

Evolution of human pitch: Preliminary analyses in the French population using INA audiovisual archives of Vox Pops

Alexandre Suire¹ & Melissa Barkat-Defradas²

¹Post-doctoral Fellow – University of the Ryukyus, Japan

²CNRS Researcher – University of Montpellier, France

Introduction

Beyond its linguistic content, the human voice is a formidable tool to convey a wide range of emotions and affective states. Moreover, it also reflects some biological information about the emitter, such as sex (with women having twice as high vocal height than men), age (voice quality changing over lifespan) as well as some personality traits. Another important information inferred from voice is the speakers' socio-cultural background. Indeed, as vocal registers can vary from one language to another, it has been suggested that beyond the tonal variations due to the the well known physio-anatomical differences between sexes and populations, some vocal specificities can also be due to the environment [1] and that part of what makes male vs female speech variation is socially “learned” [2].

Indeed, a large body of literature shows that tonal registers can be highly variable from one language to another. For instance, Traunmüller & Eriksson (1995) [3] reported in a largely supported description of examples that there is a wide variability in voice pitch (measured by the fundamental frequency F_0 , which accounts for vocal folds vibration) between languages. For instance in a group of men matched in age and recorded in the same conditions, Bulgarian and/or Polish males attest a higher F_0 than German and/or British English (160 Hz and 163 Hz vs. 118 Hz and 128 Hz, respectively [4]). Other multilingual studies have shown that British English speakers have lower F_0 than Mandarin Chinese speakers (107 Hz and 129 Hz respectively, [5]) and differences in F_0 between American English and Spanish and Japanese speakers have also been highlighted (105.6 Hz, 124.6 Hz and 130.6 Hz respectively, [6]). As for women and under similar experimental conditions, American women have an lower F_0 than Japanese (211 Hz vs. 224 Hz respectively, [7]), while Bulgarian and Polish women have a F_0 higher than German and English women (272 Hz and 266 Hz vs. 210 Hz and 217 Hz respectively, [4]). Other examples include those of Rose (1991) [8] who showed that in the Chinese dialects of Wù, the F_0 in men and women are respectively 170 Hz and 187 Hz. According to the authors one possible explanation for such variation may lie in a socially learned “vocal behavior”, influenced by the preferences and social practices of a given culture or society. Such an idea has been extensively investigated in sociolinguistics, which explores the ways in which language and society are interrelated and more particularly the ways in which social representations and cultural expectations act upon language.

van Bezooijen [9] showed that the significant difference in voice pitch observed between Japanese and Dutch women (~230 Hz vs. ~190 Hz respectively) are explained by the fact that they comply to social expectations linked to their gender, which vary considerably

between the two cultures. In Japan for example, traits traditionally associated with femininity include “modesty”, “innocence” and “psychological dependence”. In order to conform to the preferences of men sharing the same cultural representations, Japanese women vocally express these traits by using a higher pitch, which is known to be perceptually associated with impressions of “docility” and “weakness” [10]. Conversely, in a culture where the social roles attributed to men and women are relatively more egalitarian like, for example, the Netherlands, Dutch men tend to value traits related to female “independence” and “hard-working”, which result in a marked preference for women with lower-pitched voices, that are known to be aurally associated with “self-confidence” and “leadership” [11]. Thus, Dutch men, in the same way as Japanese men, show preferences towards culturally congruent voices related to the values attributed to femininity. More generally, this illustrates the idea that voice pitch reflects how a given culture considers men and women social roles. As a matter of fact, voice is a crucial feature for the definition of both sexual and social identity. The way men and women talk (and/or the way we think they should speak) reveals the evolution of our beliefs about masculinity and/or femininity that is, at the same time, shaped by these conceptions since voice quality relates the speaker to his/her social role; namely it is important to talk “as a man” or “as a woman” so as to be recognized and accepted as such. As a matter of fact it would be spurious to assume male vs. female voices differ only because of biological factors. Culture also intervenes in the definition of what makes a voice feminine vs. masculine by establishing contrastive norms for both sexes [12]. A significant part of vocal behavior far from being “natural” is conceived to reinforce and emphasize sexual differences. This can be demonstrated by the fact that it is possible to determine the gender of prepubertal children prior to voice break. Indeed, although 5 to 6 year-old girls’ and boys’ voices often overlap in height (i.e. F_0), acoustic differences remain too remarkable to be solely imputed to their anatomical/physiological dissimilarities [13]. As a matter of fact, it has been shown children also learn how to talk through imitation according to the social and cultural expectations attributed to their sex/gender by replicating their same-sex parents’ vocal register [14].

