

Use of paradata derived from vocal properties of interviewers and respondents in researching survey participation

Frauke Kreuter

June 26, 2014

Acknowledgement

- **Support:** Charles Cannell Fund; NSF Grants No. SES-0819734 and No. SES-0819725
- **Co-Investigators:** José Benkí, Jessica Broome, Frederick G. Conrad, Robert M. Groves
- **Statistical Consulting:** Susanne Hüttinger, Alexandra Birg, Helmut Küchenhoff
- **Coding, Transcribing, Acoustic Analysis, Sampling:** P. Batra, R. Benner, K. Franckowiak, H. Gu, B. Jarvi, E. Kordupel, P. Kotvis, a. Lincoln, L. Linstrom, M. Littlefield, D. Lopez, C. McClain, J. Matuzak, C. McCarty, P. Maher, G. Moss, K. Mull, D. Nielsen, D. Perkins, F. Pacheco, D. Popielarz, C. Sheffler, A. Tatro, D. Vanette, and D. Vollans

The Puzzle in CATI Interviewing

- CATI introductions are usually fairly scripted.
- Very little information is conveyed in the first few seconds or minutes of a CATI call.
- We see variation across interviewers in their success in gaining cooperation over the phone (both within a survey and across surveys).
- **Hypothesis:** Differences in interviewers verbal attributes play an important part in outcomes (Oksenberg, Coleman & Cannell, 1986; Oksenberg & Cannell, 1988).

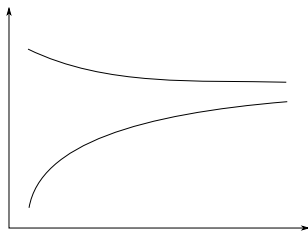
Past Findings of Speech Patterns on Participation

Positive effects on sample unit participation:

- Higher pitch for interviewers' voices (Sharf & Lehman, 1984; Groves, O'Hare, Gould-Smith, Benki, Maher 2008).
- Lower pitch for male interviewers (Benkí, Broome, Conrad, Groves, and Kreuter 2011).
- Less scripted, more extemporaneous deliveries of survey introductions (Groves, et al. 2008).
- Moderate levels of disfluency, rate, and pausing in interviewers' speech (Conrad, et al. 2013; Benkí, Broome, Conrad, Groves, and Kreuter 2011).

Open Research Questions

Speech researchers have observed convergence in conversations lasting 10+ minutes. It is unknown if these effects can be observed during the short duration of the survey invitation.



- Is there convergence of speech patterns between interviewer and answerer?
- Does convergence increase survey participation?

- Paradata on speech behavior in telephone interviews
- Sample of households/individuals selected from five studies
- 100 different interviewers (≤ 40 cases agree; ≤ 40 non-agree)
- Corpus of 1.380 audio recorded survey invitations
- Data available until the moment when the respondent ultimately agrees to the invitation or refuses to participate
- Data available on turn-level

What is a turn?

- Each contact is split into conversational turns taken by the interviewer and answerer
- Example conversation
 - Answerer: “Hello” (first turn)
 - Interviewer: “My name is . . . and I would like to . . .” (second turn)
 - Answerer: “What is the study about?” (third turn)
 - Interviewer: “It is about . . .” (fourth turn)
 - Answerer: . . .
- Differing number of turns

Variables

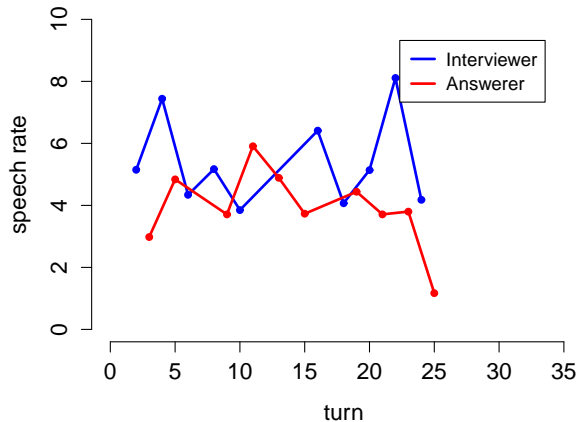
- Outcome
 - Answerers decision relating to the survey invitation (agree, refusal, scheduled call-back, hang-up, others)
- Controls (known to matter from Conrad et al. 2013)
 - **filler** number of fillers (e.g. “um”, “uh”) per 100 words, produced by interviewer
 - **answerer backchannel** proportion of backchannels (e.g. “uh huh”, “I see”) per contact
 - **overspeech** proportion of simultaneous speech between interviewer and answerer
 - **studies** Gujarati Community Survey (Gujarati), National Study on Medical Decisions (NSMD), Interests of the General Public (IGP), Mississippi Community Study (MCS), Survey of Consumer Attitudes (SCA)

