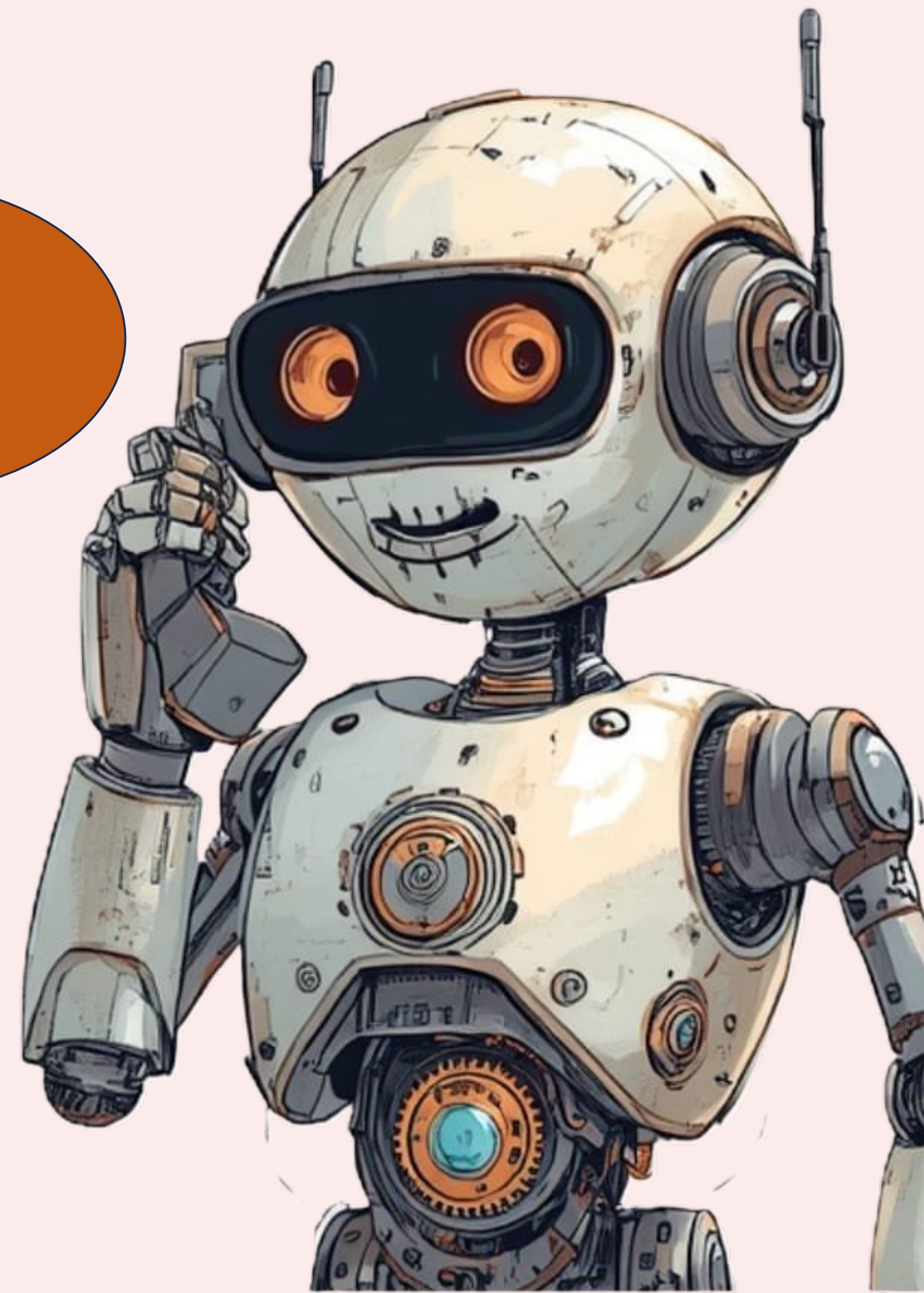




„Hallo, mit wem spreche ich denn?“ Die Rolle des Alters bei der Erkennung synthetischer Stimmen

von Nora Hibbe, Helen Jäschke, Thalia Kilimann, Julie Opitz & Sophia Schmidt
Leitung: Dr. Christine Nussbaum



