and likability on said speaker's pitch. In previous studies on attraction, male speakers were found to lower their pitch when directing their speech to a more attractive listener of the opposite sex [3]. While [4] found that female speakers raised their pitch when speaking to a more attractive male listener, other studies found females to lower their pitch [3]. In a previous study on global pitch effects of perceived attractiveness and likability in speed dating conversations, participants of both sexes expanded their pitch range when talking to a more attractive speaker of the opposite sex and raised their register when talking to a more likable speaker of the opposite sex [1]. Aside from [1], pitch features specifically connected to the perceived impression of likability have not been explicitly studied so far. However, likability or related social meanings have been associated with the phenomenon of prosodic entrainment [cf. 5, 6, 7]. Prosodic entrainment describes the interdependency of prosodic features of two interlocutors [5, 6, 7]. Levitan [5] distinguishes two dimensions and three types of entrainment. *Global entrainment* regards features spanning the entire interaction (e. g. f0 mean) while *local entrainment* describes entrainment at turn breaks. Regarding the types of entrainment, speakers can either become more similar in absolute terms (*similarity*), increase their similarity over the course of the conversation (*convergence*), and/or adjust their

prosodic features relative to each other thereby imitating dynamic changes without necessarily becoming more similar in absolute terms (*synchrony*). Entrainment has been found to correlate with stronger collaboration or rapport [8] and is generally linked to more successful conversations. In this study we focus on pitch synchrony. Different to Levitan [6] we distinguish between global and local synchrony. *Local synchrony* covers the original definition of synchrony and refers to speakers dynamically changing their pitch at turn breaking points [6, 7]. As *global synchrony* we define speakers adjusting their mean pitch features over the course of a whole conversation to their respective interlocutors' pitch features compared across different interlocutors and different conversations. So far, the interaction between local synchrony and perceived attractiveness or likability has not been investigated. In a previous study, [1] found effects of global synchrony for the f0 mean, the median, and the pitch range, but no interaction with perceived attractiveness or likability.

The interdependence of attractiveness and likability is a problem rarely considered in the respective studies. Since it can be assumed that an interlocutor's likability can influence the impression of attractiveness and that visual attractiveness can influence the impression of likability, we explicitly treat both as separate factors. We aim to find phonetic features that are exclusive to one or the other, thereby making effects of attractiveness and likability distinguishable. As mentioned above, [1] showed that attractiveness and likability can affect different phonetic features which supports the necessity of separating the two. We seek to answer the following research questions: 1) Does the perception of attractiveness/likability correlate with global or local features of the speaker's pitch? 2) Does the perception of attractiveness/likability affect the presence or degree of global or local synchrony of a speaker's pitch features?

## 2. Method

### 2.1 Speakers

Ten female and 10 male students from the University of Oldenburg participated in the study as paid volunteers. All subjects were monolingual speakers of High German aged between 19 and 28 years. They all grew up in (northern) Germany, i.e. share a common cultural background. Only heterosexual singles were included in the study. All subjects were unacquainted and had no interactions prior to the experiment.

### 2.2 Procedure

We investigate effects of attractiveness and likability in spontaneous natural conversations. The subjects participated in a speed dating setting, which was altered to meet the research objective. To this end, each participant was paired with each of the 10 participants of the opposite sex resulting

in a total of 100 opposite-sex combinations from which 98 could be included in the acoustic analysis. The subjects were placed in a quiet room and engaged in a conversation with no restrictions to the topics. A note with sample topics was provided in case participants had difficulties starting the conversation. Each conversation lasted between 15 and 20 minutes. Prior to the first verbal interaction as well as immediately after each conversation, participants received a questionnaire and were asked to evaluate their interlocutor in terms of purely visual attractiveness as well as perceived likability on a 10-point Likert scale. The participants were given the necessary privacy and their ratings were not revealed to the respective interlocutor. Recordings were made in stereo using a portable digital recorder (Tascam HD P2) at a sampling rate of 48 kHz and 24-bit resolution with one head-mounted microphone (DPA 4065 FR) per speaker.

### 2.3 Acoustic analysis

The acoustic analysis was carried out using Praat [9]. For the analysis of global effects, audio tracks were separated for each speaker. Subsequently, all filled pauses, laughter, overlaps as well as the interlocutor's speech parts were manually silenced to preserve the time structure of the recordings. For local effects, the speech parts of all speakers were segmented into interpausal units [cf. 6]. The two interpausal units adjacent to a turn break with a speaker transition were used for the acoustic analysis. While all 98 conversations were included in the analysis of global effects, the analysis of local effects is based on a subset of 38 conversations. This subset consists of all conversations of four female speakers and about four conversations of each of the ten male speakers. F0 features were extracted from the f0 track for the variables shown in table 1 of the respective speech parts from both speakers and all measurements were calculated in semitones.

