Real-Time Evaluations of Accent and Professional Competence

Erez Levon, Devyani Sharma, Yang Ye, Amanda Cardoso & Dominic Watt
“The moment an Englishman opens his mouth, another Englishman despises him.”

(George Bernard Shaw, Pygmalion, 1916)

- Long-standing patterns of **inequality** in professional hiring in the UK
  (Ashley et al. 2015)

- Accent is a **key signal** of social background and can impede access to elite professions
  (Giles et al. 1975; Kalin et al. 1980; Giles et al. 1981; Alemoru 2015; Roberts et al. 1992)

- Differences in perceptual evaluation of accents widely reported
  - Historical studies of RP and other dialects in the UK
    (Giles 1971; Bishop & Coupland 2005)
  - Recent nationwide survey of attitudes 5 major British accents
    (Levon et al. 2019)
Phases of Project

Verbal guise survey with **general UK public** and with **legal professionals** (with latter quality of response is manipulated)

Testing influence of “**accentedness**” ratings

Examining perceptual evaluations in **real-time**

Designing and testing different anti-bias **interventions**
Phases of Project

Verbal guise survey with **general UK public** and with **legal professionals** (with latter quality of response is manipulated)

Testing influence of “**accentedness**” ratings

Examining perceptual evaluations in **real-time**

Designing and testing different anti-bias **interventions**
How do attitudes form **during** the time course of listening?

How do these **compare** to commonly-used retrospective ratings?

What does real-time analysis **add** to our understanding of accent perception?
Methods — Stimuli

- Real-time rating of verbal guise stimuli (N=160)
- Listeners evaluated interview performance of “candidates” for trainee solicitor position at a corporate law firm
- Stimuli included one native speaker each for 2 English accents:
  - **Received Pronunciation** (RP)
    - Traditional standard, prestige accent
  - **Multicultural London English** (MLE)
    - New working class, multi-ethnic urban accent
Presence of bias in Nationwide Survey (retrospective ratings)
Methods — Stimuli

- Real-time rating of verbal guise stimuli (N=160)
- Listeners evaluated interview performance of “candidates” for trainee solicitor position at a corporate law firm
- Stimuli included one native speaker each for 2 English accents:
  - Received Pronunciation (RP) — Traditional standard, prestige accent
  - Multicultural London English (MLE) — New working class, multi-ethnic urban accent
- Stimuli consisted of identical responses to two common types of job interview question:
  - Expert — “Explain the difference between contract and tort.” (requires legal expertise)
  - Non-expert — “Why do you want to be a lawyer?”
- Counter-balanced for speaker order and question order
Methods — Procedure

- Listeners heard a pair of responses from one candidate, then the other
- Listeners provided two types of evaluation
  - **Real-time evaluation** of candidate’s performance during each recording via a graphical sliding scale on computer screen.
  
  ![Graphical sliding scale example](image)

  - **Retrospective assessment** of the quality of candidate’s response after each recording
    
    How would you rate the overall quality of the candidate’s answer?
    Does the candidate’s answer show expert knowledge?

- Listeners provided information about their social and linguistic background at the end
Methods — Modelling

- Listener evaluations modelled at 1 second intervals across time-course of stimuli
- GAMMs (mcgv package in R 3.6.1) examined influence of:
  - **Linguistic factors** — accent, Dialect Density (difference from RP, 3 second window)
  - **Social/Psychological factors** — age, Motivation to Control Prejudiced Response
  - **Contextual factors** — question order (expert/non-expert first), accent order (RP/MLE first)
- Separate models built for expert and non-expert questions
- Models manually stepped down from maximal models including all parametric terms, smooth terms (by relative time) and factor-smooth interactions
- Models also included random smooths for participants
**Hypothesis 1  Absence of bias**
There may be no influence of accent on listeners’ perceptions of competence in a job interview response — Fluctuations in ratings may correspond to other factors, e.g. content
→ ratings of MLE and RP will not differ over the time course of listening

**Hypothesis 2  Early broad categorisation**
Listeners may make an early categorisation of a speaker based on general recognition of the accent (Levon 2014; Pharao et al. 2014)
→ ratings of MLE and RP will diverge early in the time course of listening

**Hypothesis 3  Accent-specific shibboleths**
Listeners’ perceptions of competence in a job interview may be sensitive to specific accent features or clusters of features in an utterance (Montgomery & Moore 2018)
→ ratings of MLE and RP will fluctuate differently throughout the time course of listening
Results — General accent effect

EXPERT QUESTION

*Explain the difference between contract and tort.*

- Identical content
- But accents rated differently
- ✗ Hypothesis 1
- MLE downrated early
- Looks like early categorisation (Hypothesis 2)
Results — Expert vs. Non-Expert

EXPERT QUESTION
Explain the difference between contract and tort.

NON-EXPERT QUESTION
Why do you want to be a lawyer?
• Dialect density fluctuates
• But no clear correspondence between DD and rating

✗ Hypothesis 3

• What about Hypothesis 2?
Results — Age

- Younger listeners rate MLE higher than older, but same pattern of late rise
- For RP, all listeners show early rise
- Older listeners have a steeper slope as they get to a higher endpoint

(EXPERT QUESTION)
Results — Motivation to Control Prejudice

- Still the same late/early contrast between MLE and RP
- MCPR does not have significant affect for RP
- For MLE, higher motivation improve ratings, both steeper and a bit sooner.
Results — Expert vs. Non-Expert

EXPERT QUESTION

NON-EXPERT QUESTION
Summary

- Different real-time response profiles for RP and MLE
  - **standard variety**: early fast rise, continued increase across response
  - **non-standard variety**: long low plateau, rapid rise toward end

- Evidence for early classification followed by updating of priors (content mitigates stereotype, as do age and Motivation to Control Prejudice)

- Attitudes evolve after early classification (retrospective ratings correlate with endpoint values)

→ Non-standard varieties have enhanced **burden of proof** (trajectory bias) as compared to standard varieties

→ Real-time examination provides additional insight into the process of online attitude formation and the potential role of stereotypes in real-time processing
Thank You!

www.accentbiasbritain.org
accentbiasbritain@qmul.ac.uk
Appendix — Exposure to Expertise

- When the expert answer is heard first, we see the standard trajectory difference in accent ratings.

- When the expert answer follows the non-expert answer, MLE receives an ‘expertise’ boost that reduces the primary pattern.

(EXPERT QUESTION)
Appendix — Order Expert vs. Non-Expert

**EXPERT QUESTION**

**NON-EXPERT QUESTION**