Moreover and interestingly enough, some evidence suggests that gender representations have significantly evolved over time, at least for women from Western’s societies. Indeed, several studies have shown that women are increasingly moving away from feminine stereotypes by adopting traits that were up to there generally associated with masculinity, such as competitiveness, independence, leadership, as well as declaring themselves as being more orientated towards higher professional careers [15,16]. In turn, such changes would shift the use of their voices towards a lower-frequency pitch in order to project those impressions, to which men seem to be more attracted nowadays [17]. Besides, two studies have suggested a change in pitch over time. Pemberton et al. (1998) compared the recordings of 28 women between 18 and 25 years old in 1945 to recordings of a similar group of women in 1993. Their results show a difference in F_0 of about 23 Hz (which is far above the frequency just-noticeable difference¹), voice pitch being significantly lower in the second group. In a second study, it was shown that the voice pitch of French male journalists had also lowered over the past fifty years [18]. However, besides these two studies, there is currently a

¹ Knowing that the minimal change in the frequency of a tone that can be detected by the ear is about 0.5% of the

strong lack of information on whether voice pitch has evolved or not across a larger sample size to confirm such results.

In this context, audiovisual archives offer a unique opportunity to tackle this issue and the present project aims at exploring the evolution of both men and women voice pitch in the French population as a first study case. For this purpose, we focused on Vox Pops recorded for television news that are most representative for how individuals vocally behave within a given society at a given time. Therefore, the relevance of this work is twofold: (1) offer an original systematic analysis of how the human vocal pitch has evolved over time and (2) highlight how audiovisual archives can shed light on the evolution of human social behaviors such as speech.

Methods

Voice samples

Voice samples were selected from French speakers interviewed during TV news Vox Pops between 1940 and 2019. These recordings generally took place in public areas, and involved anonymous passers-by who were asked to give their opinion on a particular subject. Recordings were provided by the INA (Institut National de l'Audiovisuel, France) as online video clips from which we selected a series of suitable ones for our purpose. The reason for choosing this kind of material is that average citizens' spontaneous speech was felt to be more representative of the way of speaking of the general population than that, for instance, of interviewers, TV presenters, comedians or singers who have generally a better control of their vocal apparatus.

Clips with high background noise or chaotic vocalizations (e.g. mechanical noises, such as engine noises; street rumors and/or screaming, shouting, laughing, crying...) were excluded. Eventually, 96 video clips were selected. Then, we extracted the audio signals associated from each selected video and saved them as .WAV files for further analyses.

Voice analysis

Speech segmentation and acoustic analyses of audio signals were conducted using Praat© v. 6.0.43 [19]. Segmentation consisted in isolating speech samples (i.e. speakers talking) by removing fragments of external noises (i.e. surrounding noises), multi-voicing (i.e. voice overlaps) and non-linguistic vocalizations (e.g. laughter, crying...). For some speakers, multiple speech samples were selected since the former could appear several times in the same interview. Then, we used a batch-processing script to measure four acoustic parameters of fundamental frequency ($F0$, Hz): mean ($F0_{\text{mean}}$), range ($F0_{\text{min}}$ and $F0_{\text{max}}$) and $F0$ contour with its standard variation ($F0_{\text{sd}}$). Then, we measured the coefficient of variation $F0_{\text{CV}}$, which corresponds to $F0_{\text{sd}} / F0_{\text{mean}}$. This measure represents the logarithmic perception² of voice pitch (i.e. controlling for the effect of magnitude on variability [15,17]). All $F0$ parameters were measured using Praat's autocorrelation algorithm with a search range of 100–500 Hz for

² The frequency response of the ear is likewise nonlinear.

women speakers and 75–300 Hz for men speakers. We also measured the duration (s) of each voiced segments. For speakers with multiple voice samples, we averaged all five $F0$ parameters.

