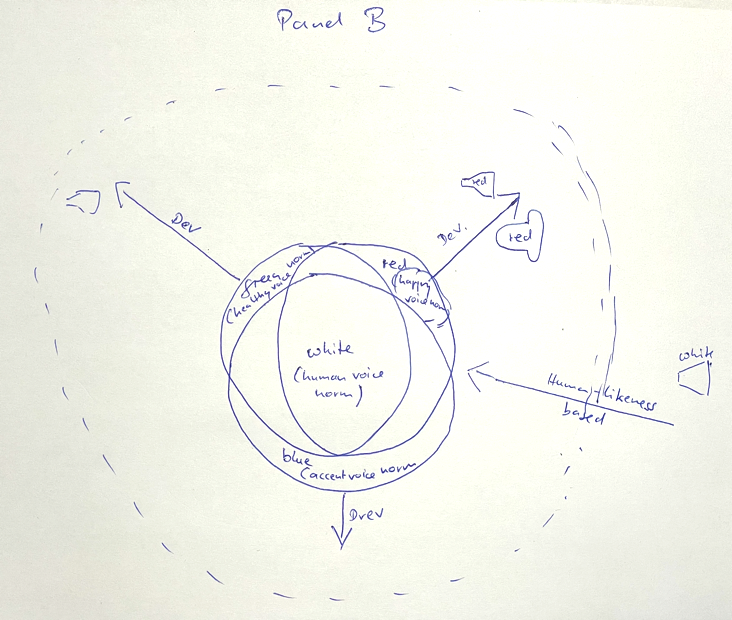
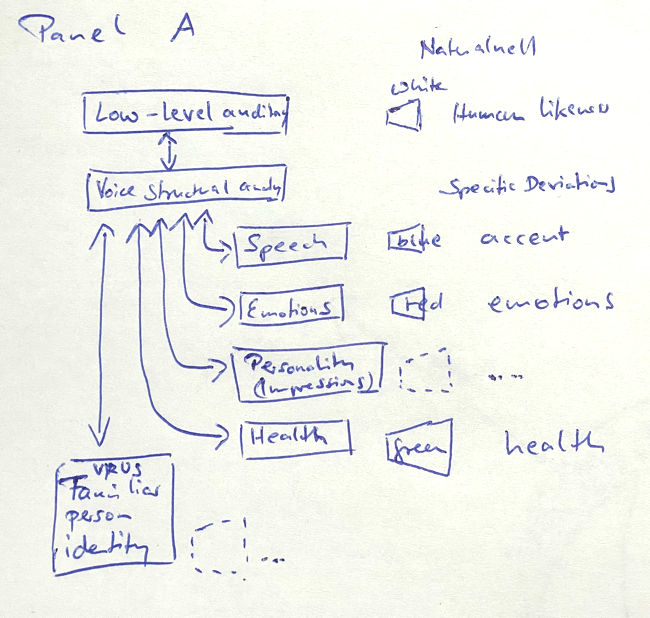
Notes on Naturalness paper, June 11, 2024, key Figure Sketch



Here´s my drawing for the Figure – Panel A inspired by Sascha´s and Panel B inspired by Tine´s go. Sorry for the awful quality – hope you can read my claw (zoom in might help a bit). On a side: It was challenging to get paper at U York in non-busy times, and no coloured pens. Nevertheless, here´s the idea.

* **Panel A** (model, Sascha) uses much the design of Fig 1 of Young, Frühholz & Schweinberger, 2020, TICS, but also incorporates strategic boxes considered by Lavan & McGettigan, 2023, Communications Psychology. For the latter, I picked Health and Impressions as the most relevant ones, and I think I´d rather avoid the cloud of the wider cognitive system which always strikes me as being somewhere between ridiculous and obvious but can be disputed.
* Panel A then further uses colour codes to sketch that naturalness can come it at various levels of processing, and with various references/norms in mind. The most fundamental level and (as per today´s discussion and the Roswandowitz 24 paper) perhaps to relate to low-level acoustic analysis and/or structural processing would be human likeness based. (We´d potentially also have the Charest et al. reference to back this point).
* I´ve picked 3 others, with 3 different colours RGB to relate to panel B, and I´ve picked accent and health (the latter particularly obvious) and emotions (if controversial, I´ll certainly not insist on this one, but to have 3 would be nice for aesthetic reasons under this idea). I also thought in Panel A perhaps we should sketch via dotted lines that we can imagine more to emerge in the future.
* **Panel B** (naturalness visualization relative to references, Tine) is extended to circular. At the center we borrow from additive mixture of colour, with a wide center of overlap (white) as a general prototype/reference that applies to all (human) voices, and individual specific prototypes/references that apply to healthy voices (green), “normal” accent (blue), and natural versions of a specific emotion as per example happy (red)
* Many other aspects of Tine´s sketch are worth preserving (I just could not paint this so nicely…
* On second look, and provided this leads somewhere: Perhaps one could rotate Panel B by 180 deg so that there is maximal spatial proximity between the human likeness aspects in Panel A and B
* Hope that is helpful, alternatively it was a try…

References

1. Charest, I., Pernet, C. R., Rousselet, G. A., Quinones, I., Latinus, M., Fillion-Bilodeau, S., Chartrand, J. P., & Belin, P. (2009). Electrophysiological evidence for an early processing of human voices. *Bmc Neuroscience*, *10: 127*, 1-11. ISI:000271324800001 (NOT IN FILE)