Effects of Speech Rate, Pitch, and Pausing on Survey Participation Decisions

José Benkí,¹ Jessica Broome,¹ Frederick Conrad,¹,² Robert Groves,³ and Frauke Kreuter²

¹ University of Michigan
² University of Maryland
³ US Census Bureau

Acknowledgements

• Support:
  – NSF Grants # SES-0819734 and # SES-0819725;
  – Survey Research Center, University of Michigan
  – Dept. of Communicative Sciences & Disorders, Michigan State University

• Coding, Transcribing, Acoustic Analysis, Sampling
  – Pete Batra, Rachel Benner, Kelly Franckowiak, Haley Gu, Ben Jarvi, Emily Kordupel, Peter Kotvis, Abby Lincoln, Lacie Linstrom, Melissa Littlefield, Daniela Lopez, Colleen McClain, Joe Matuzak, Colleen McCarty, Patty Maher, Gabe Moss, Kirsten Mull, Danny Nielsen, Dana Perkins, Fernando Pacheco, Danielle Popielarz, Christine Sheffler, Amanda Tatro, Dave Vanette, and Dylan Vollans
Voice, language, and participation

• Telephone interviewers’ success obtaining interviews is due, at least in part, to what they communicate over the phone
  – How they Sound -- vocal attributes
  – How and What they Speak -- manner and content of speech
  – How they Interact with potential respondents

• Over their careers, some Iwers are more and others less successful; this strongly implies that differences in their verbal attributes play an important part in outcomes

• What are the verbal attributes of Iwers that affect participation decisions and how do they have their effect?

• How do verbal attributes of Iwers interact with the verbal attributes of potential respondents in affecting participation?
Example Survey Introductions

Invitation 1: male pitch variation

Invitation 2: fluent but script-like

Invitation 3: disfluency

Invitation 4: noticeable pause

Note that we are focusing in this paper in the invitation and decision to participate, not on the interview itself.
Literature is Mixed About Attributes of “Successful” Iwers

- Lower pitch (Oksenberg, Coleman & Cannell, 1986; Oksenberg & Cannell, 1988)
- Higher pitch (Sharf & Lehman, 1984; Groves, O’Hare, Gould-Smith, Benki, Maher 2008)
- More variation in pitch (Sharf & Lehman, 1984; Oksenberg, Coleman & Cannell, 1986)
- Non-linear effect of pitch variation (Steinkopf, Bauer, & Best, 2010)
- Falling intonation (Oksenberg & Cannell, 1988)
- Rising intonation (Groves, et al. 2008)
- Greater fluency (Van der Vaart, Ongena, Hoogendoorn, & Dijkstra, 2005)
- Less scripted, i.e., less fluent (Groves, et al. 2008)
Our Project

• Examines impact of *Iwers’* voice, speech and interaction with phone answerer on decision to
  – Participate (*Agree*)
  – Refuse to Participate (*Refuse*)
  – Defer (*Scheduled Callback*)

• Measures:
  – Transcription and turn marking
  – Acoustic (e.g., pitch or *f0*, speech rate, vowel quality)
  – Paralinguistic (e.g., fillers, i.e., *uh, um*)
  – Move coding (e.g., self-identification, duration question)
  – Global ratings: (e.g., animation, similarity of *I* and *A* accent)

• 1380 audio recorded telephone introductions from 5 studies at University of Michigan SRC

• 100 interviewers
Multilevel Data Set

- Interviewer is cross-classified with study.
- Interviewer is cross-classified with cases/samples.

1380 contacts, transcribed and coded

- Each level is completely nested in the subsequent upper level.
Transcription and acoustic analysis

Hello my name is Bill. I’m calling from the University of Michigan in Ann Arbor and we’re doing a nationwide study on the economy. We just want to include some questions on other countries the world and a few other things.
Today’s Talk

• Three phenomena in interactive speech, data collected at the *turn* level:
  – Speech rate
  – Pitch
  – Pauses

• Outcomes:
  – 263 Agrees
  – 333 Refusals
  – 537 Scheduled Callbacks
  – 233 Hangups, 14 Other

• (Individual contacts, not particular *Iwers*)
Hypotheses: Speech rate

• *Faster speech rates* are associated with credibility and authority (e.g., Miller et al., 1976; Buller & Aune, 1988)

• We hypothesize that participation will increase when interviewers speak more quickly.
Hypotheses: Pitch

• **Mean pitch** is associated with vocal attractiveness for both sexes but in different directions (Collings & Missing, 2003; Hughes et al., 2008).