Einleitung

- Synthetische Stimmen klingen zunehmend menschenähnlich und können oft kaum noch von menschlichen Stimmen unterschieden werden [1]
- Es ist jedoch unklar, welche Stimmerkmale die Diskriminationsfähigkeit beeinflussen [2]
- Hypothese:** Es gibt Unterschiede in der Diskriminationsleistung zwischen menschlichen und synthetischen Stimmen in Abhängigkeit vom **Alter** des **Sprechers** und des **Hörers**.

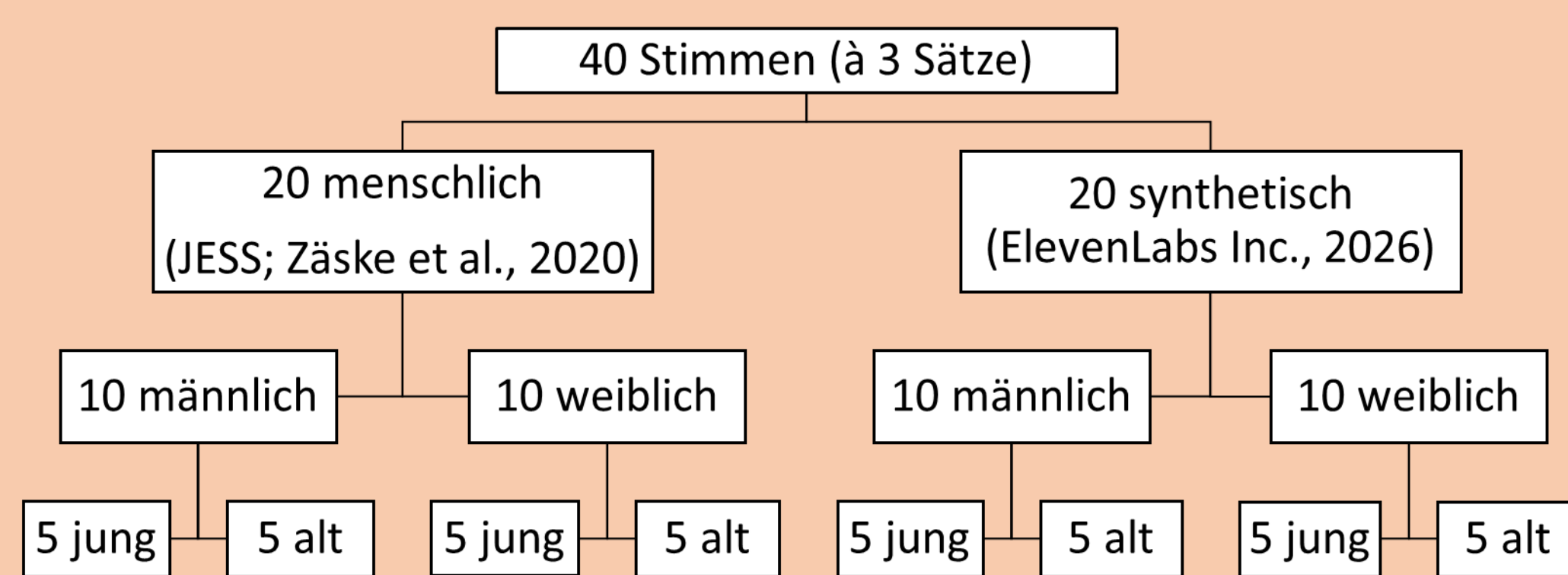
2.

