8

the british psychological society
promoting excellence in psychology

British Journal of Psychology (2020), 111, 823–839

www.wileyonlinelibrary.com

Does he sound cooperative? Acoustic correlates of cooperativeness

Arnaud Tognetti[1,2]*, Valerie Durand[3], Melissa Barkat-Defradas[3,†] and Astrid Hopfensitz[4,†]

[1]Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden
[2]Institute for Advanced Study in Toulouse, France
[3]ISEM, CNRS, IRD, University of Montpellier, France
[4]Toulouse School of Economics, Université Toulouse 1 Capitole, France

The sound of the voice has several acoustic features that influence the perception of how cooperative the speaker is. It remains unknown, however, whether these acoustic features are associated with actual cooperative behaviour. This issue is crucial to disentangle whether inferences of traits from voices are based on stereotypes, or facilitate the detection of cooperative partners. The latter is likely due to the pleiotropic effect that testosterone has on both cooperative behaviours and acoustic features. In the present study, we quantified the cooperativeness of native French-speaking men in a one-shot public good game. We also measured mean fundamental frequency, pitch variations, roughness, and breathiness from spontaneous speech recordings of the same men and collected saliva samples to measure their testosterone levels. Our results showed that men with lower-pitched voices and greater pitch variations were more cooperative. However, testosterone did not influence cooperative behaviours or acoustic features. Our finding provides the first evidence of the acoustic correlates of cooperative behaviour. When considered in combination with the literature on the detection of cooperativeness from faces, the results imply that assessment of cooperative behaviour would be improved by simultaneous consideration of visual and auditory cues.

The manifestation of cooperation and trust across contexts has been the focus of much research in economics, biology, and psychology. Decisions about interpersonal interactions that require cooperation are influenced by relatedness (kin selection), spatial constraints (e.g., spatial selection, multi-level selection), and reputation (Hamilton, 1964; Nowak, 2006; Rand & Nowak, 2013; West, Griffin, & Gardner, 2007). It remains unknown, however, how cooperative partners can be identified in the absence of these contextual factors. One possibility is that people use phenotypic cues to assess the cooperative intent of potential partners.

Studies that examined the existence of such cues focus primarily on facial traits. Discriminating between high and low altruistic individuals, and estimating their trustworthiness and cooperation, is fast, spontaneous, and intuitive processes (Bonnefon,

*Correspondence should be addressed to Arnaud Tognetti, Division of Psychology, Department of Clinical Neuroscience, Karolinska Institutet, Nobels väg 9, 171 65 Solna, Stockholm, Sweden (email: arnaud.tognetti@gmail.com).

†These authors jointly supervised this work.

DOI:10.1111/bjop.12437

Hopfensitz, & De Neys, 2013, 2017; Dzhelyova, Perrett, & Jentzsch, 2012), which are based on both static and dynamic facial cues. For example, only a few seconds are needed to predict the trustworthiness or cooperation of a target individual from a picture of his/her face (Bonnefon *et al.*, 2013; De Neys, Hopfensitz, & Bonnefon, 2015) or a silent video clip (Fetchenhauer, Groothuis, & Pradel, 2010; Oda, Yamagata, Yabiku, & Matsumoto-Oda, 2009). Individuals' assessments and decisions regarding a cooperative other rely on facial cues, including width-to-height ratio (Little, Jones, DeBruine, & Dunbar, 2013; Stirrat & Perrett, 2010, 2012) and the presence of a genuine (i.e., Duchenne) smile (Brown, Palameta, & Moore, 2003; Centorrino, Djemai, Hopfensitz, Milinski, & Seabright, 2015; Johnston, Miles, & Macrae, 2010; Oda, Naganawa, Yamauchi, Yamagata, & Matsumoto-Oda, 2009; Reed, Zeglen, & Schmidt, 2012). In addition, facial cues of cooperativeness are reliable and valid across cultures (Tognetti, Berticat, Raymond, & Faurie, 2013; Tognetti, Yamagata-Nakashima, Faurie, & Oda, 2018).

Nevertheless, evolutionary ecology highlights the adaptive advantage of *multimodal* signalling in animal communication (Hebets & Papaj, 2005). For example, the *redundant signal hypothesis* suggests that multiple cues considered in combination provide a better estimate than any single cue (Groyecka *et al.*, 2017; Moller & Pomiankowski, 1993). Accordingly, cooperativeness might also be best identified by considering multiple cues.

One likely contributing factor is the voice. As with faces, vocal cues influence the perception of a speaker's social traits. For example, digital manipulation of speech recordings (raising or lowering vocal pitch) alters ratings related to trustworthiness, although this association seems to be context-dependent. Lower-pitched male voices are perceived as *less* trustworthy in economic trust games (Montano, Tigue, Isenstein, Barclay, & Feinberg, 2017; O'Connor & Barclay, 2017; Torre, White, & Goslin, 2016) and mating contexts (i.e., more likely to cheat; O'Connor & Barclay, 2017; O'Connor, Re, & Feinberg, 2011), but *more* trustworthy in general (Oleszkiewicz, Pisanski, Lachowicz-Tabaczek, & Sorokowska, 2017; Schirmer, Feng, Sen, & Penney, 2019; Tsantani, Belin, Paterson, & McAleer, 2016), or when trust is linked to the political context (Klofstad, Anderson, & Nowicki, 2015; Klofstad, Anderson, & Peters, 2012; Tigue, Borak, O'Connor, Schandl, & Feinberg, 2012). Vocal cues also influence ratings of cooperativeness. For example, male voices are perceived as *less* likely to 'contribute to a mutually beneficial goal' when they have a low pitch, but *more* likely to cooperate when they exhibit wide pitch variations in combination with a low pitch (Knowles & Little, 2016).

The voice's influence on the perception of a speaker's social traits could stem from the pleiotropic effect of testosterone on both vocal pitch and cooperative or trust-related behaviours (O'Connor & Barclay, 2017). Indeed, the surge of testosterone during puberty causes the vocal folds to grow longer and thicker through androgen receptors in the epithelial cells of the vocal folds, thus lowering vocal pitch (larger vocal folds mechanically vibrate more slowly than smaller ones; Harries, Hawkins, Hacking, & Hughes, 1998). Testosterone continues to influence vocal pitch through adolescence (Hodges-Simeon, Gurven, & Gaulin, 2015; Pedersen, Møller, Krabbe, & Bennett, 1986) and adulthood (Dabbs & Mallinger, 1999; Evans, Neave, Wakelin, & Hamilton, 2008; Puts, Apicella, & Cárdenas, 2012). Testosterone is also associated with cooperative or trust-related behaviours, but its influence is context-dependent. For example, men with higher levels of testosterone report lower cooperativeness on the self-report altruism scale (Harris, Rushton, Hampson, & Jackson, 1996), a greater likelihood of exploiting their partner's trust in a trust game (Takagishi, Takahashi, & Yamagishi, 2011), and more punitive acts towards other players in ultimatum games (Burnham, 2007). Men with artificially raised testosterone levels are also less generous in ultimatum games (Zak *et al.*, 2009, but see Cueva *et al.*, 2017). Nevertheless, during intergroup competition, men who have high testosterone levels are more cooperative and generous with in-group members (Diekhof, Wittmer, & Reimers, 2014; Reimers & Diekhof, 2015).

