Our theoretical considerations on the processing of voice naturalness calls for the empirical investigation of the time-course and underlying brain mechanisms, relative to authenticity assessments but also in relation to other voice characteristics. Initial evidence suggests that voice naturalness affects the brain response as early as 200 ms past voice onset and interacts with the processing of vocal emotions (Duville, Lattner 2003 HBM). Comparably early effects have been found for authenticity assessments (Conde 2022, Kesilo 2021, Sarzedas 2024). Although interpretability of these findings is limited due to potential acoustic confounds, they suggest that naturalness and authenticity assessments are fast processes that form part of fundamental voice perception. However, electrophysiological insights directly comparing the time-course of naturalness and authenticity are elusive, as is their interplay with impressions of age, gender, or personality traits. A recent EEG study suggests that many first impressions formed from voices are highly intercorrelated (Lavan 2024), but for naturalness we are currently limited to behavioral data that point towards interactions with age, gender and emotion perception (Quellen). In a broad sense, naturalness impressions are always formed against the backdrop of a specific context, whether that refers to the micro-level of the voice itself, or the macro-level i.e. the properties of the interaction. Therefore, it could make a difference if the same voice is assessed in a human-human conversation or human-machine interaction.

Context