Methoden

Stichprobe

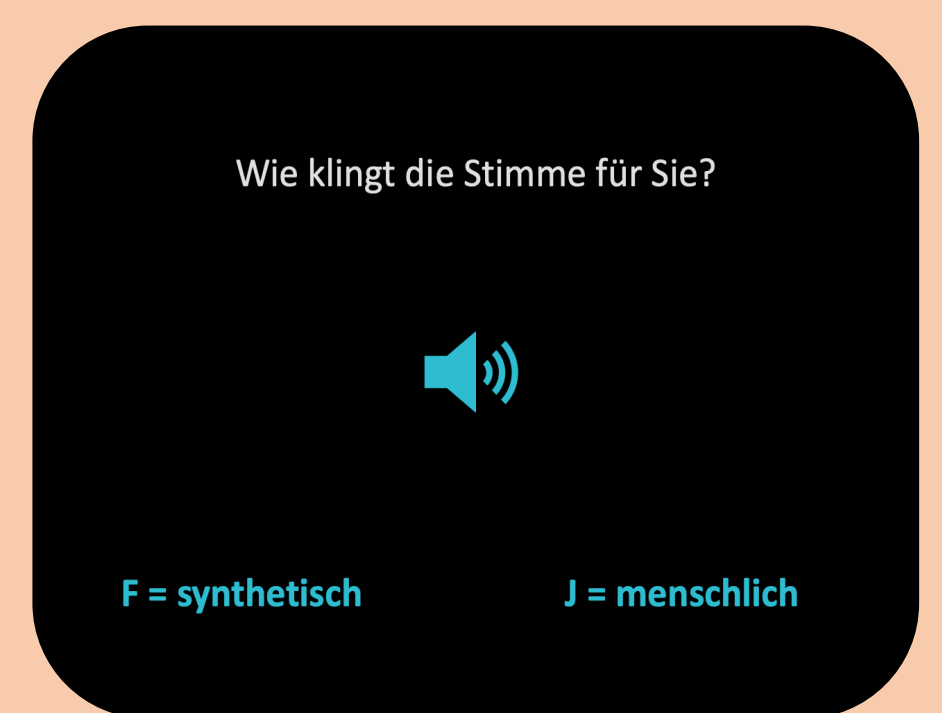
- 70 Personen** (38 männlich, 31 weiblich, 1 keine Angabe)
- Junge Gruppe:** 18 bis 32 Jahre ($n = 39$, $M = 23.56$ Jahre, $SD = 4.29$)
- Alte Gruppe:** 60 bis 80 Jahre ($n = 31$, $M = 66.51$ Jahre, $SD = 6.68$)

Stimuli: 2x2x2 Design



Durchführung

- Online-Erhebung über **PsyToolkit** [3]
- 4 Blöcke à 30 Stimuli**
- Entscheidung per Tastendruck:
menschliche vs. synthetische Stimme
- Am Ende Fragebogen zu **Kontakthäufigkeit** und **Akzeptanz** synthetischer Stimmen