Findings that link testosterone with vocal pitch and cooperative behaviour motivated previous research to examine how fundamental frequency (the acoustic correlate of vocal pitch) influences the perception of a speaker's social traits. Such perceptions, could, however, be influenced by other acoustic features such as pitch variations, jitter, and harmonics-to-noise ratio (HNR), which are proxies for intonation, vocal roughness, and breathiness, respectively (Hillenbrand & Houde, 1996; Rabinov, Kreiman, Gerratt, & Bielamowicz, 1995). Indeed, all of these acoustic criteria are sexually dimorphic and thus potentially influenced by sexual hormones such as testosterone. Relevant evidence indicates that, compared with women, men have less variance in fundamental frequency (more monotonous; Puts *et al.*, 2012), but higher and lower values of jitter and HNR, respectively (Graddol, 1989; Van Borsel, Janssens, & De Bodt, 2009). One study showed a weak association between testosterone and pitch variations (Puts *et al.*, 2012), whereas some authors suggest that jitter and HNR are sensitive to hormonal influx because both relate to oscillations of the vocal folds, which are influenced by circulating androgens (Pisanski *et al.*, 2016). Moreover, pitch variations, jitter, and HNR influence listeners' perceptions of cooperativeness or trustworthiness (Belin, Boehme, & McAleer, 2017; Knowles & Little, 2016; McAleer, Todorov, & Belin, 2014; Ponsot, Burred, Belin, & Aucouturier, 2018; Schirmer *et al.*, 2019; Weirich, 2008). State-of-the-art vocal manipulations of pitch highlight the major influence that pitch variations have on perceptions of a speaker's social traits (Belin *et al.*, 2017; Ponsot *et al.*, 2018), such that voices with greater variations in pitch (i.e., more dynamic, less monotonic) are rated as more cooperative than voices with less variations (Knowles & Little, 2016). One interesting finding indicated that women perceive low-pitched male voices as the *least* cooperative when they have small pitch variations, but as the *most* cooperative when they have greater variations (Knowles & Little, 2016). It remains unknown, however, whether acoustic features co-vary in the way they influence the perception of a speaker's social traits. Most importantly, it is unknown whether acoustic features of the voice are associated with *actual* cooperative behaviours, and not only with *perceived* cooperativeness. Thus, it is crucial to disentangle whether inferences of traits from voices are based on dubious stereotypes, or facilitate the detection of cooperative partners.

In the present study, we investigated the existence of acoustic correlates of cooperativeness, by (1) relating cooperative behaviour in an economic social-dilemma game to acoustic parameters measured from recordings of spontaneous speech and (2) examining the potential role of testosterone in this association. Specifically, we quantified the cooperativeness of native French-speaking men in a standard public good game. This game represents a stylized model of a community in which each individuals' well-being depends on their own and others' contributions (Ledyard, 1995). Individually, each individual is best off if s/he contributes nothing and relies on others' efforts to create social benefits by behaving cooperatively. The external validity of the public good game was demonstrated previously by linking individuals' contributions to the public good with cooperative behaviours in naturally occurring situations (Fehr & Leibbrandt, 2011; Rustagi, Engel, & Kosfeld, 2010). We also measured mean fundamental frequency, pitch variations, roughness, and breathiness from spontaneous speech recordings of the same men, and we collected saliva samples to measure their testosterone levels. In line with the previous literature, we hypothesized that men's contributions to the public good would be associated with vocal-acoustic characteristics due to the influence of testosterone on both cooperative behaviours and acoustic features.

Methods

The protocol concerning recruitment procedures and data collection was approved by the ethical committee of Toulouse School of Economics - Institute for Advanced Study in Toulouse (#2016-10-001). We obtained written informed consent from all participants.

Participants

Origin (Ordin & Mennen, 2017; Zimmerer, Jügler, Andreeva, Möbius, & Trouvain, 2014), language (Andreeva *et al.*, 2014; Pépiot, 2014), sexual orientation (Baeck, Corthals, & Borsel, 2011; Gaudio, 1994; Munson, McDonald, DeBoe, & White, 2006), and smoking (Sorensen & Horii, 1982) influence acoustic characteristics. Hence, in the present study we only used the voice recordings of participants that were native French speakers, born in metropolitan France, had French parents and European grandparents, and who neither smoked, nor reported to be homosexual. We expected that approximately 15% of show-ups would not fulfil these prerequisites. As a result, a total of 81 men were recruited as participants in our study, leading to an estimated number of approximately 70 useable observations for speech analysis. Numbers were based on estimated sample size requirements ($N = 62$) for a linear model with 4 explanatory variables with 80% power, $\alpha = 0.05$, and $R^2 = 0.15$ (*pwr.f2.test* function of the *pwr* package in R). Male participants were randomly selected from a pool of more than 4,000 volunteers from the University of Montpellier. None of them had previously participated in a public good game experiment.

From November 2016 to February 2017, 12 experimental sessions took place at the Laboratory of Experimental Economics in Montpellier (University of Montpellier, France). All sessions started at 2 p.m. Each session started in a computer laboratory where participants first played a public good game. Upon completion, they were invited to a separate room for speech recordings.

Measure of cooperativeness

To quantify cooperation, we used the public good game (Ledyard, 1995). This game, which is considered the benchmark for experimental research on social dilemmas, represents a stylized model of a community in which each individuals' well-being depends on own and others' contributions. Individually, each member is best off if s/he contributes nothing and relies on others' efforts to create social benefits by behaving cooperatively.

The public good game was run on a computer network (interface programmed with LE2M – Software for the Economic Experiments of Montpellier developed by D. Dubois and JM. Rousselle). To prevent visual contact, each participant was seated in an individual cubicle containing a computer terminal. Communication between participants was not allowed. At the beginning of each session, participants received a written copy of the instructions (for details see the Supporting Information). To implement common knowledge of the game and the task, the principal investigator (AT) also read the instructions aloud. Questions were allowed and were answered privately. We checked participants' understanding of the instructions by a computerized questionnaire. To guarantee experimenter–subject and subject–subject anonymity, a subject number was assigned to each participant.

In each session, participants played, in groups of two, a one-shot linear public good game followed by a conditional contribution in the same game (we followed procedures of Fischbacher *et al.* (2001); for details see the SM). The one-shot public good game is classically used to measure cooperativeness and has the advantage to provide a continuous measure of cooperativeness but does not control for the participants' beliefs. Conditional contributions do control for individuals' beliefs and allow categorization of individuals according to type (as defined by Fischbacher *et al.* (2001); see also the SM). However, only two categories are generally highly represented: *conditional cooperators* (individuals who are willing to contribute more to a public good the more others contribute) and *free-riders* (individuals who do not contribute to a public good regardless of others' contributions).

At the beginning of the one-shot public good game, participants were randomly assigned into pairs. Each participant received an initial endowment of 20 tokens. Then, each player independently decided how to allocate his endowment between a private and a common account. Allocation decisions yielded payoffs in euros. Each token allocated by a subject to his private account paid off 1.50€ for himself while the common account paid off 0.90€ to each member of the pair (marginal per capita return, MPCR = 0.6). It was made clear that each token allocated to the common account would provide exactly the same payoff to each member of the pair regardless of the contributor. From these parameters, it follows that the utilitarian optimum and the efficient symmetric outcome is for all group members to contribute their entire endowments to the public account. However, even under these specifications, it still remains in each individual's self-interest to contribute zero.

After the one-shot public good game, participants were asked, in an incentive compatible way, how much they would contribute if they could condition their contributions on their partner's one (conditional choice).

All participants were informed about their final payoff at the end of the entire experimental session to avoid any potential influence on speech recordings. Average earnings were 32 $\pm$ 5.60 € (Mean $\pm$ *SD*), and each subject was paid in private.

Speech recordings and analyses

Recordings

At the end of the public good game, each participant was invited to enter individually another room for the speech recordings. A single investigator conducted all the recordings (MBD) which took place in a quiet room of the laboratory of Experimental Economics in Montpellier. We recorded spontaneous speech due to its stronger ecological validity compared with scripted speech (Puts, Hodges, Cárdenas, & Gaulin, 2007; Suire, Raymond, & Barkat-Defradas, 2018). Thus, participants were asked to describe a picture regarding global warming (Figure S1). Once they finished describing the picture, spontaneous speech was elicited by asking: 'Do you think that ecological actions by individual citizens are sufficient to reduce the impact of global change or even to halt it?' (our translation from the original French question).

