

Assessing Survey Questions through a Machine Learning Pipeline: Emotions and Paralinguistic Behaviors

Anil Battalahalli, Ting Yan, Hanyu Sun, *Westat*

10/16/2023

Background

- › Paralinguistic utterances can serve as cues about the speaker's current state of mind including their understanding of a question and their affective state or mode (e.g., Brown 1977, Hancock 2004).
- › Previous research has studied the paralinguistic utterances to explore the interviewer-respondent interaction, such as, disfluencies (e.g., "uh" or "um") and pauses:
 - The presence of such behaviors can be used to flag low-quality responses (e.g., Schaeffer and Dykema, 2011; Schober, Conrad, Dijkstra, and Ongena, 2012; Min, Schaeffer, Garbarski, and Dykema, 2020).

Research Question

- › RQ: As exploratory research, can we use paralinguistic utterances and respondent emotions for question assessment?
 - Respondent behaviors:
 - Long pause: Respondent may have trouble understand or answer the question
 - Pitch variation: Sharp increase in pitch may indicate angry or upset.
 - Positive emotions: Good flow between the interviewer and the respondent
 - Interviewer-respondent interaction:
 - Overlapping speech: Respondent may interrupt the interviewer's reading of the question to provide a response or ask for clarification

Data and Methods (1)

- › 20 questions selected from a large-scale cross-sectional study of a nationally representative sample:
 - 479 question-answer recordings from 53 cases
 - 13 closed questions, 7 open-ended questions
 - 6 single choice questions, 7 multiple choice questions
 - 9 questions with showcards
- › Expert review: Difficult vs. not difficult questions

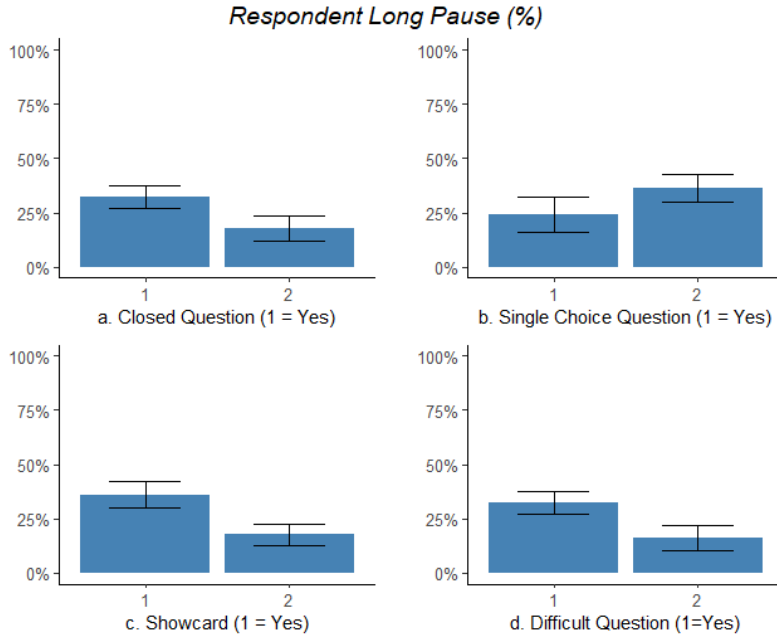
Data and Methods (2)

- › Automated processing of recordings through a machine-learning audio pipeline.
- › Respondent behaviors:
 - Long pause immediately after the interviewer asked the question
 - Pitch variation
 - Positive emotions
- › Interviewer-respondent interaction:
 - Overlapping speech

Data and Methods (4)

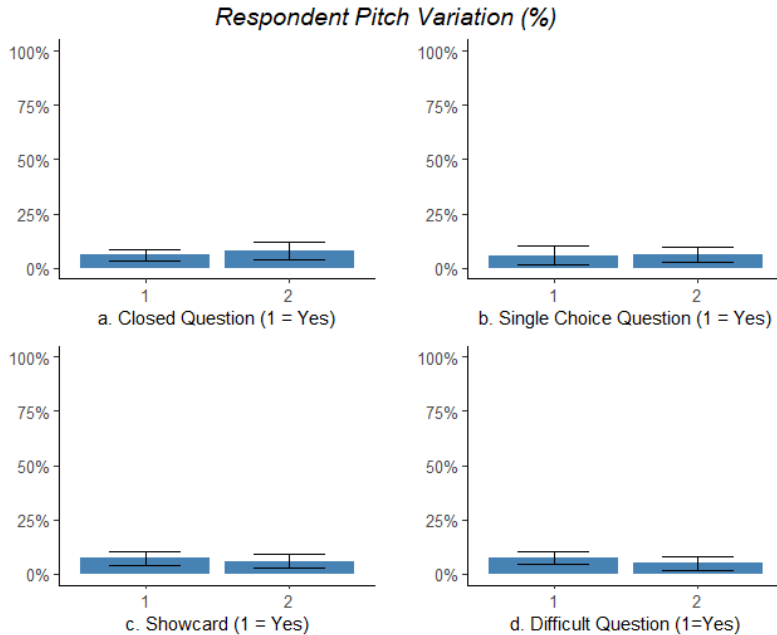
- › Trained models using the audio-only files from the RAVDESS (Livingstone and Russo 2018) to predict emotions
 - Extracted acoustic parameters using openSMILE (Eyben et al. 2013 and 2015).
- › Applied the final model to the recordings to predict the respondent's emotions (positive, neutral, negative) at the turn immediately after the interviewer read the question