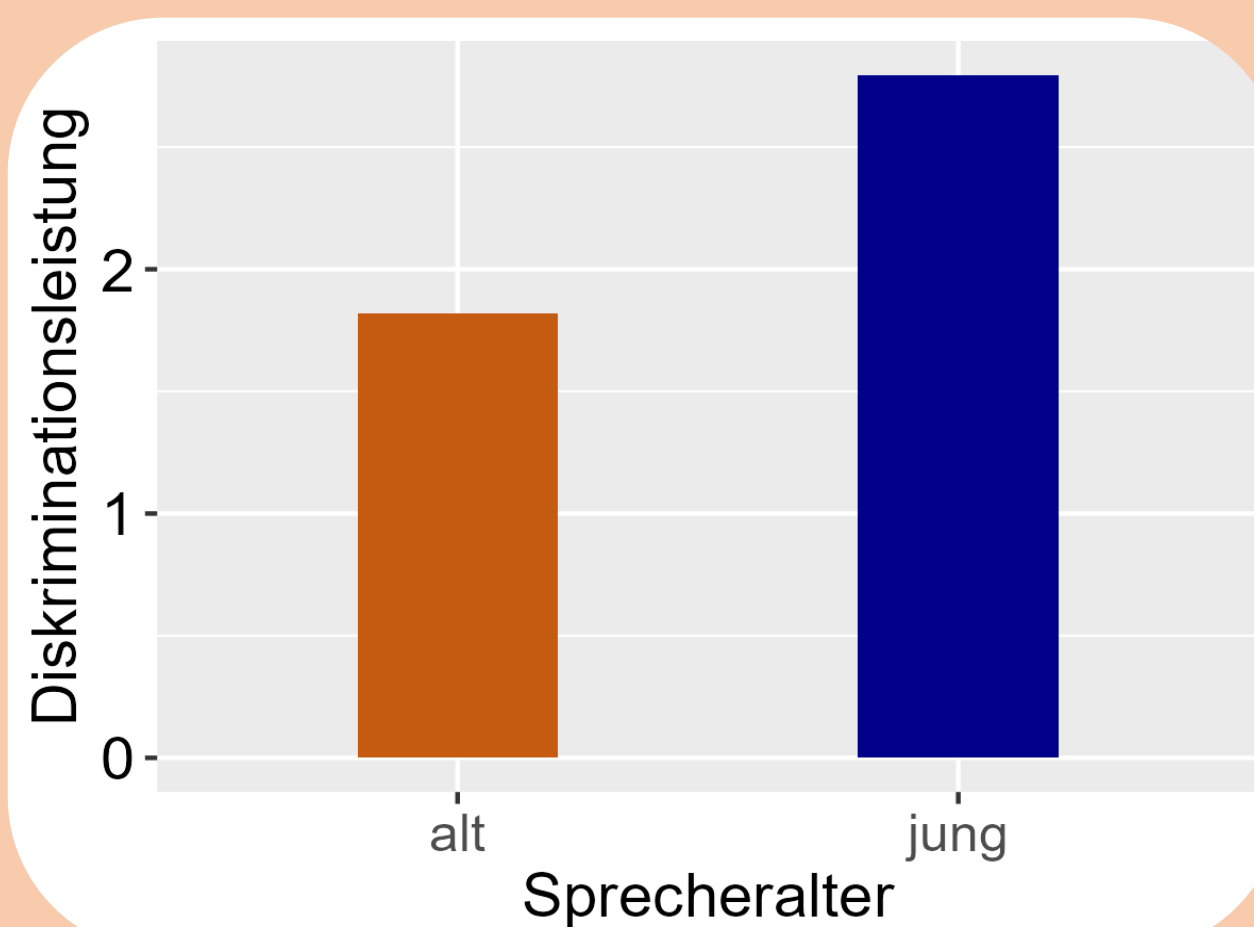
Auswertung

- AV: **Diskriminationsfähigkeit (d')**
- UVs: within: **Sprecheralter**
between: **Höreralter**
- Ziel: Unterschiede in der **Diskrimination** **menschlicher und synthetischer Stimmen** analysieren