We used the complete answer produced by participants as material for acoustic analysis. Answers to semi-directive open questions limit content variation while inducing participants to use a more natural voice, as their attention is concentrated on the expression of their own personal opinions.

We recorded speech samples using a linear PCM recorder with a sampling rate of 22 kHz, 16bits, mono, which were saved into.wav files. To control for intensity, participants were instructed to speak within a constant distance of 15 cm from the recorder.

Speech analyses

Participants who were native French speakers, born in metropolitan France, who had native French parents and European grandparents, and who neither smoked, nor reported to be homosexual were eligible for speech analysis. Our final sample consisted of 64 men (Mean age ± *SD* = 21.6 ± 3.3 years), a majority of which were students (55 out of 64 participants).

The 64 speech samples were analysed using the speech analysis software Praat© version 5.1.45. (Paul Boersma and David Weenink, Phonetic Sciences, University of Amsterdam, www.praat.org). Pitch floors were set to 70 Hz with a ceiling of 250 Hz. Other settings were kept as default. For each participant, we extracted values for four acoustic parameters via the voice report menu: mean fundamental frequency (mean F0, the acoustic correlate of vocal height, in Hz), variations of fundamental frequency (proxy of intonation, F0-SD in Hz), jitter (proxy of vocal roughness, in %), and the HNR (proxy of vocal breathiness, in dB).

Saliva collection and testosterone assays

Testosterone concentrations (pg/mL) were measured in saliva samples. This non-invasive technique yields testosterone levels that are highly correlated with free testosterone levels (Ellison, 1988; Vittek, L'Hommedieu, Gordon, Rappaport, & Louis Southren, 1985). At the start of the experiment, each participant was given one labelled tube and straw (Salicaps kits; IBL-Hamburg, Hamburg, G) to collect saliva. Contamination of saliva samples was minimized by instructing participants not to eat, drink (except plain water), smoke, chew gum, or brush their teeth for one hour before the session. Because the change of testosterone rate in the diurnal cycle is the lowest in early afternoon (Dabbs, 1990b), saliva samples were collected at 2.00 p.m. Samples were kept refrigerated during the experiment and then stored at −80°C. Testosterone levels were analysed by luminescence immunoassay (LIA) technique, using LIA Testosterone kits (IBL, Hamburg). The assay of each sample was replicated twice, and only measures whose intra-assay CV was lower than 10% were used.

Additional data

Sociodemographic questionnaires were also administered (see SM) at the end of the session (after public good game and speech recording sessions). We obtained self-reports on nationality, mother tongue, country of birth of parents and grandparents, sexual orientation, and smoking habits, to use this information as exclusion criteria during data analysis (see above). Moreover, since age, level of education and body mass index (BMI) affect cooperative behaviours, acoustic characteristics, or testosterone levels (e.g., Dabbs, 1990a, 1990b, 1992; Ellison *et al.*, 2002; Gachter & Herrmann, 2009; Jasienska, Jasienski, & Ellison, 2012; Lukas, Campbell, & Ellison, 2004; Schötz & Müller, 2007; Sutter & Kocher, 2007; Torre & Barlow, 2009), we also collected self-reports on age and education and measured and weighted each participant, to allow statistical control for these potentially confounding variables.

Statistical analyses

To examine whether acoustic characteristics were associated with unconditional contributions to the public good, we used a censored regression model (Tobit model, *censReg* function of the *censReg* package in R). This type of model enables us to control for the fact that contributions had an upper and a lower bound, since players could not contribute less than 0 or more than 20 tokens to the public good. Our explanatory variables were the four acoustic parameters discussed above (mean F0, F0-SD, jitter, HNR). We also added testosterone level to examine its potential influence on contributions to the public good.

We, then, examined whether conditional cooperators ($N = 44$) and free-riders ($N = 11$), the two main categories measured by the conditional contributions (Figure S2), differed in terms of acoustic parameters, by using a logistic regression (*glm* function with a binomial error structure of the *stats* package in R). We used the four acoustic parameters and testosterone level as explanatory variables.

Because co-occurring acoustic parameters are likely to be intercorrelated (e.g., Leongómez *et al.*, 2014; Werth, Voigt, Döllinger, Eysholdt, & Lohscheller, 2010), we checked for multicollinearity by calculating the variance inflation factor (VIF) for all explanatory variables (*vif* function of the *car* package in R). We ruled out potential bias from multicollinearity as all variables demonstrated a low value of VIF for both models (VIF < 2.46).

Finally, in order to examine the influence of testosterone level on men's speech, we performed four linear models, one for each acoustic parameter studied (i.e., mean F0, F0-SD, jitter, and HNR). Each acoustic parameter was used as a response variable, and we included testosterone level as an explanatory variable.

We also controlled for the robustness of our results by doing the same models but controlling for several confounding factors known to affect acoustic characteristics, testosterone levels, and cooperative behaviours such as age, BMI, level of education, and duration of the speech.

Statistical analyses were performed using R, version 3.4.2 (R Core Team, 2017). Following the recommendations found in Schielzeth (2010), we centred every continuous variable (including the explanatory variables) in all the models in order to make the effects more easily biologically interpretable.

Results

Average contributions to the public good, testosterone levels, and acoustic characteristics were similar to results from previous studies using similar subject pools (Suire *et al.*, 2018; Tognetti, Dubois, Faurie, & Willinger, 2016). Participants allocated slightly less than half of their endowment to the public good (Mean ± SD = 8.45 ± 6.96 tokens). The mean testosterone level was 122.91 ± 44.08 pg/mL. Mean fundamental frequency ranged between 82.45 and 144.78 Hz (Mean ± SD = 113.59 ± 13.86 Hz), pitch variations ranged between 7.14 and 31.30 Hz (14.37 ± 4.88 Hz), jitter ranged from 1.52% to 5.41% (2.68 ± 0.73 %), and HNR ranged from 9.10 to 15.52 dB (12.06 ± 1.68 dB). Also in line with the existing literature (e.g., Leongómez *et al.*, 2014; Werth *et al.*, 2010), we observed a significant positive correlation between fundamental frequency and pitch variations ($r = .54, N = 64, p < .0001$) and a significant negative correlation between jitter and HNR ($r = -.68, N = 64, p < .0001$). No other significant correlation was observed between the acoustic characteristics ($-.17 > r > .16$, $N = 64$, $.18 > p > .55$ for all correlations).

The censored regression model examining the potential association between acoustic characteristics and cooperative behaviours showed that both mean F0 and F0-SD were significantly associated with men's contributions to the public good (see Table S1.A). Lower-pitched voice men ($\beta = -0.26$, $SE = 0.12$, 95% CI = [−0.49, −0.03], χ^2 (1, $N = 64$) = 5.19, $p = .02$, Figure 1a) and men displaying more pitch variations ($\beta = 0.77$, $SE = 0.32$, 95% CI = [0.14, 1.40], χ^2 (1, $N = 64$) = 5.74, $p = .02$, Figure 1b) were more cooperative in a social-dilemma situation. Jitter ($\beta = -3.42$, $SE = 2.41$, 95% CI = [−8.15, 1.31], χ^2 (1, $N = 64$) = 2.00, $p = .16$), HNR ($\beta = 0.02$, $SE = 1.09$, 95% CI = [−2.11, 2.15], χ^2 (1, $N = 64$) < 0.01, $p = .98$), or testosterone level ($\beta = -0.01$, $SE = 0.03$, 95% CI = [−0.06, 0.05], χ^2 (1, $N = 64$) = 0.02, $p = .88$) were not significantly associated with men's contributions to the public good (see Table S1.A).

