Description of the variables contained in the data\_speech\_features.csv file, used for the study described in ‘Speech Acoustic Features: A Comparison of Gay Men, Heterosexual Men, and Heterosexual Women’ by Suire/Tognetti et al. (Archives of Sexual Behavior, 2020, in press). For additional information on the dataset, please contact A. Suire (alexandresuire@hotmail.fr) or A. Tognetti (arnaud.tognetti@gmail.com).

Each row corresponds to the characteristics of one individual:

* ID = the anonymous ID given to the participant
* sex = sex of the participant

1. M = male
2. F = female

* age = age of the participant (in years)
* orientation = sexual orientation of the participant (heterosexual/homosexual)
* Testosterone = testosterone level (pg/ml)
* time = speaking duration (s)
* meanf0 = mean fundamental frequency (Hz)
* f0sd = variation of the fundamental frequency (Hz)
* jitter = amplitude perturbations of the fundamental frequency (%)
* HNR = Harmonics-to-Noise Ratio (dB)
* country\_birth = birth country of the participant
* mother\_country = birth country of the mother
* gp\_father\_maternal\_country = birth country of the father of the mother
* gp\_mother\_maternal\_country = birth country of the mother of the mother
* father\_country = birth country of the father
* gp\_father\_paternal\_country = birth country of the father of the father
* gp\_mother\_paternal\_country = birth country of the mother of the father
* relationship = whether the participant was in a romantic relationship at the moment of the study (yes/no)
* education = education level of the participant. ‘Bac’ corresponds to high school graduation. The number following refers to the number of year in the higher education level. ‘Bep’ and/or ‘Cap-Bep’ corresponds to an individual that has not completed high school but has obtained a degree. The specific case ‘3\_eme’ refers to an individual that has not even gone to high school and is thus considered as the lowest education level in France. From lowest to highest education: 3\_eme > bep > bac > bac+1 > bac+2 > bac+3 > bac+4 > bac+5 > bac+6 > bac+7 > bac+8.
* monthly\_income = income of the participant, rated on a Likert-Scale, from 1 (very low income) to 6 (highest income).

Please send an email to the authors for additional information.