Perceptually, $F0_{\text{mean}}$, $F0_{\text{min}}$ and $F0_{\text{max}}$ refer to voice pitch (i.e. vocal height) while $F0_{\text{sd}}$ and $F0_{\text{CV}}$ refer to intonation (i.e. vocal modulation). As such, voices with relatively lower $F0_{\text{mean}}$, $F0_{\text{min}}$ and $F0_{\text{max}}$ will sound lower-pitched or “deeper” whereas voices exhibiting relatively higher $F0_{\text{mean}}$, $F0_{\text{min}}$ and $F0_{\text{max}}$ will sound higher-pitched. Moreover, voices with relatively lower $F0_{\text{sd}}$ and $F0_{\text{CV}}$ will sound more monotonous whereas voices exhibiting relatively higher $F0_{\text{sd}}$ and $F0_{\text{CV}}$ will sound more dynamic and expressive.

Note that remarkable differences are perceptually noticeable between the two sexes. Women’s voices are significantly higher in pitch than men (approximately twice as much) as well as more expressive. These characteristics are consistent across languages. For instance, mean pitch ($F0_{\text{mean}}$) of French speaking women vary between 140 and 240 Hz and between 100 and 150 Hz in men [20], while pitch variation $F0_{\text{sd}}$ is around 30 Hz in women and 13 Hz in men [21].

Analyses and results

Except for the period 1940 – 1959, women and men’s voices were analyzed separately by decade, from 1960 – 1969, 1970 – 1979, etc. to 2010 – 2019 (see Table 1 for details concerning the number of speakers analyzed for each decade). In total, we analyzed 341 women and 438 men unique speakers across 7 decades. Speech samples averaged duration is about 4.7 seconds (± 4.6).