**Table 1:** Acoustic measurements of f0 features.

Feature	Description
mean	overall f0 mean
median	overall f0 median
max	95 <sup>th</sup> percentile (to exclude outliers)
range1	difference between upper quartile and lower quartile
range2	difference between f0 max and f0 mean

### 2.4 Statistical analysis

We conducted linear mixed effects models using R [10], the lme4-package [11] as well as the lmerTest-package [12]. Model fit was determined by maximum likelihood ratio tests. *P*-values were calculated using the Satterthwaite approximation. For primary effects we used SPEAKER SEX (SEX), perceived attractiveness (ATTRACTIVENESS\_POST (ATTR)) and likability (LIKABILITY\_POST (LIKE)), using the ratings obtained immediately after the respective conversation, as well as all interactions. For the effects of synchrony, we excluded SPEAKER SEX due to model complexity and included the respective interlocutor's pitch features (INTERLOCUTOR PITCH (PITCH)) as a continuous fixed effect. We included random intercepts for *speaker* as random

effects. The dependent variables were the f0 features listed in table 1.

## 3. Results

### 3.1 Primary global effects

The differences in natural register cause significant speaker sex effects for the register features. Perceived attractiveness does not affect the speakers' register but is positively correlated with *range1* ( $b=0.05$ ,  $SE=0.02$ ,  $df=183.54$ ,  $t=2.59$ ,  $p<.05$ ). Perceived likability however is positively correlated with *mean* ( $b=0.06$ ,  $SE=0.03$ ,  $df=177.67$ ,  $t=1.99$ ,  $p<.05$ ) but does not show any significant effects for pitch range (s. table 2).

**Table 2:** Primary effects for global pitch features.

	Mean	Median	Max	Range1	Range2
SEX	<.001	<.001	<.001	n.s.	n.s.
ATTR	n.s.	n.s.	n.s.	<.05	n.s.
LIKE	<.05	n.s.	n.s.	n.s.	n.s.

### 3.2 Global pitch synchrony

We find global synchrony for all register features including *mean* ( $b=0.10$ ,  $SE=0.02$ ,  $df=178.70$ ,  $t=4.16$ ,  $p<.001$ ), *median* ( $b=0.05$ ,  $SE=0.02$ ,  $df=177.87$ ,  $t=2.81$ ,  $p<.01$ ), and *max* ( $b=0.08$ ,  $SE=0.02$ ,  $df=178.60$ ,  $t=3.42$ ,  $p<.001$ ), and *range1* ( $b=0.13$ ,  $SE=0.05$ ,  $df=181.30$ ,  $t=2.37$ ,  $p<.001$ ). All four features are positively correlated. Global synchrony is significantly affected by perceived attractiveness for *range1* ( $b=1.16e-02$ ,  $SE=5.40e-03$ ,  $df=183.70$ ,  $t=2.15$ ,  $p<.05$ ) and by perceived likability for *mean* ( $b=6.94e-04$ ,  $SE=3.10e-04$ ,  $df=175.90$ ,  $t=2.25$ ,  $p<.05$ ). Both interactions increase the slope of global synchrony (s. table 3).

**Table 3:** Effects for global synchrony.

	Mean	Median	Max	Range1	Range2
PITCH	<.001	<.01	<.001	<.001	n.s.
PITCH X ATTR	n.s.	n.s.	n.s.	<.05	n.s.
PITCH X LIKE	<.05	n.s.	n.s.	n.s.	n.s.

### 3.3 Primary local effects

**Table 4:** Primary effects for local pitch features.

	Mean	Median	Max	Range1	Range2
SEX	<.001	<.001	<.001	n.s.	<.01
ATTR	<.001	<.001	<.001	<.05	n.s.
LIKE	<.001	<.001	<.001	<.01	n.s.
ATTR X LIKE	<.001	<.001	<.001	n.s.	n.s.
SEX X LIKE	<.001	<.001	<.001	<.01	<.01
SEX X ATTR	n.s.	n.s.	n.s.	<.001	n.s.