Variables cont'd

Variables to create measures of convergence

- **speech rate** spoken words per sec (calculated from transcript and Praat timestamps)
- **gap** pause at the beginning of the turn
- **gap duration** duration of the gap in sec
- **pause** number of pauses during the turn: excluding gaps and logistic pauses
- **pause duration** duration of pauses in sec per turn
- **vocal pitch** median f0 in Hertz: the pitch of someones voice

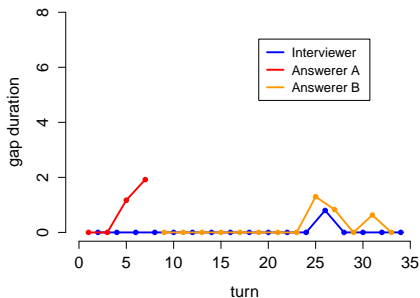
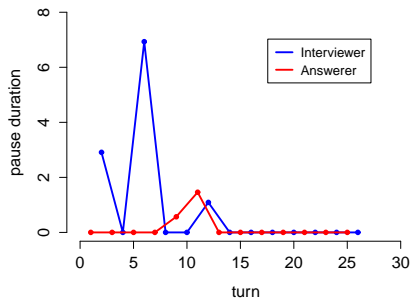
Speech Rate



- measured by Praat program
- range: from 0.43 to 9.99 words per second

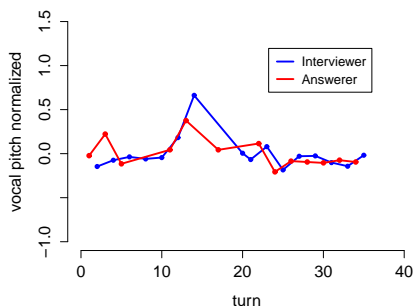
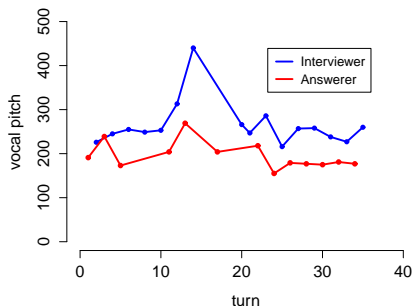
Gap Duration and Pause Duration

- range of pause duration from 0 to 17.41 seconds
- range of gap duration from 0 to 11.85 seconds
- duration of all pauses per turn added up
- many turns with no gaps and pauses



Vocal Pitch (normalized)

- vocal pitch calculated as median of vocal pitch per turn
- vocal pitch normalized by the mean of the speakers' vocal pitch values
- range of vocal pitch from 74 to 490 Hertz
- range of vocal pitch normalized from -0.66 to 2.26 Hertz



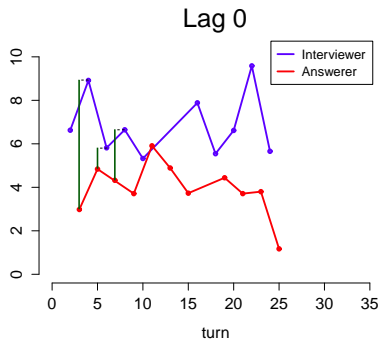
Measures of Convergence

Based on linguistic knowledge: convergence applies to speech rate, pauses and gaps

- calculate the differences between interviewer and answerer in consecutive turns
- five different ways of difference-calculation as the interviewer may not react immediately and exclusively to the last turn

Measures of Convergence

- **Lag0** (δ - interviewer turn and the former answerer turn)
- **Lag1** (δ - interviewer turn and the second to last answerer turn)
- **Lag2** (δ - interviewer turn and the third to last answerer turn)
- **Int1** (δ - interviewer turn and the mean of the last two answerer turns)
- **Int2** (δ - interviewer turn and the mean of the last three answerer turns)



Measures of Convergence

Separate the differences in begin / end and compare these values

- first half and last half differences
- first five and last five differences
- both methods for all five difference calculations to get the best fit
- general example:

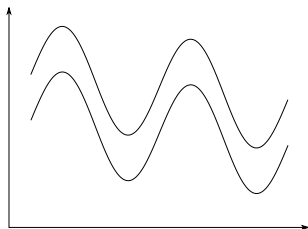
	begin				end			
turn	1	2	3	4	5	6	7	8
answerer	3		6		6		4	
interviewer		9		4		3		3
diff. (Lag0)	6		2		3		1	
mean	4				2			
	$4 - 2 = 2$							

→ the larger the difference of the means, the stronger is convergence

→ negative values imply divergence

Measures of Synchrony

Based on linguistic knowledge
not convergence but synchrony
applies to vocal pitch.