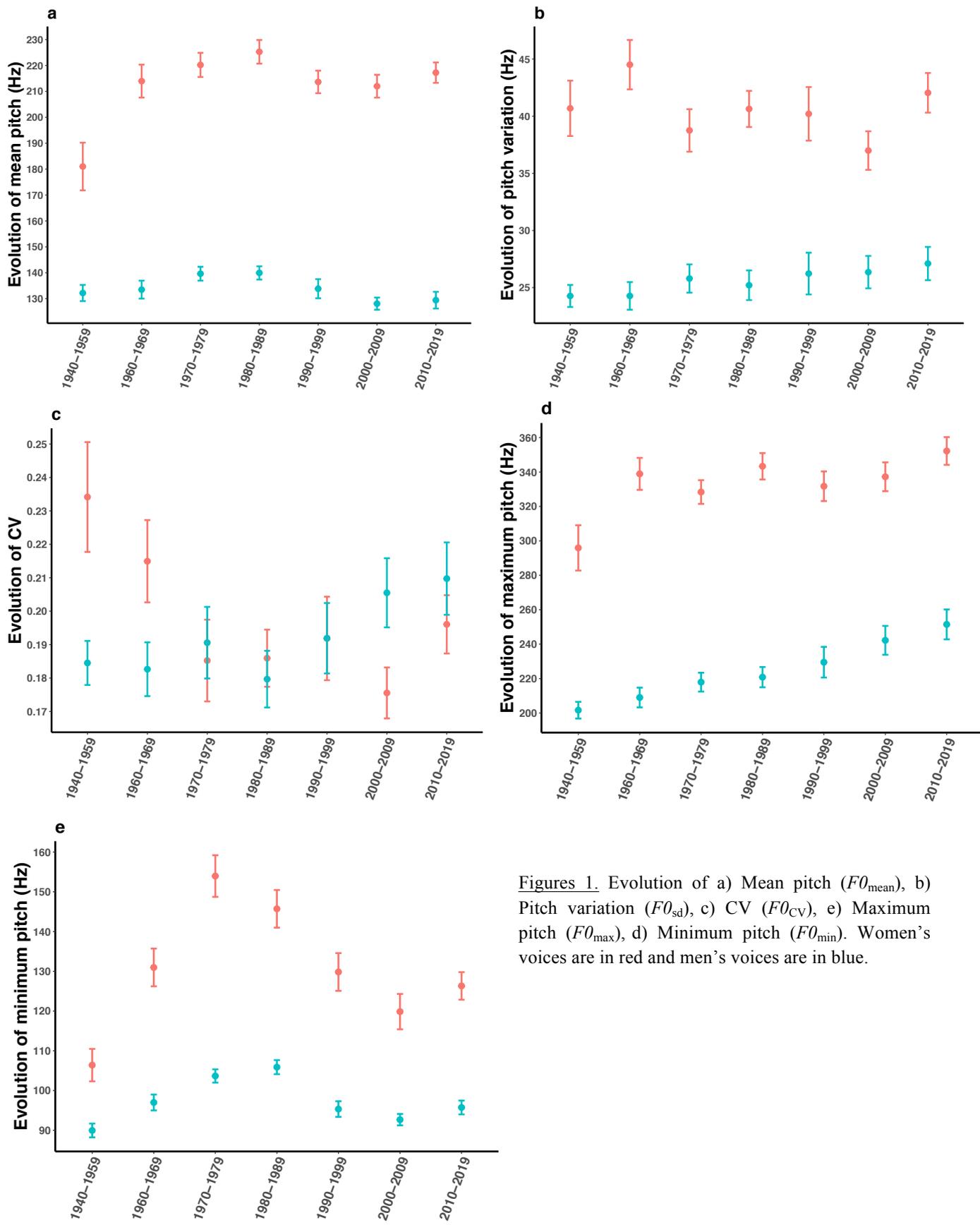
Table 1. Number of analyzed speakers per sex and decade

	Women	Men
1940 – 1959	27	48
1960 – 1969	49	37
1970 – 1979	50	80
1980 – 1989	62	83
1990 – 1999	42	62
2000 – 2009	49	56
2010 – 2019	62	72

ANOVAs (Analyses of Variance) were used to examine the evolution of all $F0$ parameters over time (i.e. per decade) and for each sex. Tukey post-hoc comparisons with Bonferroni’s correction were then performed to look for specific differences between each decade. Results were considered significant at the $\alpha = .05$ level. All results are summarized in Figure 1.

Mean pitch ($F0_{\text{mean}}$) was significantly higher in women after 1940 – 1959 (all $p < .05$) but remained constant after that specific period. In men, no significant increase or decrease was observed between each decade and thus mean pitch remained constant over almost ~80 years (all $p > .05$). Pitch variation ($F0_{\text{sd}}$) did not significantly vary in both women and men