Again, the results show natural differences between the two sexes for the register features but also for *range2*. All three register features are affected by perceived attractiveness and likability as well as their interaction. Perceived attractiveness is positively correlated with *mean* ( $b=0.21$ ,  $SE=0.03$ ,  $df=7911.00$ ,  $t=6.74$ ,  $p<.001$ ), *median* ( $b=0.19$ ,  $SE=0.03$ ,  $df=7910.00$ ,  $t=6.14$ ,  $p<.001$ ), and *max* ( $b=0.26$ ,  $SE=0.04$ ,  $df=7912.00$ ,  $t=7.11$ ,  $p<.001$ ). Perceived likability is also

positively correlated with *mean* ( $b=0.13$ ,  $SE=0.02$ ,  $df=7909.00$ ,  $t=7.64$ ,  $p<.001$ ), *median* ( $b=0.11$ ,  $SE=0.02$ ,  $df=7908.00$ ,  $t=6.99$ ,  $p<.001$ ), and *max* ( $b=0.14$ ,  $SE=0.02$ ,  $df=7910.00$ ,  $t=7.28$ ,  $p<.001$ ). The interaction between perceived attractiveness and likability is significantly negative for *mean* ( $b=-0.03$ ,  $SE=4.15e-03$ ,  $df=7911.00$ ,  $t=-7.49$ ,  $p<.001$ ), *median* ( $b=-0.03$ ,  $SE=4.09e-03$ ,  $df=7910.00$ ,  $t=-6.70$ ,  $p<.001$ ), and *max* ( $b=-0.04$ ,  $SE=4.82e-03$ ,  $df=7912.00$ ,  $t=-7.79$ ,  $p<.001$ ). For *range1* there are also significant effects for perceived attractiveness ( $b=0.04$ ,  $SE=0.02$ ,  $df=5876.00$ ,  $t=2.43$ ,  $p<.05$ ) and likability ( $b=-0.04$ ,  $SE=0.01$ ,  $df=7772.00$ ,  $t=-3.03$ ,  $p<.01$ ). However, while attractiveness is positively correlated, likability is negatively correlated. Lastly, speaker sex significantly interacts with perceived likability on all relevant parameters including *mean* ( $b=0.15$ ,  $SE=0.02$ ,  $df=7913.00$ ,  $t=7.45$ ,  $p<.001$ ), *median* ( $b=0.13$ ,  $SE=0.02$ ,  $df=7912.00$ ,  $t=6.52$ ,  $p<.001$ ), *max* ( $b=0.18$ ,  $SE=0.02$ ,  $df=7913.00$ ,  $t=7.72$ ,  $p<.001$ ), and *range1* ( $b=0.06$ ,  $SE=0.02$ ,  $df=4194.00$ ,  $t=2.84$ ,  $p<.01$ ), and *range2* ( $b=0.11$ ,  $SE=0.03$ ,  $df=4556.00$ ,  $t=3.17$ ,  $p<.01$ ). Accordingly, we conducted post-hoc linear mixed effects models on the subsets for the two speaker sexes to investigate the nature of these interactions (s. table 4).

The analysis of the speaker sex subsets generally reinforces the effects found for the whole sample. However, a major difference can be seen in the direction of the main effects. While perceived attractiveness is positively correlated with *mean* ( $b=0.24$ ,  $SE=0.03$ ,  $df=3980.00$ ,  $t=7.22$ ,  $p<.001$ ) and *range1* ( $b=0.04$ ,  $SE=0.02$ ,  $df=1247.00$ ,  $t=2.20$ ,  $p<.05$ ) for the female speakers, it is negatively correlated with *mean* ( $b=-0.04$ ,  $SE=0.01$ ,  $df=3901.00$ ,  $t=-2.71$ ,  $p<.01$ ) and *range1* ( $b=-0.03$ ,  $SE=0.01$ ,  $df=3563.00$ ,  $t=-2.10$ ,  $p<.05$ ) for the male speakers. Furthermore, the negative correlation between *range1* and perceived likability is only significant for the female speakers ( $b=-0.04$ ,  $SE=0.01$ ,  $df=3430.00$ ,  $t=-2.87$ ,  $p<.01$ ) and not for the male speakers. Lastly, attractiveness and likability interact for the mean in female speakers ( $b=-0.03$ ,  $SE=4.69e-03$ ,  $df=3998.00$ ,  $t=-7.45$ ,  $p<.001$ ) but not in the male speakers.