• **Pitch variation** is associated with liveliness and animation (Hincks 2005).

• Mean pitch is predicted to be be **positively** correlated with participation for contacts with **female** interviewers but **negatively** correlated for contacts with **male** interviewers.

• Pitch variation is predicted to be positively correlated with participation for all contacts.
Hypotheses: Pauses

• Previously we reported that a moderate rate of fillers (e.g., um or er) is associated with higher participation rates than either zero or high filler rates (Conrad et al., 2010).

• We predict that a moderate pause rate will also be associated with the highest participation rates.

• Invitations with moderate filler and pause rates sound neither scripted nor overly disfluent.
Results: Speech rate

- Interviewer speech rates of 3.5 words per second (WPS) are most successful.
- Participation falls off in the high and low speech rate quintiles.
Results: Mean pitch

- 374 male *Iwer* contacts: mean pitch of agrees was 13 Hz lower than in refusals.

- 768 female *Iwer* contacts: mean pitch of agrees was 7.5 Hz lower, opposite of our prediction.

All contacts also include 193 contacts with no interviewer sex information.

Error bars are standard errors.

Values are for the first 13 turns of each contact, the average length of a refusal.
Results: Pitch variation

- Average $F_0$SDQ for the female /werk contacts is 2.8 Hz higher in agrees.
- Average $F_0$SDQ for the male /werk contacts is 2.7 Hz lower in agrees, opposite our prediction.

$F_0$SDQ is the range of middle third of the distribution of pitch values in a turn.
$F_0$SDQ is correlated with overall pitch.
Results: Pauses

- 40% of contacts had no pauses in the first 13 turns (PPT=0).
- Contacts with some pauses, up to a little less than 1 pause every other turn, were most successful.
- Contacts with greater pause rates were less successful.
Conclusions

• The way telephone Iwers speak and interact when inviting household members to be interviewed is related to success of contact
  – At least in this corpus
  – Not just content (i.e., what they speak) but style (i.e., how they speak)

• **Speaking Rate:** Iwers are most successful at recruitment when they speak around 3.5 wps, and are least successful when speaking less than 3.0 wps.

• Compare with Cannell’s 2 wps recommendation for interviews.
Conclusions: Pitch

• Agrees with male Iwers show significantly lower mean pitch than refusals, consistent with attractiveness, but pitch variation was slightly hurtful for participation.
  
  Inappropriate interviewer intonation patterns, or perhaps sing-song deliveries while adhering closely to a script.
  High pitch variation conflicts with attractiveness for male talkers.

• Agrees with female Iwers have slightly lower mean pitch than refusals, inconsistent with attractiveness prediction, but pitch variation was found to be slightly helpful.
  Attractiveness is more weakly associated with mean pitch in females than in males.
  Females are generally more flexible and adaptable speakers.
Conclusions: Pausing

• Contacts with no pauses may sound overly fluent and perhaps scripted, but excessive pausing may be perceived as overly disfluent.
• Excessive pausing may also be symptomatic of an interviewer’s effort to salvage a call going badly.
• A moderate amount of pausing appears to be beneficial in recruitment.
  – Pauses may be perceived as part of a tailored invitation (Groves & Couper, 1998).
  – Pauses may be helpful in establishing a conversational context or rapport instead of a persuasive or sales context.
Some general conclusions

- Our research methodology is to analyze objective measures in relatively large numbers of actual contacts.
- This observational approach facilitates the detection of non-monotonic relationships.
- Our approach is complementary to others: Studies that examine relatively fewer contacts in great detail, often with more subjective measures. Controlled experimental methodologies that can more directly assess causation.
Next Steps