Consistently, the logistic regression showed that conditional cooperators displayed significantly more pitch variations ($\beta = 0.24$, $SE = 0.13$, 95% CI = [0.02, 0.54], χ^2 (1, $N = 55$)= 4.70, $p = .03$) than free-riders (Figure 2) but no difference concerning vocal pitch ($\beta = -0.03$, $SE = 0.03$, 95% CI = [−0.10, 0.03], χ^2 (1, $N = 55$) = 0.94, $p = .33$), jitter ($\beta = 0.44$, $SE = 0.64$, 95% CI = [−0.79, 1.79], χ^2 (1, $N = 55$) = 0.48, $p = .49$), HNR ($\beta = 0.55$, $SE = 0.33$, 95% CI = [−0.05, 1.29], χ^2 (1, $N = 55$) = 3.23, $p = .07$), or testosterone level ($\beta = 0.01$, $SE = 0.01$, 95% CI = [−0.007, 0.03], χ^2 (1, $N = 55$) = 1.36, $p = .24$) was found (see Table S2.A).

Finally, testosterone level did not significantly influence men's mean F0 ($\beta < -0.01$, $SE = 0.04$, $F(1,62) = 0.04$, $p = .84$), F0-SD ($\beta < 0.01$, $SE = 0.$ 01, $F(1,62) = 0.23$,

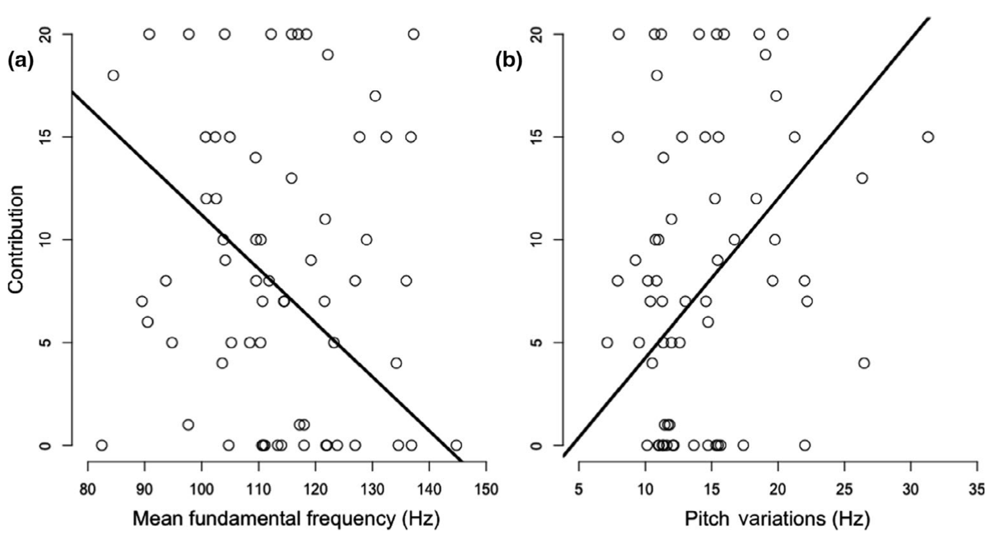

Figure 1. Contribution to the public good as a function of (a) the mean fundamental frequency (the acoustic correlate of vocal pitch) and (b) the pitch variations (acoustic correlate of intonation). Both the raw data ($N = 64$) and the predicted values from the censored regression model are shown. A censored regression model indicated that men's contributions to the public good are significantly associated with both mean F0 ($\beta = -0.26$, $SE = 0.12$, $p = .02$) and F0-SD ($\beta = 0.77$, $SE = 0.32$, $p = .02$).

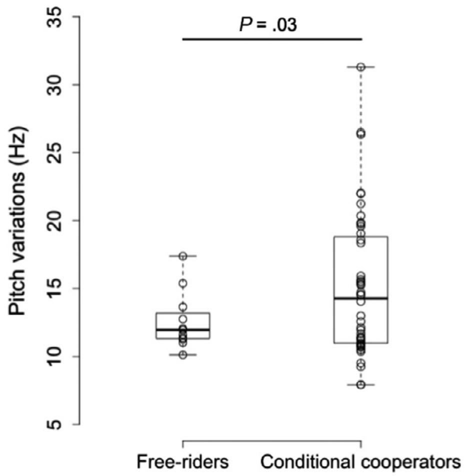

Figure 2. Box plot representation of the pitch variations exhibited by the two main categories measured by the conditional contributions: free-riders (N = 11) and conditional cooperators (N = 44). Medians (thick lines), first and third quartiles, and whiskers representing the entire data range, and individual data points are indicated. A logistic regression indicated that pitch variations significantly differ between free-riders and conditional cooperators (p = .03).

p = .63), jitter ($\beta < -0.01$, SE = 0. 01, $F(1,62) = 1.05$, p = .31), or HNR ($\beta < -0.01$, SE = 0.01, $F(1,62) = 0.24$, p = .63) (Table S3.A).

The results found were robust to the inclusion of various confounding variables such as age, BMI, level of education, and duration of speech. Specifically, qualitatively similar effects of both F0 and F0-SD were obtained (Tables S1.B S2.B), and none of the confounding variables significantly influenced the outcome variables (see panels B of Tables S1–S3).

Discussion

Several acoustic features influence the perception of how trustworthy and cooperative the speaker is (Belin *et al.*, 2017; Knowles & Little, 2016; Montano *et al.*, 2017; O'Connor & Barclay, 2017; Oleszkiewicz *et al.*, 2017; Ponsot *et al.*, 2018; Tigue *et al.*, 2012; Tsantani *et al.*, 2016). Their influence could stem from the pleiotropic effect of testosterone on both acoustic features and cooperative behaviours (O'Connor & Barclay, 2017). It is unknown, however, whether acoustic features are associated with actual cooperative behaviour and whether testosterone mediates this association. In this study, we present evidence that both vocal pitch and its variations are related to cooperative behaviour in an incentivized social-dilemma game: the public good game. However, no effect of testosterone level on cooperation, or on any of the other acoustic features studied, was found. Overall, our study provides the first evidence of the existence of acoustic correlates of cooperativeness.

Specifically, our results indicate that men's contributions to the public good are significantly and negatively associated with fundamental frequency and significantly and positively with its variations. When we compared the acoustic traits between conditional cooperators and free-riders (the two main categories as defined in Fischbacker *et al.* (2001)), we found that conditional cooperators exhibit significantly higher pitch variations than free-riders. Taken together, our results suggest that highly cooperative men have deeper voices and exhibit greater variations in their intonation compared to less cooperative men.

The present results are consistent with the only previous study examining jointly the influence of vocal pitch and its variations on the perception of a speaker's cooperativeness (Knowles & Little, 2016). Indeed, Knowles and Little (2016) found that male voices were perceived as the most likely to cooperate when they exhibited high pitch variations in combination with a low pitch (although it was found for women's but not for men's ratings). Vocal pitch and its variation are, thus, associated in the same way with both cooperative behaviours and perceived cooperativeness. It, therefore, indicates that inferences of cooperativeness from voices might actually facilitate the detection of cooperative partners. This sets the ground for future research, namely whether particular combinations of acoustic traits influence ratings of cooperativeness and to which degree these acoustic cues of cooperativeness are reliable or could be manipulated through conscious control of the speaker (e.g., by lowering voice pitch or by increasing speech's intonation).