### 3.4 Local pitch synchrony

In contrast to global pitch synchrony, there are no significant main effects for local synchrony for any of the investigated pitch features. Perceived attractiveness does not show significant effects on local synchrony for *mean* or *range1* but reaches significance for *median* ( $b=5.40e-03$ ,  $SE=2.58e-03$ ,  $df=7904.00$ ,  $t=2.09$ ,  $p<.05$ ), *max* ( $b=6.44e-03$ ,  $SE=2.86e-03$ ,  $df=7906.00$ ,  $t=2.26$ ,  $p<.05$ ), and *range2* ( $b=3.38e-02$ ,  $SE=1.69e-02$ ,  $df=7718.00$ ,  $t=2.01$ ,  $p<.05$ ), increasing the degree of entrainment in all three cases. Perceived likability, however, significantly increases local synchrony for all three register features, (*mean*:  $b=8.10e-03$ ,  $SE=1.65e-03$ ,  $df=7901.00$ ,  $t=4.93$ ,  $p<.001$ ; *median*:  $b=7.94e-03$ ,  $SE=1.66e-03$ ,  $df=7901.00$ ,  $t=4.78$ ,  $p<.001$ ; *max*:  $b=8.37e-3$ ,  $SE=1.83e-03$ ,  $df=7902.00$ ,  $t=4.57$ ,  $p<.001$ ) but not for the range features (s. table 7).

Table 7: Effects for local synchrony.

	Mean	Median	Max	Range1	Range2
PITCH	n.s.	n.s.	n.s.	n.s.	n.s.
PITCH X ATTR	n.s.	<.05	<.05	n.s.	<.05
PITCH X LIKE	<.001	<.001	<.001	n.s.	n.s.

## 4. Discussion

The results suggest that perceived attractiveness and likability systematically affect a speaker's pitch features as well as the presence and the degree of pitch synchrony on a global and a local level. Furthermore, the two impressions and the two levels can be distinguished to a certain degree regarding the pitch features involved.

On the global level, the effects of perceived attractiveness and likability were smaller and less diverse than on the local level. However, this results in a clearer distinguishability between the two impressions. Speakers were found to increase their interquartile pitch range with an increase of perceived attractiveness. An increase in perceived likability, however, led to a raise in register regarding the overall mean. The effects were consistent for both speaker sexes. Accordingly, the found effects do not support the previous findings [cf. 1]. Furthermore, perceived likability did not systematically affect the pitch range [cf. 1, 2].

Significant main effects of synchronous entrainment were found on the global level for all pitch features excluding the range from maximum to mean, which supports previous assumptions about prosodic entrainment [5, 6, 7]. Furthermore, the degree of entrainment was significantly affected by both impressions in the direction expected from the primary effects. Perceived attractiveness increased the degree of pitch synchrony for the interquartile range, while perceived likability increased the degree of pitch synchrony for the overall mean. Thus, global synchrony not only depends on the interactive nature of the conversational situation but also on individual perceptions of social variables.

For the primary effects on the local level, perceived attractiveness and likability both led to an increase in all of the register features and significantly interacted with each other. The interquartile range also showed significant effects of both impressions but no interaction. An investigation of the interactions for sex draws a slightly different picture. Perceived likability showed consistent effects across speaker sexes but perceived attractiveness differed. While female speakers raised their register and expanded their interquartile range with an increase of perceived attractiveness, the male speakers narrowed their interquartile range and lowered their register. These effects are compatible with the speaker sex differences expected from the research literature [cf. 1]. However, those effects were not found on the global level.

Another peculiarity arises for local pitch synchrony. While speakers adjust to their respective interlocutor's pitch features globally, they did not do so dynamically at turn breaks during a conversation as a main effect. However, local synchrony did occur in interaction with both perceived attractiveness and likability. The effects for attractiveness are less clear since the two features most frequently affected

in this study, the mean and the interquartile range, showed no effects, while the median, the maximum, and the range from maximum to mean did. However, perceived likability systematically caused local synchrony for all three register features. Accordingly, the degree to which speakers adjust their register to their interlocutor at turn breaks is dependent on how likable they perceive their interlocutor. This is also consistent with the effects found for global synchrony [cf. 1].

In summary, while both the primary effects and the effects on entrainment are fairly consistent globally and locally for perceived likability, the effects of attractiveness differ across the two levels and are much less systematic for entrainment. How could this finding be explained?