Ergebnisse

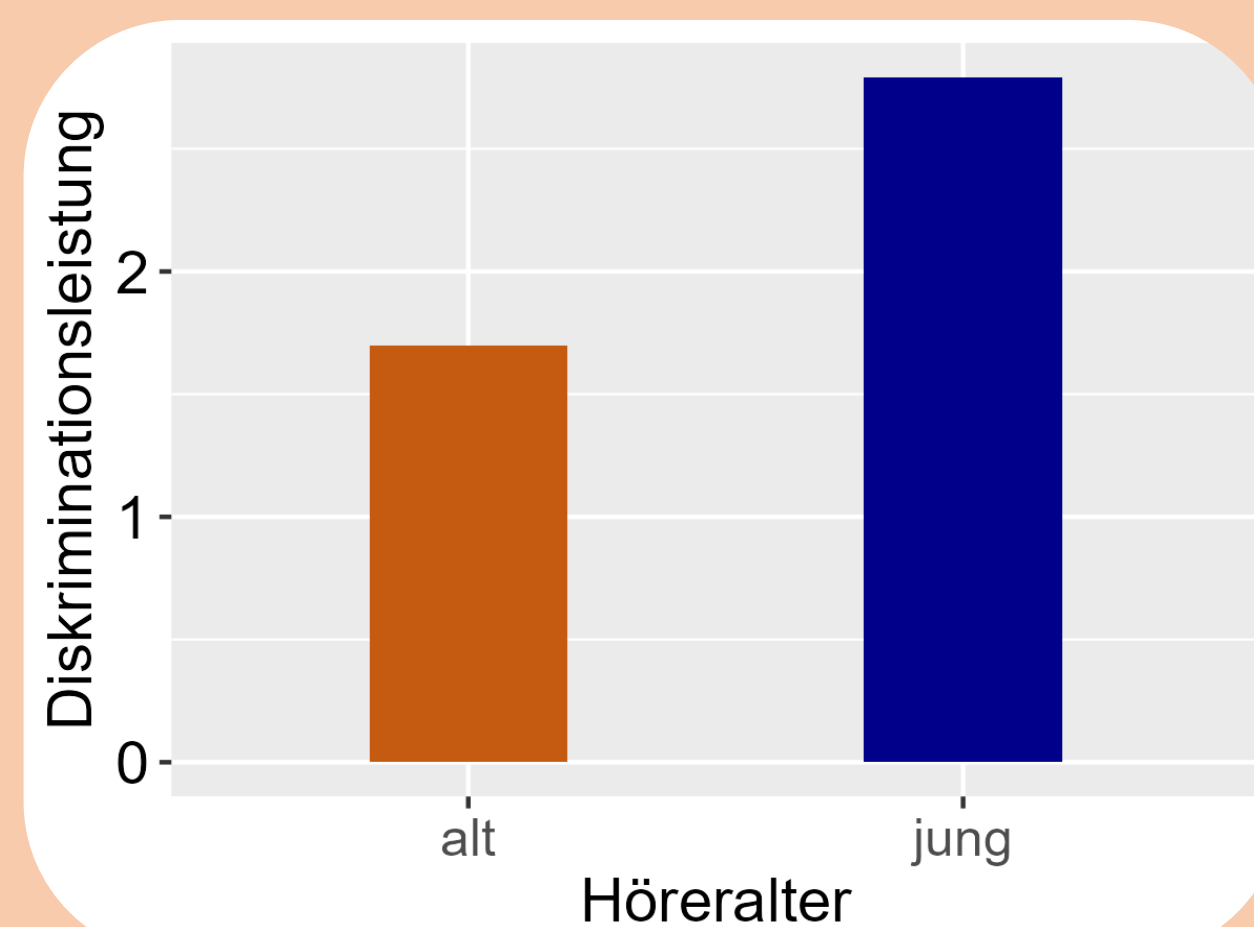
3.