Behavioural decisions in the public good games (contributions to the public good) and in the trust game (amounts sent to the other player) are highly correlated (Galizzi & Navarro-Martínez, 2018; Peysakhovich, Nowak, & Rand, 2014), which suggests a strong association between cooperativeness and trustworthiness. Hence, similarly to cooperativeness, vocal pitch and its variations are also likely to be used as cues of trustworthiness. In fact, both acoustic features influence perception of how trustworthy a speaker is. For example, lower-pitched male voices and voices with high pitch variations are perceived as more trustworthy in general (Belin *et al.*, 2017; Oleszkiewicz *et al.*, 2017; Schirmer *et al.*, 2019; Tsantani *et al.*, 2016) or when trust is linked to the political context (Klofstad *et al.*, 2015, 2012; Tigue *et al.*, 2012). It remains unknown, however, whether these acoustic features correlate with *actual* trustworthiness, and not only with *perceived* trustworthiness.

The existence of vocal cues of cooperativeness could stem from the pleiotropic effect of testosterone on both cooperative behaviours (Burnham, 2007; Diekhof *et al.*, 2014; Reimers & Diekhof, 2015; Takagishi *et al.*, 2011) and vocal pitch (Dabbs & Mallinger, 1999; Evans *et al.*, 2008; Puts *et al.*, 2012). As testosterone has immunosuppressive effects (*immunocompetence handicap hypothesis*: Folstad & Karter, 1992; Rantala *et al.*, 2012; but see: J. Nowak, Pawłowski, Borkowska, Augustyniak, & Drulis-Kawa, 2018), men with (costly) lower pitch might benefit from a higher biological quality (Arnocky, Hodges-Simeon, Ouellette, & Albert, 2018; Hodges-Simeon *et al.*, 2015). In addition, they may also be more socially dominant (Puts *et al.*, 2012; Puts, Gaulin, & Verdolini, 2006; Puts *et al.*, 2007). Accordingly, because of their underlying qualities, including access to resources, men with lower pitch would be more cooperative than men with higher pitch because they could better afford the costs associated with cooperative behaviours while receiving reputational benefits (Raihani & Smith, 2015; Sylwester & Roberts, 2010; Tognetti, Berticat, Raymond, & Faurie, 2012; Tognetti *et al.*, 2016). This condition-dependent mechanism could ensure the reliability of the vocal cues of cooperativeness. However, in the present study testosterone levels did not seem to affect cooperation or any of the acoustic features studied. Testosterone is a multiple-effect hormone which is influenced by numerous biological and environmental factors and pathways. As such, it is generally difficult to correlate testosterone levels to other biological or behavioural traits. In addition, we could only collect one sample of saliva for hormonal assays, which might not accurately reflect a participant's basal testosterone level.

Although the present study retains many strengths, it is also subject to several limitations. In particular, it is the first to investigate the existence of acoustic correlates of cooperativeness in speech production. However, the investigation is restricted to French men. To provide broader conclusions, it should not only be extended to women, but to other populations as well. In addition, we did not conduct a perceptual study using our recordings to examine whether listeners use acoustic features as a social cue in a behavioural economic task. Indeed, we recorded individuals' free speech due to its stronger ecological validity (Puts *et al.*, 2007; Suire *et al.*, 2018), but this type of recordings is not suitable for perceptual studies, as recordings roughly differ in duration and semantic content. Finally, we used state-of-the-art methodology in economics to quantify and categorize individuals according to type (Fischbacher *et al.*, 2001). However, we compared two categories with unbalanced sample sizes ($N_{\text{Free-rider}} = 11$ vs. $N_{\text{conditional cooperator}} = 44$) and the sample size of the free-riders was limited (although its proportion (18%) mimics the proportion found in the general French population; Frey, 2017). Hence, we cannot exclude the possibility that the acoustic differences found between free-riders and conditional cooperators arose from this specific and particular sample of 11 free-riders. It seems, nevertheless, unlikely as the results found using this categorization are qualitatively similar to the ones we found using a continuous measure of cooperativeness (i.e., contributions to the public good).

To conclude, the present study provides evidence that at least two acoustic features (vocal pitch and its local variations) could be used as cues of cooperativeness. Facial cues enable individuals to discriminate between high and low cooperative individuals with an above chance accuracy (Bonnefon *et al.*, 2013, 2017; Fetchenhauer *et al.*, 2010; Little *et al.*, 2013; Oda, Naganawa, *et al.*, 2009; Reed *et al.*, 2012; Stirrat & Perrett, 2010, 2012; Tognetti *et al.*, 2013) but the accuracy of face-based cooperation detection is rather low (Bonnefon *et al.*, 2017). Hence, by highlighting the fact that cooperativeness is advertised by several cues across multiple sensory modalities, our findings pave the way for further investigations examining whether the assessment of cooperative behaviour is improved by simultaneous consideration of both visual and auditory cues.

Acknowledgements

We are very grateful to D. Dubois for his help during experimental sessions and the Laboratory of Experimental Economics of Montpellier (LEEM, University of Montpellier) for hosting the experiment. We also thank A. Suire, A. Vergès, and J. Gourragne for their help in acoustic analyses and all of the participants in our study. Financial support by PEPS CNRS HomoVox, SINT-ANR-15-CE33-0005-01 and ANR Labex IAST is gratefully acknowledged. This is contribution ISEM 2019-254.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Andreeva, B., Demenko, G., Wolska, M., Möbius, B., Zimmerer, F., Jügler, J., . . . Trouvain, J. (2014). Comparison of pitch range and pitch variation in Slavic and Germanic languages. In N. Campbell,

D. Gibbon & D. Hirst (Eds.), *Proceedings to the 7th Speech Prosody Conference* (pp. 776–780). Dublin, Ireland: International Speech Communication Association.

Arnocky, S., Hodges-Simeon, C. R., Ouellette, D., & Albert, G. (2018). Do men with more masculine voices have better immunocompetence? *Evolution and Human Behavior*, *39*(6), 602–610. https://doi.org/10.1016/j.evolhumbehav.2018.06.003

Baeck, H., Corthals, P., & Borsel, J. V. (2011). Pitch characteristics of homosexual males. *Journal of Voice*, *25*(5), e211–e214. https://doi.org/10.1016/j.jvoice.2010.10.019

Belin, P., Boehme, B., & McAleer, P. (2017). The sound of trustworthiness: Acoustic-based modulation of perceived voice personality. *PLoS ONE*, *12*(10), e0185651. https://doi.org/10.1371/journal.pone.0185651

Bonnefon, J.-F., Hopfensitz, A., & De Neys, W. (2013). The modular nature of trustworthiness detection. *Journal of Experimental Psychology: General*, *142*(1), 143. https://doi.org/10.1037/a0028930

Bonnefon, J.-F., Hopfensitz, A., & De Neys, W. (2017). Can we detect cooperators by looking at their face? *Current Directions in Psychological Science*, *26*(3), 276–281. https://doi.org/10.1177/0963721417693352

Brown, W. M., Palameta, B., & Moore, C. (2003). Are there nonverbal cues to commitment? An exploratory study using the zero-acquaintance video presentation paradigm. *Evolutionary Psychology*, *1*, 42–69. https://doi.org/10.1177/147470490300100104

Burnham, T. C. (2007). High-testosterone men reject low ultimatum game offers. *Proceedings of the Royal Society of London B*, *274*(1623), 2327–2330. https://doi.org/10.1098/rspb.2007.0546

Centorrino, S., Djemai, E., Hopfensitz, A., Milinski, M., & Seabright, P. (2015). Honest signaling in trust interactions: Smiles rated as genuine induce trust and signal higher earning opportunities. *Evolution and Human Behavior*, *36*(1), 8–16. https://doi.org/10.1016/j.evolhumbehav.2014.08.001

Core Team, R. (2017). *R: A language and environnement for statistical computing*. Retrieved from https://www.R-project.org/

