Title: Vehicle City Voices (VCV) Corpus – Part I: An Oral History and Linguistic Survey of Flint, Michigan

Author: Erica Britt
Language: English
Applications: discourse analysis, narrative, prosody
Data Types: Audio, Text, and Excel Spreadsheet
Data Size: 1.99 GB
Hours of Audio: approximately 16.7 hours
Number of Words: 168,715 words
Character Encoding: Unicode
Format/Structure: XML
Sample Rate: 44100
Format: FLAC
Bit Depth: 16
Channel Count: 1

The Vehicle City Voices (VCV) corpus, compiled at the University of Michigan-Flint, is an ongoing oral history project and survey of English language variation in Flint, Michigan. This corpus was designed with the aim of providing researchers access to high-quality recordings that would be useful for acoustic analysis. In addition, the oral history interview, a semi-structured interview format that focuses on the interviewee's experience growing up and living in Flint, provides an important context for the examination of narrative structure as well as the discursive construction of individual and collective identity in urban spaces.

The 21 interviews in this corpus were collected between 2012 and 2015 by undergraduate and graduate students for civic engagement projects in linguistics courses and by a graduate student research assistant. Residents were recruited through references from community pride organizations, as well as through friends, family members, and community contacts of the interviewers. Residents were either born in or have lived in Genesee County, Michigan for most of their lives. The residents in this corpus, who represent a range of ages, genders, and ethnicities, were born between 1935 and 1991. Of the 21 interviewees, 11 are Black/African American, 8 are White/Caucasian, and 2 are biracial/mixed ethnic heritage. This release of the corpus contains 11 female and 10 male participants.

The interviews were recorded using Marantz PMD661 portable SD recorders with accompanying Audio-Technica AT831B lavalier condenser microphones. The original recordings were uncompressed (PCM-16) sound files stored in WAV format recorded at a sampling rate of 44,100 Hz. These files were then converted to FLAC format.

Face-to-face interviews took place in a number of quiet locations in the Flint community including university spaces, community meeting spaces, and a church meeting room. Interview questions focused on resident recollections of important

community events and festivals, their remembrances of important buildings, community centers, and neighborhoods that may or may not still exist, their (or their family's) relationship to the automotive industry, and their reflections on the economic, demographic, and physical transformation of the city. In addition, some residents were also asked specific questions about their sense of community pride, their own activism, and the negative perceptions of Flint that they have encountered.

Metadata for this corpus includes (where provided by participants) information on the interviewee's gender, ethnicity, year of birth, level of education, field of employment, average income, length of time living in Flint and its surrounding areas as well as interviewer age, gender, and ethnicity. In addition, original interview durations, edited interview durations, interview year, and transcript word counts are also provided in the metadata file.

Each speech file is accompanied by a transcript. Personal names, full addresses, phone numbers, etc., in the transcripts have been removed where necessary in order to preserve the anonymity of the interviewees and their acquaintances and the portions of the original audio files containing this information have been removed. Names of public figures have not been removed from these interviews. Transcriber commentaries on elements of the interview situation or descriptions of deleted information are presented in parentheses in the transcripts.