across time (all $p > .05$). Although there was considerable variation, a look at the evolution of $F0_{CV}$ ($F0$ coefficient of variation), which corrects for the logarithmic perception of voice pitch [11], shows that it was only significantly lower between 1940 – 1959 and 2000 – 2009 ($p < .05$). Women's $F0_{CV}$ also overlapped the values observed for men between 1970 – 1979 and 1990 – 1999. In other terms, it appears that women's voices became relatively more monotonous over time. Values of $F0_{CV}$ for men did not significantly change over time (all $p > .05$). Maximum pitch ($F0_{max}$) was significantly higher in women after 1940 – 1959, with another important increase in 2010 – 2019 (all $p < .05$), while in men it only started to be significantly higher between 2000 – 2009 and 2010 – 2019 compared to 1940 – 1959 (all $p < .05$). Minimum pitch ($F0_{min}$) significantly increased after 1940 – 1959 in 1970 – 1979 and 1980 – 1989, but significantly decreased after those decades between 1980 – 1989 and 2010 – 2019 (all $p < .05$). In men, a similar pattern was observed (all $p < .05$). In summary, these results reveal that women's voices have evolved higher pitched voices after 1940 – 1959 but remained constant after that. We also observed an increase in vocal range (i.e. simultaneous increase in $F0_{max}$ and decrease in $F0_{min}$) and a more monotonous voice after 1970 – 1979. For men, voice pitch has not significantly changed over time but vocal range also increased after the same period.



Figures 1. Evolution of a) Mean pitch ($F0_{\text{mean}}$), b) Pitch variation ($F0_{\text{sd}}$), c) CV ($F0_{\text{CV}}$), e) Maximum pitch ($F0_{\text{max}}$), d) Minimum pitch ($F0_{\text{min}}$). Women's voices are in red and men's voices are in blue.

Discussion and future perspectives

Our study shows that voice pitch ($F0_{\text{mean}}$) in both men and women has little evolved over time (Figure 1a). Those results seem to contradict the hypothesis of a female vocal masculinization, as initially suggested. In this context, we suggest two possible explanations. First, $F0_{\text{mean}}$ may not be the most reliable acoustic parameter to study the evolution of voice over time as it captures relatively little information on how an individual behaves vocally (i.e. static cue). Second, lack of significant changes could also be explained by an important factor we could not control for: individuals could use their voice pitch differently depending on their social positions such as a lower vs. a higher voice pitch in high- vs. low-ranking individuals, respectively [11,23,24]. Indeed, some studies showed the existence of a link between vocal characteristics and social class [24–27]. Here, the relevance of INA archives (or, for that matter, any other audiovisual archives) should be underlined since a detailed analysis of available visual data could be particularly useful to verify the validity of this hypothesis. For instance, a fine-grained analysis of clothes could allow us to infer the interviewed individual's social background and account for an original method to further investigate this topic. On the linguistic dimension we suggest to rely on the study of lexical richness and/or syntactic complexity (that are known to vary depending on the speakers' socio-economic and/or educative background) to correlate linguistic observation to visual subjective judgments. Eventually, it could be interesting to compare these inferences to naïve subjects' socio-economic judgments of recorded people.

Nonetheless, considering more dynamic acoustic criteria (i.e. $F0_{\text{sd}}$, $F0_{\text{CV}}$, $F0_{\text{min}}$, $F0_{\text{max}}$), such features revealed interesting patterns. Figure 1d and 1e (i.e. $F0_{\text{min}}$ and $F0_{\text{max}}$) show that both men and women's vocal range increased (simultaneous increase of $F0_{\text{max}}$ and decrease of $F0_{\text{min}}$) from the 1970's to nowadays. Considering that $F0_{\text{mean}}$ and $F0_{\text{sd}}$ did not change significantly, we could interpret these results by suggesting that two “vocal profiles” coexist within both sexes in the French population. All things being equal, this would mean that a part of the population progressively spoke over time in a more lower pitch register and another in a higher pitch register. Again, one possible explanation we suggest is that different social classes are using different voice registers; that is to say socially distinctive vocal behaviors, which became more marked over time. This hypothesis could also partially confirm that, at least in some parts of the population, there has been a “*female vocal masculinization*” accompanied by an unexpected “*male vocal feminization*”.