Cueva, C., Roberts, R. E., Spencer, T. J., Rani, N., Tempest, M., Tobler, P. N., . . . Rustichini, A. (2017). Testosterone administration does not affect men's rejections of low ultimatum game offers or aggressive mood. *Hormones and Behavior*, *87*, 1–7. https://doi.org/10.1016/j.yhbeh.2016.09.012

Dabbs, J. M. (1990a). Age and seasonal variation in serum testosterone concentration among men. *Chronobiology International*, *7*(3), 245–249. https://doi.org/10.3109/07420529009056982

Dabbs, J. M. (1990b). Salivary testosterone measurements: Reliability across hours, days, and weeks. *Physiology and Behavior*, *48*(1), 83–86. https://doi.org/10.1016/0031-9384(90)90265-6

Dabbs, J. M. (1992). Testosterone and occupational achievement. *Social Forces*, *70*(3), 813–824. https://doi.org/10.1093/sf/70.3.813

Dabbs, J. M., & Mallinger, A. (1999). High testosterone levels predict low voice pitch among men. *Personality and Individual Differences*, *27*(4), 801–804. https://doi.org/10.1016/S0191-8869(98)00272-4

De Neys, W., Hopfensitz, A., & Bonnefon, J.-F. (2015). Adolescents gradually improve at detecting trustworthiness from the facial features of unknown adults. *Journal of Economic Psychology*, *47*, 17–22. https://doi.org/10.1016/j.joep.2015.01.002

Diekhof, E. K., Wittmer, S., & Reimers, L. (2014). Does competition really bring out the worst? Testosterone, social distance and inter-male competition shape parochial altruism in human males. *PLoS ONE*, *9*(7), e98977. https://doi.org/10.1371/journal.pone.0098977

Dzhelyova, M., Perrett, D. I., & Jentzsch, I. (2012). Temporal dynamics of trustworthiness perception. *Brain Research*, *1435*, 81–90. https://doi.org/10.1016/j.brainres.2011.11.043

Ellison, P. T. (1988). Human salivary steroids—Methodological considerations and applications in physical-anthropology. *Yearbook of Physical Anthropology*, *31*, 115–142. https://doi.org/10.1002/ajpa.1330310507

Ellison, P. T., Bribiescas, R. G., Bentley, G. R., Campbell, B. C., Lipson, S. F., Panter-Brick, C., & Hill, K. (2002). Population variation in age-related decline in male salivary testosterone. *Human Reproduction*, *17*(12), 3251–3253. https://doi.org/10.1093/humrep/17.12.3251

Evans, S., Neave, N., Wakelin, D., & Hamilton, C. (2008). The relationship between testosterone and vocal frequencies in human males. *Physiology & Behavior*, *93*(4–5), 783–788. https://doi.org/10.1016/j.physbeh.2007.11.033

Fehr, E., & Leibbrandt, A. (2011). A field study on cooperativeness and impatience in the Tragedy of the Commons. *Journal of Public Economics*, *95*(9–10), 1144–1155. https://doi.org/10.1016/j.jpubeco.2011.05.013

Fetchenhauer, D., Groothuis, T., & Pradel, J. (2010). Not only states but traits—Humans can identify permanent altruistic dispositions in 20 s. *Evolution and Human Behavior*, *31*(2), 80–86. https://doi.org/10.1016/j.evolhumbehav.2009.06.009

Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters*, *71*(3), 397–404. https://doi.org/10.1016/S0165-1765(01)00394-9

Folstad, I., & Karter, A. J. (1992). Parasites, bright males, and the immunocompetence handicap. *The American Naturalist*, *139*(3), 603–622. https://doi.org/10.1086/285346

Frey, U. J. (2017). Cooperative strategies outside the laboratory—Evidence from a long-term large-N-study in five countries. *Evolution and Human Behavior*, *38*(1), 109–116. https://doi.org/10.1016/j.evolhumbehav.2016.07.006

Gachter, S., & Herrmann, B. (2009). Reciprocity, culture and human cooperation: Previous insights and a new cross-cultural experiment. *Philosophical Transactions of the Royal Society B-Biological Sciences*, *364*(1518), 791–806. https://doi.org/10.1098/rstb.2008.0275

Galizzi, M. M., & Navarro-Martínez, D. (2018). On the external validity of social preference games: A systematic lab-field study. *Management Science*, *65*(3), 976–1002. https://doi.org/10.1287/mnsc.2017.2908

Gaudio, R. P. (1994). Sounding gay: Pitch properties in the speech of gay and straight men. *American Speech*, *69*(1), 30. https://doi.org/10.2307/455948

Graddol, D. (1989). *Gender voices*. Hoboken, NJ: Blackwell Publishing.

Groyecka, A., Pisanski, K., Sorokowska, A., Havlíček, J., Karwowski, M., Puts, D., . . . Sorokowski, P. (2017). Attractiveness is multimodal: beauty is also in the nose and ear of the beholder. *Frontiers in Psychology*, *8*, 778. https://doi.org/10.3389/fpsyg.2017.00778

Hamilton, W. D. (1964). Genetical evolution of social behavior II. *Journal of Theoretical Biology*, *7*, 17–52. https://doi.org/10.1016/0022-5193(64)90039-6

Harries, M., Hawkins, S., Hacking, J., & Hughes, I. (1998). Changes in the male voice at puberty: Vocal fold length and its relationship to the fundamental frequency of the voice. *The Journal of Laryngology & Otology*, *112*(5), 451–454. https://doi.org/10.1017/S0022215100140757

Harris, J. A., Rushton, J. P., Hampson, E., & Jackson, D. N. (1996). Salivary testosterone and self-report aggressive and pro-social personality characteristics in men and women. *Aggressive Behavior*, *22(5)*, 321–331. https://doi.org/10.1002/(SICI)1098-2337(1996)22:5<321::AID-AB1>3.0.CO;2-M

Hebets, E. A., & Papaj, D. R. (2005). Complex signal function: Developing a framework of testable hypotheses. *Behavioral Ecology and Sociobiology*, *57*(3), 197–214. https://doi.org/10.1007/s00265-004-0865-7

Hillenbrand, J., & Houde, R. A. (1996). Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. *Journal of Speech, Language, and Hearing Research*, *39*(2), 311–321. https://doi.org/10.1044/jshr.3902.311

Hodges-Simeon, C. R., Gurven, M., & Gaulin, S. J. C. (2015). The low male voice is a costly signal of phenotypic quality among Bolivian adolescents. *Evolution and Human Behavior*, *36*(4), 294–302. https://doi.org/10.1016/j.evolhumbehav.2015.01.002

Jasienska, G., Jasienski, M., & Ellison, P. T. (2012). Testosterone levels correlate with the number of children in human males, but the direction of the relationship depends on paternal education.

Evolution and Human Behavior, *33*(6), 665–671. https://doi.org/10.1016/j.evolhumbehav.2012.05.001

Johnston, L., Miles, L., & Macrae, C. N. (2010). Why are you smiling at me? Social functions of enjoyment and non-enjoyment smiles. *British Journal of Social Psychology*, *49*(1), 107–127. https://doi.org/10.1348/014466609X412476

Klofstad, C. A., Anderson, R. C., & Nowicki, S. (2015). Perceptions of competence, strength, and age influence voters to select leaders with lower-pitched voices. *PLoS ONE*, *10*(8), e0133779. https://doi.org/10.1371/journal.pone.0133779

Klofstad, C. A., Anderson, R. C., & Peters, S. (2012). Sounds like a winner: Voice pitch influences perception of leadership capacity in both men and women. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1738), 2698–2704. https://doi.org/10.1098/rspb.2012.0311

Knowles, K. K., & Little, A. C. (2016). Vocal fundamental and formant frequencies affect perceptions of speaker cooperativeness. *The Quarterly Journal of Experimental Psychology*, *69*(9), 1657–1675. https://doi.org/10.1080/17470218.2015.1091484

Ledyard, J. O. (1995). Public goods: A survey of experimental research. In J. H. Kagel & A. E. Roth (Eds.), *The handbook of experimental economics* (pp. 111–181). Princeton, NJ: Princeton University Press.

