Estimating Type of Print Exposure across Aging through Author Production

Mengyang Qiu (Presenter)
Nichol Castro (Buffalo), Brendan Johns (McGill)

TRENT UNIVERSITY

Background & Motivation

- Individual differences in language exposure crucial for cognitive research
- Print exposure: important source of knowledge accumulation
- Challenges in measuring print exposure:
  - Self-reports: subjective, potentially biased
  - Author Recognition Test (ART): difficulties in item selection
- Need for objective measures of print exposure type/genre

Study Objectives

- Introduce novel approach to quantify individual print exposure, Author Production Test (APT)
- Integrate distributional semantics with APT to objectively estimate genre preferences
- Validate approach across age groups and reading frequencies

Methodology Overview

- Author Production Test (APT): participants list fiction authors they've read
- Universal Sentence Encoder (USE): generate vector representations of authors
- Construct:
  - Participant vectors: aggregated author vectors produced in APT
  - Genre vectors: representative characteristics of each literary genre
- Estimate genre preferences: cosine similarity between participant and genre vectors

Distributional Estimation of Genre Preferences

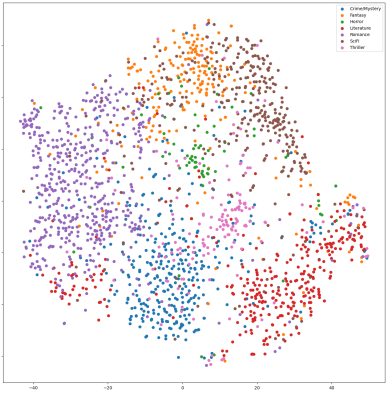
- Genre vectors:
  - Medoid approach: most centrally located author vector in each genre
  - Captures 'typical' features of authors within a genre
- Participant vectors:
  - Sum of author vectors produced in APT
  - Unit normalized to preserve distributional information
- Objective genre preference: cosine similarity between participant and genre vectors
  - $\text{cosine}(\mathbf{A}, \mathbf{B}) = \frac{\mathbf{A} \cdot \mathbf{B}}{||\mathbf{A}|| \times ||\mathbf{B}||}$

Data Collection

- Participants:
  - 294 older adults (50-70 years)
  - 309 younger adults (18-30 years)
- Measures:
  - Language background survey (including reading frequency)
  - Self-reported genre preferences (7-point Likert scale)
  - Author Production Test (APT)

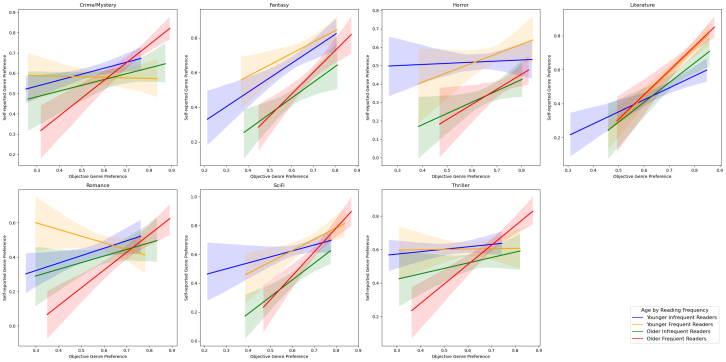
Analysis & Results - Genre Vectors

- Authors within genres cluster together



Analysis & Results - Correlations

- Stronger correlations for older frequent readers across all genres
- Suggests effectiveness in capturing print exposure, especially for experienced readers



Key Findings

- Distributional approach effectively captures individual print exposure
- Stronger correlation between objective and subjective measures in older frequent readers
- Importance of objective measures to complement potentially biased self-reports
- Interaction between amount and type of print exposure

Implications & Future Directions

- Potential for developing personalized models of lexical behavior
- Applications in education and neuropsychology:
  - Tailoring teaching materials
  - Optimizing neuropsychological evaluations