Along with results of a respective decrease and increase in $F0_{\text{CV}}$ in women and to a lesser extent in men (respectively considered as a masculine vs. feminine vocal trait [28]), such results occurring from the 1970's could be interpreted by the fact that some men and women started to depart themselves vocally from the social expectations attributed to their sex and gender at the very beginning of the 1960's because of the relaxation of social taboos and norms [29]. Indeed, the culture of decades prior to the 1960's was described as being unable to contain the demands for greater individual freedom, which saw people breaking free of the social constraints through extreme deviation from the norm. Such a social phenomenon seems to corroborate the hypothesized “masculinization” of women. Indeed, a large amount of studies have shown that they tend to adopt traits generally attributed to men such as competitiveness, independence and leadership. Likewise, they declare themselves as being

more orientated towards higher professional careers [15,16]. Although this major socio-historical phenomenon has been studied in the context of Humanities (i.e. sociology, history...), our experimental study is, to our knowledge, the first one to investigate the effect of socio-historical change on the evolution human voice and, therefore, and allows us hypothesizing that the evolution of vocal characteristics may be considered as a reliable sign of cultural evolution.

Although results are few, this work offers a more systematic analysis on the evolution of voice pitch in both native French speaking men and women. Cross-sectional studies are particularly important because they answer questions regarding the evolution of human behavior. As such, the use of audiovisual archives for that purpose is useful and extremely precious. In this context, we suggest several avenues for further research. Firstly, interdisciplinary work involving historians on sex and gender representations would be particularly welcomed to better interpret the different changes in voice that have occurred over time. Secondly, other acoustic parameters should be given attention as the human voice cannot be defined solely on the basis of voice pitch and modulation but also on timbre (vocal resonances), quality (roughness, breathiness...) and rhythm (speech tempo). Furthermore, as all languages can significantly differ across all these aspects, comparing the evolution of voices from different cultures is a particularly interesting topic. Indeed, this could give us important cues on how gender and social roles have evolved over time across each culture. Turning to evolutionary biology, our work raises an issue for which no clear-cut answer has been proposed yet. Should we consider men and women have learnt to use contrasted vocal patterns primarily because of education and social conventions or did biology give them different voices to which human culture has subsequently attributed different values?

Acknowledgments

We would like to thank the INA for providing us access to the audiovisual archives. We would also like to thank the International Federation of Television Archives for funding and allowing us to perform this original study.

Access to data

Data can be requested by sending an email to Alexandre Suire: alexandresuire@hotmail.fr

References

1. Everett C, Blasi DE, Roberts SG. 2015 Climate, vocal folds, and tonal languages: Connecting the physiological and geographic dots. *Proceedings of the National Academy of Sciences* **112**, 1322–1327. (doi:10.1073/pnas.1417413112)
2. Simpson AP. 2009 Phonetic differences between male and female speech. *Language and Linguistics Compass* **3**, 621–640. (doi:10.1111/j.1749-818X.2009.00125.x)
3. Traunmüller H, Eriksson A. 1995 The frequency range of the voice fundamental in the speech of male and female adults. *Unpublished manuscript*, 11.
4. Andreeva B, Demenko G, Mobius B, Zimmerer F, Jugler J, Oleskowicz-Popiel M. 2014

Differences of Pitch Profiles in Germanic and Slavic Languages. *Fifteenth Annual Conference of the International Speech Communication Association*.

