Human behavior is influenced by the perceived quality of objects and organisms that are encountered in our natural, social, and virtual environments. An important quality dimension concerns the perceived “naturalness”. Assessing naturalness has an evolutionary meaning, as it influences interactions, food choice, and social trust (Quelle). Naturalness, from a biological point of view, can be understood as the adaptive norm, with extreme deviations supposedly being rather “unnatural” instances (Quelle). Besides the biological context, the recent emergence of AI-generated digital and virtual contexts has brought human-machine interactions to everyday life, and therefore questions of naturalness to the forefront of scientific research.

A domain where features of (un)naturalness are of particular importance is the voice, because it is one of the prime channels for human communication (Quelle) and beyond, with synthetic voices nowadays often becoming the main carrier of communicative interactions, such as in customer service calls, gaming environments, or support platforms. When listening to voices, we form an instant impression about them (Quelle). Crucially, listeners seem to be very sensitive to (un-)natural voice features, which affects communicative quality. For **human voices**, consistent evidence from different speech-language pathologies shows that impairments in speech naturalness affect everyday interaction to a degree that can result in social isolation, reduced quality of life, and even depression (Stepp & Voitech 2019, Quellen). Similarly, deliberate acoustic manipulations and distortions disrupt effective communication (Quelle). For **synthetic voices**, one can hardly keep up with the rapid developments which make indefatigable efforts to resemble human vocal expression (Quelle). However, as of today, synthetic voices are consistently rated as less natural than human voices, which simultaneously affects perceived likeability, trustworthiness, and pleasantness (Quellen).

Given the widespread practical importance, it is crucial to put the role of voice naturalness into scientific focus. But although many recent studies provide useful empirical insights, we are currently looking at a rug rag rather than a research field. This has motivated us to take a step back and reflect on four problems in the present research: (a) conceptual underspecification, (b) inconsistent operationalization, (c) lack of exchange between research domains and (d) insufficient anchoring in voice perception theory. We argue that these problems have so far precluded a systematic understanding of vocal naturalness, impeded the visibility to a wider readership, made us overlook crucial research questions, and led to a divergence between theory and practice. In what follows, we will elaborate on each of these problems, before proposing concrete measures to address them, starting with the development of a concise conceptual framework for voice naturalness. To this end, we aim to provide a useful basis for systematic and theory-driven research on voice naturalness in the future.