• Additional Measures:
  – Multivariate models incorporating nonmonotonic relationships and random interviewer effects
  – Subjective ratings like animation
  – Content like apologies, reference to incentive and previous contact
  – Response propensity of the case

• Compare speech of most and least successful Iwers
  – All measures
  – Convergence: variation in Iwer speech and vocal attributes on the basis of particular respondents
Practical Implications

• *Iwers* can potentially be trained to:
  – Avoid scripted delivery and inappropriate intonation
  – Avoid speaking too slowly
  – Strategic short pauses

• But there are individual differences in *Iwers’* ability to attend to both what they say and how they speak (i.e., overall pitch)

• This project provides a tighter connection between interviewer research and practice
Thank You
Sampling Contacts

All contacts from 5 studies

100 interviewers

40 or all cases per / with first contact agree/other

40 or all cases per / first contact refuse

Up to 80 cases per interviewer

All contacts from cases

(27/100 l/s worked on >1 study)
Coding at the move level
Agrees by Interviewer Filler Rate

Fillers per 100 words
Conrad et al. (2010)

All adjacent points differ reliably except for quintiles 2 and 3
Contacts from 5 Studies are Included:

- Gujarati (240)
- National Study on Medical Decisions (358)
- Interests of the General Public (56)
- Mississippi Community Study (20)
- Survey of Consumer Attitudes (706)

Total # contacts = 1380
<table>
<thead>
<tr>
<th>Category</th>
<th>Duration (Mean ± Standard Error)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agree: up to turn 13</td>
<td>8.95 (.25)</td>
</tr>
<tr>
<td>Agree: turn 14+</td>
<td>9.17 (.29)</td>
</tr>
<tr>
<td>Refuse: up to turn 13</td>
<td>6.43 (.18)</td>
</tr>
<tr>
<td>Refuse: turn 14+</td>
<td>4.63 (.29)</td>
</tr>
<tr>
<td>SCB: up to turn 13</td>
<td>6.89 (.14)</td>
</tr>
<tr>
<td>SCB: turn 14+</td>
<td>4.83 (.10)</td>
</tr>
</tbody>
</table>

Iwer turn duration in sec
Example Moves

**Interviewer:**
- **d:** description of survey (This survey helps us understand what people think about the US economy)
- **f:** follow-up (We sent you a letter)
- **g:** formal greeting
- **i:** self-identification
- **l:** apology
- **m:** incentive related
- **n:** expression of gratitude
- **p:** persuasive statement (Your opinions are very important to us.)
- **r:** procedural information (This call may be recorded for quality control)
- **s:** scheduling statement (Is there a better time?)
Example Moves

**Answerer:**
- U: Backchannel
- 2: Expression of interest
- 4: Request clarification
- 5: Suspicion or misunderstanding of purpose (I don’t want to buy anything, take me off your list)
- 6: Past experience with research (I did a survey last week)
- 7: UM related
- 8: Expression of disinterest
- A: Statement of ambivalence
- D: Duration question (How long will this take)
- F: Follow-up comment (My wife said you called)
Analysis

• At 6 levels:
  – Study: overall response rate, average length of interview
  – Case: number of calls/ contacts, eventual outcome
  – Contact: day and time, length, disfluency rates, disfluency convergence between interviewer and answerer, overspeech, patterns of moves, outcome
  – Interviewer: variance across contacts
  – Turn: linguistic measures
  – Move: disfluencies, type of move
Limitations of literature

- Contradictory findings
- Studies vary greatly in
  - predictors (ratings, physical measurement)
  - measures of success (response rate, rated willingness to participate)
- Number of interviewers small, or not professional; number of contacts often small
- Little or no attention to
  - how interviewers speak, e.g., fluency
  - what interviewers say
- Most involved staged recordings so cannot examine interaction between I and A nor Is reaction to A
Other Attributes Associated with more Interviewer “success”

- Louder (Van der Vaart, Ongena, Hoogendoorn, & Dijkstra, 2005)
- Friendlier “hello” (Groves & Benki, 2006)
- Breathier, less masculine (Groves, O’Hare, Gould-Smith, Benki, Maher 2008)