

ATTACHMENT 10—SPEECH ITEMS IN THE NLSY97

As part of the Round 13 submission for OMB clearance, BLS proposed recording of open-ended respondent speech for use in coding and analyses pertaining to the relationship of wages to speech differences among non-Hispanic whites and African-Americans. At that time, OMB requested additional documentation that the proposed approach for recording and coding open-ended responses would be effective. BLS did not include speech items in the Round 13 main interview.

Since then, several events have transpired pertaining to the potential effectiveness of recording and coding open-ended speech. These include the award of an NIH grant ([project number 1R21HD065033-01A1](#)) to University of Chicago professor Jeffrey Grogger with very strong evaluation marks for the inclusion and coding of these items in the NLSY97 instrument. The proposal to include speech items in the Round 15 interview is a consequence of the award. Also, earlier work done by Grogger using NLSY97 validation interview recordings and a less formal procedure for coding speech was published in the *Journal of Human Resources* in an article entitled, “Speech Patterns and Racial Wage Inequality” (**Volume 46, Number 1, Winter 2011**). These two most recent events document the relevance of the approach to increasing understanding of racial disparities in labor market outcomes, as well as the feasibility of recording the desired responses.

A brief overview of speech items in the NLSY97 over the last several rounds may be helpful. In Round 10, validation interview recordings were used to capture raters’ perceptions of respondents’ speech patterns. Those recordings generate the data used in the *JHR* article mentioned above. For Round 13, a set of speech questions underwent cognitive testing and were tested in the pretest. The pretest indicated that the internal microphones in the NLSY97 field interviewers’ laptops were inadequate for recording open-ended responses. External microphones were adopted for all field interviewers to improve the recording quality. In response to questions from OMB, Grogger secured a research grant from NORC at the University of Chicago to conduct a small study of ‘code-switching’ items among a purposively selected sample of African-Americans in south suburban Chicago. The results of that study indicated that some (more formal) proposed questions were successful in eliciting code-switching behavior (in which individuals who had been using African-American speech switched instead to Standard American English), while some (more informal) proposed questions did not elicit code-switching.

In a related development, the Moving to Opportunity study sponsored by the U.S. Department of Housing and Urban Development, together with the MacArthur Foundation, has also administered a series of open-ended speech questions for recording. That multi-site study of families previously living in low-income housing successfully administered and recorded similarly structured open-ended items that are now being analyzed by researchers.

Through this sequence of events, we are able to propose for the Round 15 NLSY97 instrument a set of items that have been tested in a variety of contexts. We designate three subsamples for these questions: The focal group consists of all African-American respondents, all non-Hispanic white respondents living in the South at age 12, and 500 randomly selected respondents not falling in these two categories. The remaining respondents are randomly assigned to the two other subsamples.

African Americans are central to the project. Southern whites are important because Southern white dialects share many speech features with African American English. By coding these features for Southern whites, we can test whether wages are influenced by speech patterns per se, or by interactions between speech patterns and race. The non-African American, non-Southern-white group is included at the suggestion of the NIH reviewers. Reviewers were interested in whether coding protocols could be developed for other speakers in such a way that would permit a broader study of the importance of speech for the labor market. The reviewers explicitly requested a pilot study to be carried out for this purpose. The sample size of 500 for the non-African American, non-Southern-white group was chosen because Professor Grogger’s earlier pilot study published in the *JHR* included roughly 500 speakers, which was enough to significantly distinguish the speech of whites and African Americans.

The rationale for coding the speech patterns of respondents outside the focal group is more forward-looking. Professor Grogger’s current research is focused primarily on African American speech patterns, but speech patterns can affect employment opportunities and earnings for many other groups as well, such as Hispanics or Latinos,

immigrants and children of immigrants, and persons with disabilities. Coding the speech patterns of respondents outside the focal group broadens the sample size available for future analyses of speech patterns. Respondents in the focal group will be asked a question about job search that is designed to elicit formal speech. Respondents in the focal group also will be asked questions designed to elicit informal speech about what they like to do in their free time. One of the remaining subsamples will be asked the job-search question, and the other subsample will be asked about their free time. Asking respondents outside the focal group to answer one set of questions or the other but not both sets reduces respondent burden and the costs of coding speech patterns.

To elicit code-switching behavior, we ask for a self-description as one might provide to a potential employer (asked of the focal group and one of the two random halves):

YEMP-118005/ YEMP-119460

[Suppose you decided you wanted to look for a new job, and the place where you really wanted to work called you and asked you to come in for an interview./ Let's suppose you applied for a job that sounded really interesting to you and they called you and asked you to come in for an interview.] How would you describe your skills, qualifications, and experience to me if I were the person interviewing you for the job? INTERVIEWER: R SHOULD SPEAK UNTIL TIMER INDICATES 0 SECONDS REMAINING (i.e. one elapsed minute). **R'S RESPONSE WILL BE RECORDED. PLEASE DO NOT TYPE WHILE RESPONDENT IS SPEAKING.**

AS NEEDED, PROBE AS FOLLOWS:

- Tell me more

- What kind of experience have you had?

- What (other) skills do you have?

We also propose questions designed to retain African-American speech if that is the respondent's default choice (asked of respondents who are in the speech focal group, or the second randomly selected half):

TEL-87 We are also interested in what people do in their free time. Do you listen to music?

YES Go to (TEL-87a)

NO

Default: TEL-88

TEL-87a Who are your favorite musicians? INTERVIEWER: R SHOULD SPEAK UNTIL TIMER INDICATES 0 SECONDS REMAINING (i.e., one elapsed minute). **R'S RESPONSE WILL BE RECORDED. PLEASE DO NOT TYPE WHILE RESPONDENT IS SPEAKING.**

AS NEEDED, PROBE AS FOLLOWS:

- What do you like about them or their music?

- How would you describe their music?

- What are your favorite tunes?

- What about their videos? Any video you think that is really great? Why?

Default: YTEL-61

TEL-88 Do you watch TV?

YES Go to (TEL-88a)

NO

Default: TEL-89

TEL-88a What are your favorite shows? INTERVIEWER: R SHOULD SPEAK UNTIL TIMER INDICATES 0 SECONDS REMAINING. **R'S RESPONSE WILL BE RECORDED. PLEASE DO NOT TYPE WHILE RESPONDENT IS SPEAKING.**

AS NEEDED, PROBE AS FOLLOWS:

- What happened on the last show you watched?
- Which of the cast do you like best? What about them do you like?
- (For reality shows) Who do you think is going to win? Why?

Default: YTEL-61

YTEL-89 What's your favorite way to spend free time? INTERVIEWER: R SHOULD SPEAK UNTIL TIMER INDICATES 0 SECONDS REMAINING (i.e. one elapsed minute). **R'S RESPONSE WILL BE RECORDED. PLEASE DO NOT TYPE WHILE RESPONDENT IS SPEAKING.**

AS NEEDED, PROBE AS FOLLOWS:

- Is that something you usually do alone or with other people?
- When was the last time you spent time that way?
- When did you start spending time that way?
- If you had more free time how would you like to spend it?

Default: YTEL-61

In the suburban Chicago pilot, these YTEL questions performed similarly to the job question. The mean number of words elicited in response to the job question was 113. The mean number elicited by the music question was 83, and the mean number elicited by the TV question was 166. The linguist who is collaborating with Professor Grogger was pleased with these results. The interviewer probes were key to eliciting adequate amounts of speech and came easily once the interviewers gained more experience administering these questions.