

Supplemental Material S1. Scales used in the evaluation of gender perception.

When reviewing the literature about voice and gender, we found that three main methods are used for the evaluation of the gender of voice by the listeners.

The first one is *binary*, in which listeners are asked to choose between a feminine voice and a masculine voice (Bellandese, 2008).

The second asks the listener to rate the confidence of the gender perception on a scale between 1 and 5, with 1 = *lowest confidence* and 5 = *highest confidence* (Hillenbrand & Clark, 2009).

The third asks to rate the femininity/masculinity on a visual analog scale that can be numbered or not. Some use the same scale with femininity at one end and masculinity on the other end; others use two different scales: one for masculinity that can be judged from *very masculine* to *not at all masculine* and one for femininity that can be judged from *very feminine* to *not at all feminine* (Munson, 2007).

The high variety of scales available reflects the abstract nature of the gender concept based on internal representations.

We have adopted the linear analog scale with *high masculinity* on one side and *high femininity* on the other side in the self-evaluation study, because we consider it the only one subtle enough to give details and nuances on how a person can perceive him- or herself.

Concerning the perception study, we opted for the confidence rate because it does reproduce everyday life situations over the phone. When they first hear the speaker, listeners try to get hints to call him "sir" or "madam." According to Eckert,

Since the gender binaries are fundamental ideological constructs, and since at least gender correlates with variation on a large scale, it would be wrong not to do binary coding. And the fact that gender plays out differently in different places in society is precisely why it does not make sense at this point to try to code for gender beyond the binary. One might also consider that with increasing transgenderism and gender fluidity, gender orientation is fast changing, and willing participation in the binary may at some not so-distant point begin to correlate seriously with age. (Eckert, 2014)

The concept of accuracy is most frequently used in the current literature in perception studies.

References

- Bellandese, M. H. (2008). Fundamental frequency and gender identification in standard esophageal and tracheoesophageal speakers. *Journal of Communication Disorders*, 42(2), 89–99.
- Eckert, P. (2014). The problem with binaries: Coding for gender and sexuality. *Language and Linguistics Compass*, 8, 529–535. doi:10.1111/lnc3.12113
- Hillenbrand, J. M., & Clark, M. J. (2009). The role of f_0 and formant frequencies in distinguishing the voices of men and women. *Attention, Perception, & Psychophysics*, 71, 1150–1166. doi:10.3758/APP.71.5.1150
- Munson, B. (2007). The acoustic correlates of perceived masculinity, perceived femininity, and perceived sexual orientation. *Language and Speech*, 50, 125–142.