First of all, as discussed in [1], the effects of perceived attractiveness on the global interquartile range might not be an immediate effect of perceived attractiveness. A previous study found that a higher degree of perceived attractiveness correlates with a higher pitch range in the interlocutor. Accordingly, this effect can be due to a general pitch range entrainment, which coincides with perceived attractiveness. In this case, attractiveness would not show any global effects.

But why would the effects of attractiveness be restricted to the local level? As has been shown, the effects of perceived likability are not only consistent for the level, for primary effects, and for entrainment, they are also consistent across speaker sexes. In accordance with the research literature, mutual liking can be regarded as a prerequisite for social harmony. Accordingly, the phonetic strategies connected to perceived likability, such as prosodic entrainment, may be directed at achieving a collaboration with the interlocutor [6, 8]. Signaling attractiveness, however, may be a more one-directional process. This is supported by the different strategies found for the two sexes regarding the primary local effects. With the immediate effects of perceived attractiveness basically pointing in the opposite direction, the absence of entrainment is a logical consequence.

However, this does not explain why we do not find the same effects for the global level. One possible explanation is that it is either too straining or simply uneconomic to sustain an altered register or pitch range for the duration of a whole conversation. Accordingly, the turn break may not only be an essential part of conversational organization [cf. 13] but also for prosodic signalling. Thus, while the nuclear part of an intonational phrase – and thus also the last part of the last interpausal unit at turn breaks – is critical for signalling linguistic functions [cf. 14], the initial interpausal unit of the first turn of a speaker after a turn break may function as a salient point to signal paralinguistic prosodic features.

Lastly, the discrepancies between local and global effects could also be based on the differences in the amount of data. While the local effects were calculated for about 150 turns per speaker per conversation, each conversation only contributes a single value per pitch feature on the global level. Accordingly, the global effects might have disappeared statistically in the large amount of noise present in natural conversations.

## References

- [1] Michalsky, J., and Schoormann, H.: Effects of perceived attractiveness and likability on global aspects of fundamental frequency. *Proceedings of P&P12* (2016), 120–124.
- [2] Weiss, B.: Akustische Korrelate von Sympathieurteilen bei Hörern gleichen Geschlechts. 26th Konferenz Elektronische Sprachsignalverarbeitung (ESSV), Eichstätt. *Studientexte zur Sprachkommunikation*. TUD Press, Dresden 78 (2015), 165–171.
- [3] Hughes, S. M., Farley, S. D., and Rhodes, B. C.: Vocal and physiological changes in response to the physical attractiveness of conversational partners. *Journal of Nonverbal Behavior* 34 (2010), 1–13.
- [4] Fraccaro, P. J., Jones, B. C., Vukovic, J., Smith, F. G., Watkins, C. D., Feinberg, D. R., Little, A. C., and Debruine, L. M.: Experimental evidence that women speak in higher voice pitch to men they find attractive. *Journal of Evolutionary Psychology* 9 (2011), 57–67.
- [5] Levitan, R., Gravano, A., Willson, L., Beňuš, S., Hirschberg, J., and Nenkova, A.: Acoustic-prosodic entrainment and social behavior. *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (2012), 11–19.
- [6] Levitan, R. Acoustic-prosodic entrainment in human-human and human-computer dialogue, Columbia University. PhD thesis (2014).
- [7] Edlund, J., Heldner, M., and Hirschberg, J.: Pause and gap length in face-to-face interaction. *Proceedings of Interspeech* (2009).
- [8] Lubold, N., and Pon-Barry, H.: Acoustic-Prosodic Entrainment and Rapport in Collaborative Learning Dialogues. *Proceedings of the 2014 ACM workshop on Multimodal Learning Analytics Workshop and Grand Challenge*, November 12-12, 2014, Istanbul, Turkey (2014).
- [9] Boersma, P., and Weenink, D. Praat: Doing phonetics by computer (2016). <http://www.fon.hum.uva.nl/praat/>.
- [10] R Core Team. R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria (2017). <https://www.R-project.org/>.
- [11] Bates, D., Maechler, M., Bolker, B., and Walker, S.: Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67 (2015), 1–48.
- [12] Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. lmerTest: Tests in linear mixed effects models. R package version 2.0-30 (2016). <https://CRAN.R-project.org/package=lmerTest>.
- [13] Sacks, H., Schegloff, E. A., and Jefferson, G.: A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language* 50 (1974), 696–735.
- [14] Ladd, D. R. *Intonational phonology*. Cambridge University Press, Cambridge (2008).