Haupteffekt Sprecheralter



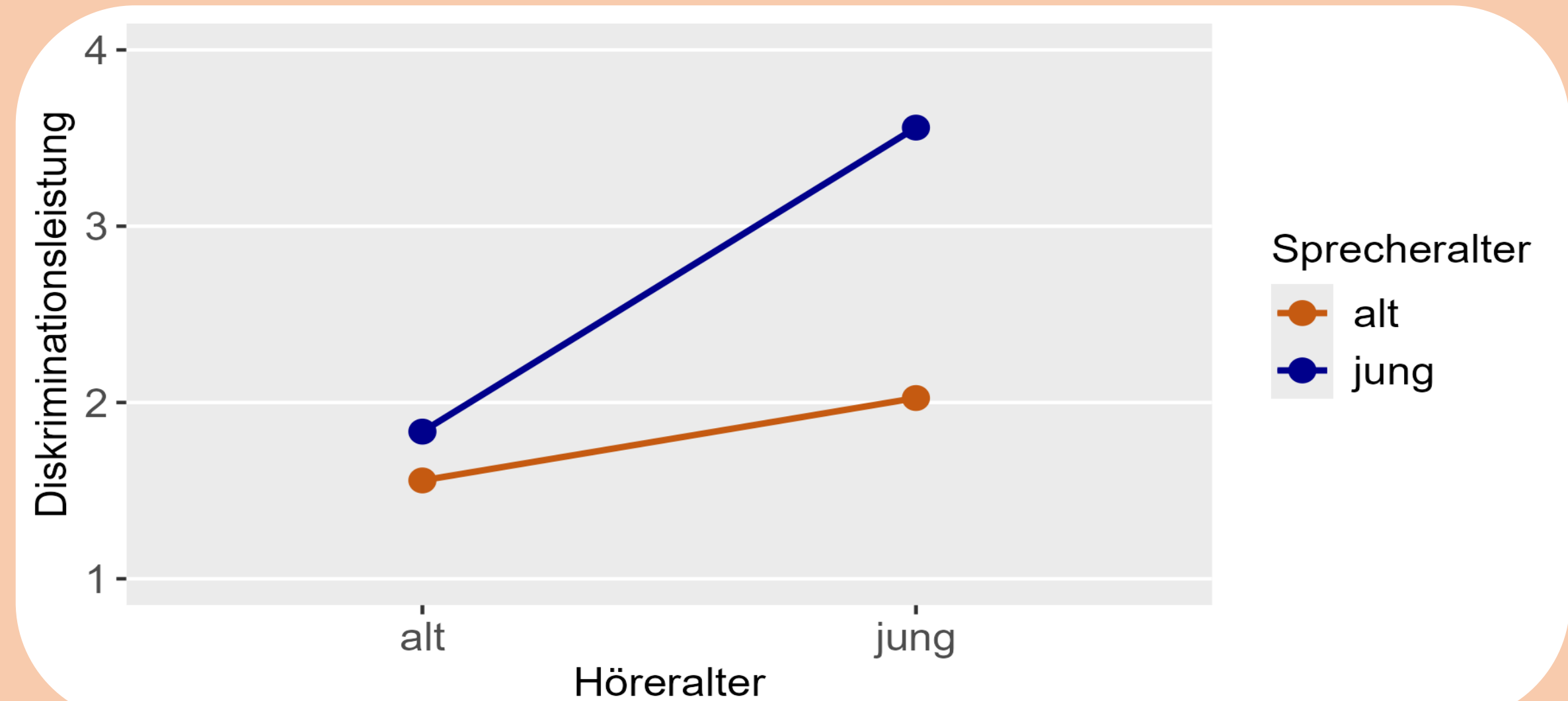
$F(1, 68) = 71.88$, $p < .001$,
 $\eta_p^2 = .514$

Haupteffekt Höreralter



$F(1, 68) = 37.45$, $p < .001$,
 $\eta_p^2 = .355$

Interaktionseffekt von Sprecher- & Höreralter



$F(1, 68) = 34.64$, $p < .001$, $\eta_p^2 = .337$
alte Hörer: $t(30) = -2.12$, $p = .042$
junge Hörer: $t(38) = -9.62$, $p < .001$

4.

Diskussion

Junge Hörer sind besser in der Diskrimination:

- Haben junge Menschen mehr Übung in Diskrimination durch häufigeren Kontakt zu synthetischen Stimmen im Alltag (Medien, Sprachassistenten etc.)?

Junge Stimmen werden besser diskriminiert:

- Höhere Medienpräsenz von jungen Stimmen?

Interaktion - Junge Menschen sind besonders bei der Diskrimination junger Stimmen besser:

- Gibt es einen gemeinsamen Effekt von mehr Kontakt mit synthetischen Stimmen & größerer Häufigkeit von jungen synthetischen Stimmen?

Limitationen:

- Realistischere junge synthetische Stimmen durch mehr Trainingsdaten in Elevenlabs?
- Selektive Überarbeitung der jungen synthetischen Stimmen als mögliche Verzerrung?