5. Keating P, Kuo G. 2012 Comparison of speaking fundamental frequency in English and Mandarin. *The Journal of the Acoustical Society of America* **132**, 1050–1060. (doi:10.1121/1.4730893)
6. Hanley TD, Snidecor JC, Ringel RL. 1966 Some acoustic differences among languages. *Phonetica* **14**, 97–107.
7. Yamazawa H, Hollien H. 1992 Speaking fundamental frequency patterns of Japanese women. *Phonetica* **49**, 128–140.
8. Rose P. 1991 How effective are long term mean and standard deviation as normalisation parameters for tonal fundamental frequency? *Speech Communication* **10**, 229–247. (doi:10.1016/0167-6393(91)90014-K)
9. van Bezooijen R. 1995 Sociocultural Aspects of Pitch Differences between Japanese and Dutch Women. *Language and Speech* **38**, 253–265. (doi:10.1177/002383099503800303)
10. Pisanski K, Rendall D. 2011 The prioritization of voice fundamental frequency or formants in listeners' assessments of speaker size, masculinity, and attractiveness. *The Journal of the Acoustical Society of America* **129**, 2201–2212. (doi:10.1121/1.3552866)
11. Klofstad CA, Anderson RC, 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**, 2698–2704. (doi:10.1098/rspb.2012.0311)
12. Weinberg, B., & Bennett, S. 1971 Speaker Sex Recognition of 5-and 6-year-old Children's Voices. *The Journal of the Acoustical Society of America*, **50**, 1210-1213.
13. Andersen, E. 2014 *Speaking With Style (RLE Linguistics C: Applied Linguistics): The Sociolinguistics Skills of Children*. Routledge.
14. Meditch, A. 1975 The development of sex-specific speech patterns in young children. *Anthropological Linguistics*, 421-433.
15. Cashdan E. 2008 Waist-to-Hip Ratio across Cultures: Trade-Offs between Androgen- and Estrogen-Dependent Traits. *Current Anthropology* **49**, 1099–1107. (doi:10.1086/593036)
16. Donnelly K, Twenge JM. 2017 Masculine and Feminine Traits on the Bem Sex-Role Inventory, 1993–2012: a Cross-Temporal Meta-Analysis. *Sex Roles* **76**, 556–565. (doi:10.1007/s11199-016-0625-y)
17. Pisanski K, Oleszkiewicz A, Plachetka J, Gmiterek M, Reby D. 2018 Voice pitch modulation in human mate choice. *Proceedings of the Royal Society B: Biological Sciences* **285**, 20181634. (doi:10.1098/rspb.2018.1634)
18. de Mareüil PB, Rilliard A, Allauzen A. 2012 A Diachronic Study of Initial Stress and other Prosodic Features in the French News Announcer Style: Corpus-based Measurements and Perceptual Experiments. *Language and Speech* **55**, 263–293. (doi:10.1177/0023830911417799)
19. Boersma, P., & Weenink, D. 2017 *Praat: Doing phonetics by computer [Computer software]*.
20. Fouquet M, Pisanski K, Mathevon N, Reby D. 2016 Seven and up: individual differences in male voice fundamental frequency emerge before puberty and remain stable throughout adulthood. *Royal Society Open Science* **3**, 160395. (doi:10.1098/rsos.160395)
21. Vaissière J. 2015 *La phonétique*. Paris: Presses universitaires de France.
22. 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**, 684–691. (doi:10.1016/j.evolhumbehav.2018.07.001)
23. Anderson RC, Klofstad CA. 2012 Preference for Leaders with Masculine Voices Holds in the Case of Feminine Leadership Roles. *PLoS ONE* **7**, e51216. (doi:10.1371/journal.pone.0051216)
24. Mayew WJ, Parsons CA, Venkatachalam M. 2013 Voice pitch and the labor market success of male chief executive officers. *Evolution and Human Behavior* **34**, 243–248. (doi:10.1016/j.evolhumbehav.2013.03.001)
25. Pittam J. 1994 *Voice in social interaction*. Sage.
26. Madsen LM, Svendsen BA. 2015 10 Stylized voices of ethnicity and social divisions. , 24.
27. Trudgill P. 1997 The social differentiation of English in Norwich. In *Sociolinguistics*, pp. 179–184. London.

28. Leongómez JD, Binter J, Kubicová L, Stolařová P, Klapilová K, Havlíček J, Roberts SC. 2014 Vocal modulation during courtship increases perceptivity even in naive listeners. *Evolution and Human Behavior* **35**, 489–496. (doi:10.1016/j.evolhumbehav.2014.06.008)
29. Debolt AA, Baugess JS. 2011 *Encyclopedia of the Sixties: A Decade of Culture and Counterculture [2 volumes]: A Decade of Culture and Counterculture*. ABC-CLIO.