**Children's Perceptions of Natural and Synthetic Voices**

Hannah Osborn
Emma Grace[1] and Jane Bickford[1]
[1] Flinders University

## Short Abstract

Synthetic voices, generated by speech-generating devices (SGDs), aim to replicate human speech, yet prior studies note their shortcomings, including a lack of expression and poor intelligibility, affecting users' communicative participation. Existing research predominantly involves adults and employs small sample sizes or single-case studies. Thus, little is known about how primary-school aged children perceive natural and synthetic voices, reflecting the need for further research. This current study employed a mixed-method design to explore how children perceive synthetic and natural voices in regard to six voice-related constructs. Fourteen participants aged between 5-12 years (girls n = 7, boys n =7) were recruited via social media, and purposive sampling. Two of these children were a child using a SGD and their sibling communication partner (SCP). The results from this dyad are presented separately as a case-study. Each child participated in a listening protocol where a natural voice and two synthetic voices including one that was matched for age and gender to the participant, and one unmatched were presented. The perceptions, revealed through rating scales and semi-structured interviews, were consider insignificant, however, they revealed participants perceived synthetic voices more negatively than natural voices. Conversely, the SCP favoured the unmatched synthetic voice, likely due to personal experiences. Perceived communicative impacts were reported. Despite limitations resulting in no statistical significance, including a small sample size, and variation within the rating scale, it is hoped that this study will inform future research and contribute to enhancing synthetic voice technology.

## Long Abstract

Synthetic voices are computer-generated voices within speech generating devices that are used to supplement and replicate human speech. Despite synthetic voices enabling positive communication outcomes, there is research that suggests the computer-based nature of synthetic voices can have a significant impact on identity, communicative successfulness, and participation. Previous studies have described synthetic voices to lack expression, and possess poor intelligibility, impacting a user's ability to participate successfully in conversations. The current research has come from small-sized participant samples, single-case studies and from the perspective of adults. Thus, little is known about how children perceive natural and synthetic voices, reflecting the need for further research. This current study employed a mixed-method design to combine and interpret elements of the quantitative and qualitative findings. It explored how children perceive synthetic and natural voices in regard to six voice-related constructs, including understanding, friendliness, emotion, participation, likability and similarity. A case study was included to explore the perceptions from a child that uses a speech

generating device and their sibling communication partner (SCP). The communicative participation impacts of synthetic voices were also explored. Two synthetic voices were used within the study including one that was matched for age and gender to the participant, and one that was unmatched for age nor gender. Fourteen participants aged between 5-12 years (girls n = 7, boys n =7) were recruited via social media platforms, such as the Facebook group, AGOSCI, which is a community group for individuals with CCN as well as their families and professionals who work with this population, and through purposive sampling, whereby, the social media posts were shared amongst the researcher's networks. Two of these children were case-study participants. The perceptions gathered through 5-point online rating scales and semi-structured interviews, revealed that the participants perceived both synthetic voices more negatively than the natural voice. Conversely, the sibling communication partner perceived the unmatched synthetic voice most positively. This is likely attributed to their personal experiences, and their familiarity with the synthetic voice, suggesting that frequent exposure to synthetic voices might influence perceptions. Additionally, communicative participation was perceived to be impacted whereby participants described that the limitations of the synthetic voice may influence successful communication. The case-study findings supported this. Whilst the findings from this current study align with findings from previous research, it is important to recognise that this study has some limitations, such as having a small sample size, and variation within the rating scale results This results in the findings being insignificant and are inconclusive retaining null hypotheses. While this is the case, it is hoped that the findings will help inform future studies and contribute to enhancing synthetic voice technology. Additionally, the findings from this present investigation reinforce the need for future research with a greater sample size to highlight the clinical and statistical significance and difference between the voices.