Review/Opinion Paper

What‘s the status quo:

* Naturalness super relevant
* Different research efforts, mainly from applied fields (Voice distortions; synthesiz, computer interaction)
* NOT from cognition and basic research
* Not rooted in voice perception models and conceptualization
* That results in a several problems:
  + These fields don’t know about each other, don’t cite each other
  + Naturalness is not uniformly defined, nor operationalized
  + Research questions are tailored to the needs of applied field, while the basic understanding of naturalness is missing

What needs to change:

* Develop theory-based research questions
* Root them in theoretical models on voice perception
* Connect insights from different fields
* Structure the definition of voice naturalness
* Pull research across fields (in human and non-human voices)

Outstanding questions:

* Voice space forms during early childhood through exposure, how is this shaped now through exposure to artificial agents?
* Is “naturalness” in the sense of human-likeness desirable in the context of artificial voices?
* Naturalness != quality for artificial voices
* Naturalness reduced due to technical artifacts
* Methodological considerations: human vs. synthetic voices (which synthetic voice, how old) -> either trying to make them as human as possible (which we almost reached now). Just showing that synthetic voices are worse at something does not really take us anywhere
* Why human voices more efficient (Rodero 2023): closeness and emotional connection to human voices; or just less load on the perceptual system