Leongómez, J. D., Binter, J., Kubicová, L., Stolařová, P., Klapilová, K., Havlíček, J., & Roberts, S. C. (2014). Vocal modulation during courtship increases proceptivity even in naive listeners. *Evolution and Human Behavior*, *35*(6), 489–496. https://doi.org/10.1016/j.evolhumbehav.2014.06.008

Little, A. C., Jones, B. C., DeBruine, L. M., & Dunbar, R. I. M. (2013). Accuracy in discrimination of self-reported cooperators using static facial information. *Personality and Individual Differences*, *54*(4), 507–512. https://doi.org/10.1016/j.paid.2012.10.018

Lukas, W. D., Campbell, B. C., & Ellison, P. T. (2004). Testosterone, aging, and body composition in men from Harare, Zimbabwe. *American Journal of Human Biology: The Official Journal of the Human Biology Association*, *16*(6), 704–712. https://doi.org/10.1002/ajhb.20083

McAleer, P., Todorov, A., & Belin, P. (2014). How do you say 'Hello'? Personality impressions from brief novel voices. *PLoS ONE*, *9*(3), e90779. https://doi.org/10.1371/journal.pone.0090779

Moller, A., & Pomiankowski, A. (1993). Why have birds got multiple sexual ornaments? *Behavioral Ecology and Sociobiology*, *32*(3), 167–176. https://doi.org/10.1007/BF00173774

Montano, K. J., Tigue, C. C., Isenstein, S. G., Barclay, P., & Feinberg, D. R. (2017). Men's voice pitch influences women's trusting behavior. *Evolution and Human Behavior*, *38*(3), 293–297. https://doi.org/10.1016/j.evolhumbehav.2016.10.010

Munson, B., McDonald, E. C., DeBoe, N. L., & White, A. R. (2006). The acoustic and perceptual bases of judgments of women and men's sexual orientation from read speech. *Journal of Phonetics*, *34*(2), 202–240. https://doi.org/10.1016/j.wocn.2005.05.003

Nowak, J., Pawłowski, B., Borkowska, B., Augustyniak, D., & Drulis-Kawa, Z. (2018). No evidence for the immunocompetence handicap hypothesis in male humans. *Scientific Reports*, *8*(1), 7392. https://doi.org/10.1038/s41598-018-25694-0

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science*, *314*(5805), 1560–1563. https://doi.org/10.1126/science.1133755

O'Connor, J. J. M., & Barclay, P. (2017). The influence of voice pitch on perceptions of trustworthiness across social contexts. *Evolution and Human Behavior*, *38*(4), 506–512. https://doi.org/10.1016/j.evolhumbehav.2017.03.001

O'Connor, J. J. M., Re, D. E., & Feinberg, D. R. (2011). Voice pitch influences perceptions of sexual infidelity. *Evolutionary Psychology*, *9*(1), 64–78. https://doi.org/10.1177/147470491100900109

Oda, R., Naganawa, T., Yamauchi, S., Yamagata, N., & Matsumoto-Oda, A. (2009). Altruists are trusted based on non-verbal cues. *Biology Letters*, *5*(6), 752–754. https://doi.org/10.1098/rsbl.2009.0332

Oda, R., Yamagata, N., Yabiku, Y., & Matsumoto-Oda, A. (2009). Altruism can be assessed correctly based on Impression. *Human Nature*, *20*(3), 331–341. https://doi.org/10.1007/s12110-009-9070-8

Oleszkiewicz, A., Pisanski, K., Lachowicz-Tabaczek, K., & Sorokowska, A. (2017). Voice-based assessments of trustworthiness, competence, and warmth in blind and sighted adults. *Psychonomic Bulletin & Review*, *24*(3), 856–862. https://doi.org/10.3758/s13423-016-1146-y

Ordin, M., & Mennen, I. (2017). Cross-linguistic differences in bilinguals' fundamental frequency ranges. *Journal of Speech, Language, and Hearing Research*, *60*(6), 1493–1506. https://doi.org/10.1044/2016_JSLHR-S-16-0315

Pedersen, M., Møller, S., Krabbe, S., & Bennett, P. (1986). Fundamental voice frequency measured by electroglottography during continuous speech. A new exact secondary sex characteristic in boys in puberty. *International Journal of Pediatric Otorhinolaryngology*, *11*(1), 21–27. https://doi.org/10.1016/S0165-5876(86)80024-6

Pépiot, E. (2014). Male and female speech: a study of mean f0, f0 range, phonation type and speech rate in Parisian French and American English speakers. *Speech Prosody*, *7*, 305–309.

Peysakhovich, A., Nowak, M. A., & Rand, D. G. (2014). Humans display a 'cooperative phenotype' that is domain general and temporally stable. *Nature Communications*, *5*, 4939. https://doi.org/10.1038/ncomms5939

Pisanski, K., Jones, B. J., Fink, B., O'Connor, J. J. M., DeBruine, L. M., Röder, S., & Feinberg, D. R. (2016). Voice parameters predict sex-specific body morphology in men and women. *Animal Behaviour*, *112*, 13–22.

Ponsot, E., Burred, J. J., Belin, P., & Aucouturier, J.-J. (2018). Cracking the social code of speech prosody using reverse correlation. *Proceedings of the National Academy of Sciences*, *115*(5), 3972. https://doi.org/10.1073/pnas.1716090115

Puts, D. A., Apicella, C. L., & Cárdenas, R. A. (2012). Masculine voices signal men's threat potential in forager and industrial societies. *Proceedings of the Royal Society B: Biological Sciences*, *279* (1728), 601. https://doi.org/10.1098/rspb.2011.0829

Puts, D. A., Gaulin, S. J. C., & Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evolution and Human Behavior*, *27*(4), 283–296. https://doi.org/10.1016/j.evolhumbehav.2005.11.003

Puts, D. A., Hodges, C. R., Cárdenas, R. A., & Gaulin, S. J. C. (2007). Men's voices as dominance signals: Vocal fundamental and formant frequencies influence dominance attributions among men. *Evolution and Human Behavior*, *28*(5), 340–344. https://doi.org/10.1016/j.evolhumbehav.2007.05.002

Rabinov, C. R., Kreiman, J., Gerratt, B. R., & Bielamowicz, S. (1995). Comparing reliability of perceptual ratings of roughness and acoustic measures of jitter. *Journal of Speech, Language, and Hearing Research*, *38*(1), 26–32. https://doi.org/10.1044/jshr.3801.26

Raihani, N. J., & Smith, S. (2015). Competitive helping in online giving. *Current Biology*, *25*(9), 1183–1186. https://doi.org/10.1016/j.cub.2015.02.042

Rand, D. G., & Nowak, M. A. (2013). Human cooperation. *Trends in Cognitive Sciences*, *17*(8), 413–425. https://doi.org/10.1016/j.tics.2013.06.003

Rantala, M. J., Moore, F. R., Skrinda, I., Krama, T., Kivleniece, I., Kecko, S., & Krams, I. (2012). Evidence for the stress-linked immunocompetence handicap hypothesis in humans. *Nature Communications*, *3*, 694. https://doi.org/10.1038/ncomms1696