- calculate the pearson correlation (ρ) between interviewer and answerer in consecutive turns
- same five ways to examine ρ (as done for convergence), for interviewers may not react immediately and exclusively to the last turn

Analysis - Generalized Linear Mixed Model

Logistic Regression Model with random intercept

$$y_{ij} = \begin{cases} 1, & \text{agree} \\ 0, & \text{refuse or scheduled call-back} \end{cases}$$

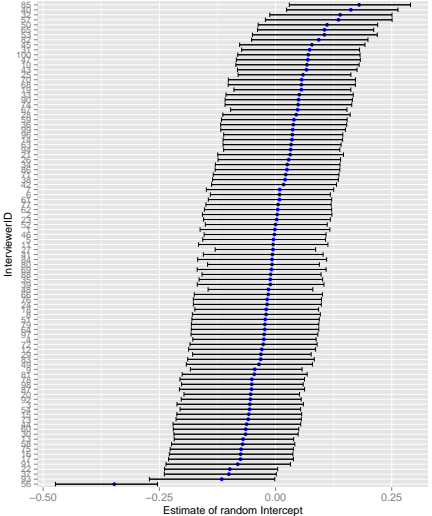
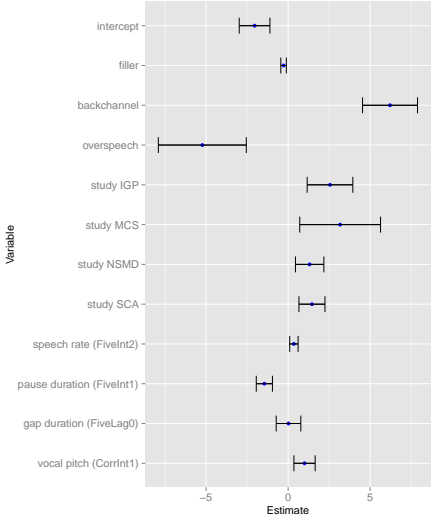
$$\log \left(\frac{\pi_{ij}}{1 - \pi_{ij}} \right) = \mathbf{x}'_{ij} \beta + \gamma_i$$

- π_{ij} probability of agreement, i.e. $\pi_{ij} = \mathcal{P}(y_{ij} = 1 | \gamma_i)$,
- \mathbf{x}_{ij} contact and interviewer covariates,
- β vector of coefficients,
- γ_i random effect, representing unobserved interviewer effect, assumed to follow normal distributions, i.e. $\gamma_i \sim N(0, \sigma_\gamma^2)$

Results

Random intercept	Variance	Std.Dev.	
Interviewer ID	0.082	0.287	
Fixed effects	Estimate	Std.Error	$Pr(> z)$
(Intercept)	-2.345	0.547	0.000018 ***
Filler	-0.278	0.086	0.001289 **
Backchannel	6.211	0.855	0.000000 ***
Overspeech	-5.230	1.368	0.000132 ***
Studies omitted	[.]	[.]	[.]
Speech rate (FiveInt2)	0.352	0.131	0.006991 **
Pause (FiveInt1)	-1.449	0.252	0.000000 ***
Gap (FiveLag0)	0.023	0.383	0.951574
Vocal pitch (CorrInt1)	1.000	0.331	0.002504 **
AIC	BIC	No. of obs.	groups
480.5	535.2	498	97

Results



Conclusion

- Summary of regressions-results:
 - **speech rate**: with increasing convergence the chance for agreement increases
 - **pause duration**: increasing convergence has a negative effect on the chance for agreement (very rare events; very short pauses systematically missing)
 - **gap duration**: increase of convergence would have a positive effect on the chance for agreement but is not significant on the 5% level
 - **vocal pitch**: with increasing synchrony the chance for agreement increases
 - Because of the missing values the results must be treated with caution
- Convergence and synchrony from I signals to A that I is adapting her speech to A. This could be micro-level instance of tailoring and thus encouraging participation.

Thank you!

fkreuter@umd.edu