Results: Respondent Long Pause (%)



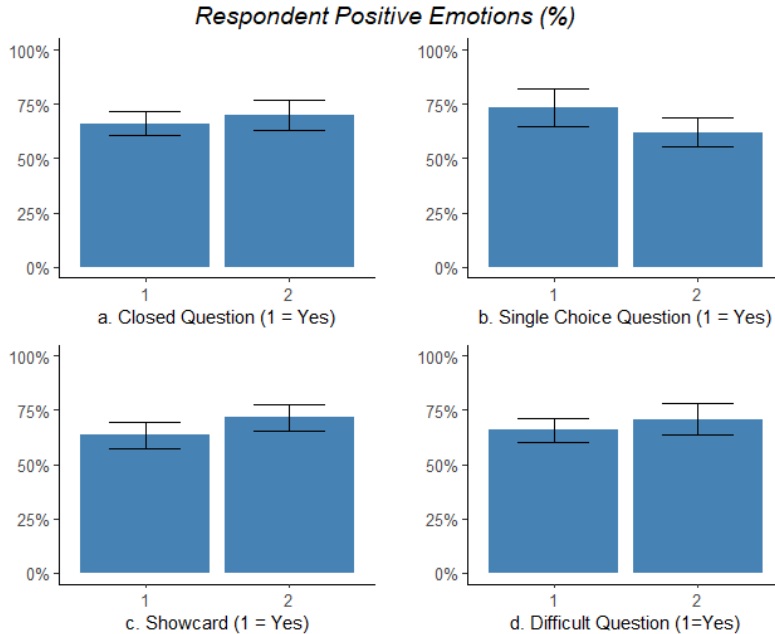
- › Multiple-choice questions, questions with showcard, and difficult questions have higher percentage of respondent long pause.
- › Closed questions have a higher percentage of respondent long pause as compared to open-ended questions:
 - 84.0% of the closed questions are considered difficult by expert review

Results: Respondent Pitch Variation (%)



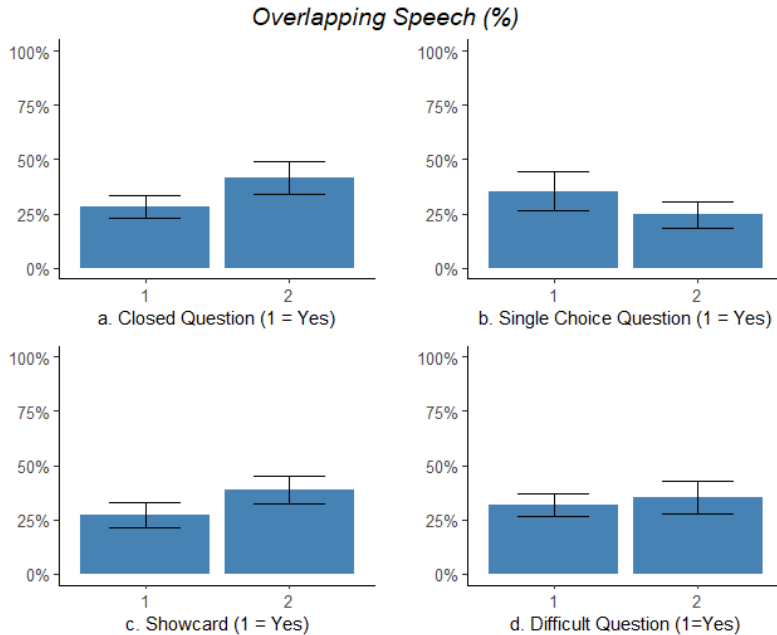
- › Not much difference between the percentages of respondent pitch variation by question characteristics:
 - Being angry or upset can be rare

Results: Respondent Positive Emotions (%)



- › Slightly higher percentages of respondent positive emotions for single choice questions, questions without showcard, and not difficult questions.
- › A slightly higher percentage of respondent positive emotions for open-ended questions as compared to closed-questions:
 - 84.0% of the closed questions are considered difficult by expert review

Results: Overlapping Speech (%)



› Open-ended questions, single choice questions, questions without showcard, and not difficult questions have higher percentages of overlapping speech:

- Respondents may interrupt the interviewer's reading of the question to provide a response

Conclusions and Discussions

- › The findings suggest that paralinguistic utterances and emotions can be used for question assessment with varied performance.
- › Future work:
 - Validate these measures with conventional behavior coding
 - Improve models with results of the conventional behavior coding
 - Understand the relationship between paralinguistic utterances and emotions and interviewer-respondent interaction
 - Derive a composite score to rank questions on difficulty/issues

Thank You

Anil Battalahalli (anilbattalahalli@westat.com)

Ting Yan (tingyan@westat.com)

Hanyu Sun (hanyusun@westat.com)