Reed, L. I., Zeglen, K. N., & Schmidt, K. L. (2012). Facial expressions as honest signals of cooperative intent in a one-shot anonymous Prisoner's Dilemma game. *Evolution and Human Behavior*, *33* (3), 200–209. https://doi.org/10.1016/j.evolhumbehav.2011.09.003

Reimers, L., & Diekhof, E. K. (2015). Testosterone is associated with cooperation during intergroup competition by enhancing parochial altruism. *Frontiers in Neuroscience*, *9*, 183. https://doi.org/10.3389/fnins.2015.00183

Rustagi, D., Engel, S., & Kosfeld, M. (2010). Conditional cooperation and costly monitoring explain success in forest commons management. *Science*, *330*(6006), 961–965. https://doi.org/10.1126/science.1193649

Schielzeth, H. (2010). Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution*, *1*(2), 103–113. https://doi.org/10.1111/j.2041-210X.2010.00012.x

Schirmer, A., Feng, Y., Sen, A., & Penney, T. B. (2019). Angry, old, male – and trustworthy? How expressive and person voice characteristics shape listener trust. *PLoS ONE*, *14*(1), e0210555. https://doi.org/10.1371/journal.pone.0210555

Schötz, S., & Müller, C. (2007). A study of acoustic correlates of speaker age. In C. Müller (Eds.), *Speaker classification II: Selected projects*, (pp. 1–9). Berlin: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-74122-0_1

Sorensen, D., & Horii, Y. (1982). Cigarette smoking and voice fundamental frequency. *Journal of Communication Disorders*, *15*(2), 135–144. https://doi.org/10.1016/0021-9924(82)90027-2

Stirrat, M., & Perrett, D. I. (2010). Valid facial cues to cooperation and trust: Male facial width and trustworthiness. *Psychological Science*, *21*(3), 349–354. https://doi.org/10.1177/0956797610362647

Stirrat, M., & Perrett, D. I. (2012). Face structure predicts cooperation: Men with wider faces are more generous to their in-group when out-group competition is salient. *Psychological Science*, *23*(7), 718–722. https://doi.org/10.1177/0956797611435133

Suire, A., Raymond, M., & Barkat-Defradas, M. (2018). Human vocal behavior within competitive and courtship contexts and its relation to mating success. *Evolution and Human Behavior*, *39*(6), 684–691. https://doi.org/10.1016/j.evolhumbehav.2018.07.001

Sutter, M., & Kocher, M. G. (2007). Trust and trustworthiness across different age groups. *Games and Economic Behavior*, *59*(2), 364–382. https://doi.org/10.1016/j.geb.2006.07.006

Sylwester, K., & Roberts, G. (2010). Cooperators benefit through reputation-based partner choice in economic games. *Biology Letters*, *6*(5), 659–662. https://doi.org/10.1098/rsbl.2010.0209

Takagishi, H., Takahashi, T., & Yamagishi, T. (2011). Testosterone diminishes reciprocity in a trust game. *Abstracts of the 34th Annual Meeting of the Japan Neuroscience Society*, *71*, e282. https://doi.org/10.1016/j.neures.2011.07.1230

Tigue, C. C., Borak, D. J., O'Connor, J. J. M., Schandl, C., & Feinberg, D. R. (2012). Voice pitch influences voting behavior. *Evolution and Human Behavior*, *33*(3), 210–216. https://doi.org/10.1016/j.evolhumbehav.2011.09.004

Tognetti, A., Berticat, C., Raymond, M., & Faurie, C. (2012). Sexual selection of human cooperative behaviour: An experimental study in rural Senegal. *PLoS ONE*, *7*(9), e44403. https://doi.org/10.1371/journal.pone.0044403

Tognetti, A., Berticat, C., Raymond, M., & Faurie, C. (2013). Is cooperativeness readable in static facial features? An inter-cultural approach. *Evolution and Human Behavior*, *34*(6), 427–432. https://doi.org/10.1016/j.evolhumbehav.2013.08.002

Tognetti, A., Dubois, D., Faurie, C., & Willinger, M. (2016). Men increase contributions to a public good when under sexual competition. *Scientific Reports*, *6*, 29819. https://doi.org/10.1038/srep29819

Tognetti, A., Yamagata-Nakashima, N., Faurie, C., & Oda, R. (2018). Are non-verbal facial cues of altruism cross-culturally readable? *Personality and Individual Differences*, *127*, 139–143. https://doi.org/10.1016/j.paid.2018.02.007

Torre, I., White, L., & Goslin, J. (2016). Behavioural mediation of prosodic cues to implicit judgements of trustworthiness. *In Proceedings of Speech Prosody*, 816–820. https://doi.org/10.21437/SpeechProsody.2016-167

Torre, P., & Barlow, J. A. (2009). Age-related changes in acoustic characteristics of adult speech. *Journal of Communication Disorders*, *42*(5), 324–333. https://doi.org/10.1016/j.jcomdis.2009.03.001

Tsantani, M. S., Belin, P., Paterson, H. M., & McAleer, P. (2016). Low vocal pitch preference drives first impressions irrespective of context in male voices but not in female voices. *Perception*, *45*(8), 946–963. https://doi.org/10.1177/0301006616643675

Van Borsel, J., Janssens, J., & De Bodt, M. (2009). Breathiness as a feminine voice characteristic: A perceptual approach. *Journal of Voice*, *23*(3), 291–294. https://doi.org/10.1016/j.jvoice.2007.08.002

Vittek, J., L'Hommedieu, D. G., Gordon, G. G., Rappaport, S. C., & Louis Southren, A. (1985). Direct radioimmunoassay (RIA) of salivary testosterone: Correlation with free and total serum testosterone. *Life Sciences*, *37*(8), 711–716. https://doi.org/10.1016/0024-3205(85)90540-5

Weirich, M. (2008). Vocal stereotypes. In A. Botinis (Ed.), *ExLing-2008* (pp. 229–232). Athens, Greece: International Speech Communication Association.

Werth, K., Voigt, D., Döllinger, M., Eysholdt, U., & Lohscheller, J. (2010). Clinical value of acoustic voice measures: A retrospective study. *European Archives of Oto-Rhino-Laryngology*, *267*(8), 1261–1271. https://doi.org/10.1007/s00405-010-1214-2

West, S. A., Griffin, A. S., & Gardner, A. (2007). Evolutionary explanations for cooperation. *Current Biology*, *17*(16), 661–672. https://doi.org/10.1016/j.cub.2007.06.004

Zak, P. J., Kurzban, R., Ahmadi, S., Swerdloff, R. S., Park, J., Efremidze, L., . . . Matzner, W. (2009). Testosterone administration decreases generosity in the ultimatum game. *PLoS ONE*, *4*(12), e8330. https://doi.org/10.1371/journal.pone.0008330

Zimmerer, F., Jügler, J., Andreeva, B., Möbius, B., & Trouvain, J. (2014). Too cautious to vary more? A comparison of pitch variation in native and non-native productions of French and German speakers. In N. Campbell, D. Gibbon & D. Hirst (Eds.), *Proceedings to the 7th Speech Prosody Conference* (pp. 1037–1041). Dublin, Ireland: International Speech Communication Association.

Received 4 April 2019; revised version received 28 October 2019

Supporting Information

The following supporting information may be found in the online edition of the article:

Figure S1. Picture shown to the participants when we recorded their spontaneous speech.
Figure S2. Conditional contributions per subject. Subjects were classified following the procedure of Fischbacher, Gächter and Fehr (2001)
Table S1. Censored regression models investigating the effect of men's speech on contributions to the public good (n = 64 men).
Table S2. Logistic regressions investigating whether free riders (n = 11) relatively to conditional cooperators (n = 44) differ in term of acoustic parameters.
Table S3. Linear regressions examining the effect of testosterone on each of the four acoustic parameters studied (n = 63).