

知的障害特別支援学校における発話明瞭度評価の自動化に向けた基礎的検討

～音声認識（ASR）による発話明瞭度の評価～

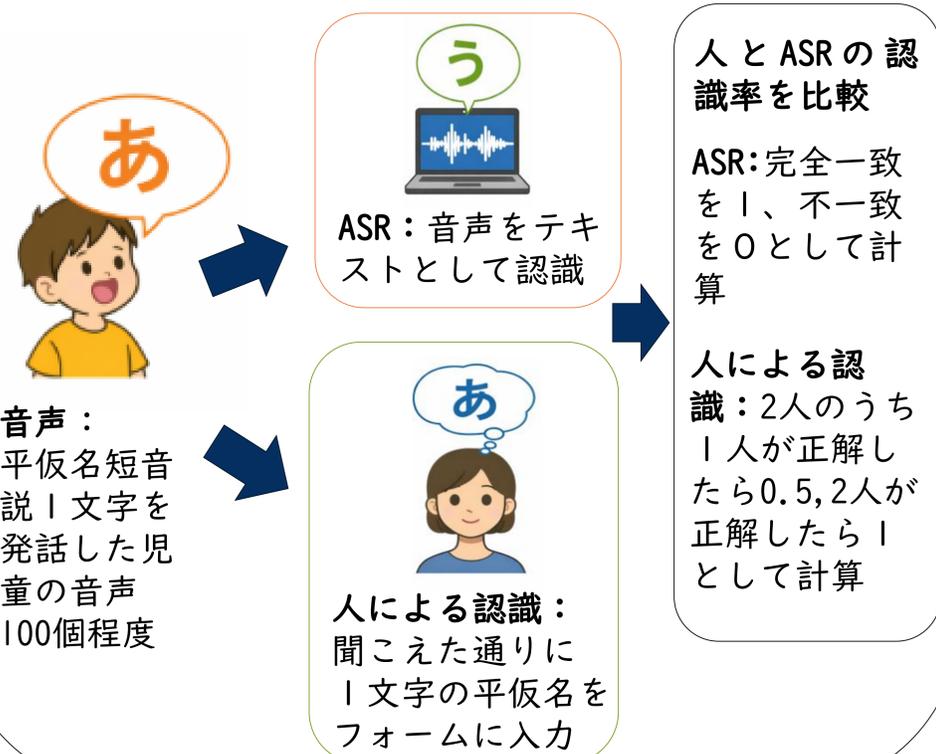
問題と目的

- 特別支援教育の現場では、児童の発話が「わかりやすくなった」と感じられることが多いが、明瞭度を客観的かつ簡便に評価する指標は十分に整備されていない。
- 本研究では、音声認識技術（ASR）を用いて、知的障害支援学級の児童の発話明瞭度を自動的に評価する方法を検討した。
- 基礎的検討として、ASRによる認識結果と人間評価者による明瞭度評価との対応関係を比較した。

方法

【対象】

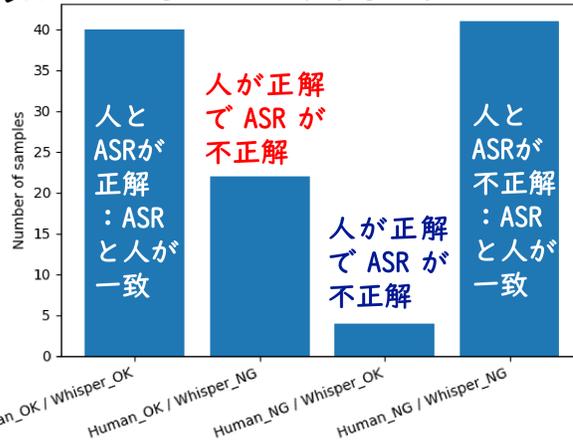
- 本校小学部高学年に所属する児童 1名
- 知的障害教育関連の専門家 2名
- 本研究は大阪教育大学倫理委員会の承認を受けて実施した。保護者に研究説明を行い、参加同意を得た（No.25139）。



結果

表1

Fig.4 Human-Whisper judgment patterns

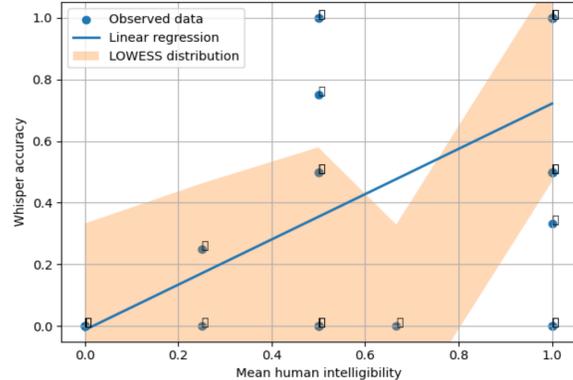


➤ ASRは摩擦音（さ、せ）や半母音（わ）の一部で高い認識を示した（ASRが正解で人が不正解）。

➤ 人は破擦音・摩擦音（ち、つ、し、す）などを認識しやすい（人が正解でASRが不正解）。

表2

Fig.5 Relationship between human intelligibility and Whisper accuracy



➤ 人による明瞭度が高い音ほど、ASRも正しく認識しやすい（人間と似ている傾向）。

考察

- 人間評価とASR（Whisper）では、音によって正誤傾向が異なった（表1, 2）。
- ASRは母音など長く安定した音を得意とし、ち・つなど短く不安定な音を不得意とする傾向がみられた（表1, 2）。
- 人間は音響的手がかりに加え、経験や学習に基づく補完により音を判断している可能性がある。これらの差は能力の優劣ではなく、判断に用いる手がかりの違いによると考えられる。
- ASRは発話明瞭度評価の補助的手法として有用であるが、人間評価との併用が不可欠である。

今後の展望

- 本研究は単音発話に限定しており、今後は文脈を含む発話への拡張と、評価者・対象数の増加による結果の安定性を検討する（Karbasi & Kolossa, 2022）。
- 相関に加えて類似度（Fleiss' kappa）で比較。
- 音声の音響的特徴量（基本周波数、持続時間、メル周波数ケプストラム係数）を分析し、認識率に影響する要因を明らかにする。
- ASR側の特徴量として、文字誤り率（Character Error Rate: CER）、認識結果の信頼度（confidence score）などを検討。