

The Nationwide Speech Project: A new corpus of American English dialects [☆]

Cynthia G. Clopper ^{a,*}, David B. Pisoni ^b

^a *Department of Linguistics, Northwestern University, 2016 Sheridan Road, Evanston, IL 60208, USA*

^b *Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN 47405, USA*

Received 11 May 2005; received in revised form 23 September 2005; accepted 27 September 2005

Abstract

Perceptual and acoustic research on dialect variation in the United States requires an appropriate corpus of spoken language materials. Existing speech corpora that include dialect variation are limited by poor recording quality, small numbers of talkers, and/or small samples of speech from each talker. The Nationwide Speech Project corpus was designed to contain a large amount of speech produced by male and female talkers representing the primary regional varieties of American English. Five male and five female talkers from each of six dialect regions in the United States were recorded reading words, sentences, passages, and in interviews with an experimenter, using high quality digital recording equipment in a sound-attenuated booth. The resulting corpus contains nearly an hour of speech from each of the 60 talkers that can be used in future research on the perception and production of dialect variation.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Speech corpus; Dialect variation; American English

1. Introduction

1.1. Phonological dialect variation in the United States

Researchers have been documenting regional linguistic variation in the United States for more than

a century. The American Dialect Society was founded in 1889 with the goal of collecting a comprehensive American English dictionary. Krapp (1925) documented regional varieties of American English based on grammar and pronunciation guides dating back to the 18th century. From this research, he identified three main dialects of American English: Eastern, Southern, and Western (or General American). Thirty years later, McDavid (1958) described the early Linguistic Atlas projects in the United States, which documented lexical and phonological variation based on fieldwork interviews conducted in predominantly rural areas. McDavid (1958) concluded that the major dialects of American English were Northern, Midland, and

[☆] Parts of this research were presented at the 147th Meeting of the Acoustical Society of America in New York, NY, May 24–28, 2004.

* Corresponding author.

E-mail addresses: c-clopper@northwestern.edu (C.G. Clopper), pisoni@indiana.edu (D.B. Pisoni).

Southern. He also acknowledged that these dialects were more sharply distinguished on the Atlantic seaboard and that more transition areas between dialects were found as one moved westward across the country. Carver (1987) also examined regional lexical variation but he described only two primary dialects of American English: Northern and Southern.

More recently, linguistic variation has been explored using acoustic–phonetic analysis techniques. Thomas (2001) obtained acoustic–phonetic vowel spaces for nearly 200 individual talkers. Although Thomas (2001) made no explicit claims about specific dialect regions, he did group his talkers into a Northern group and a Southern group. Labov and his colleagues (forthcoming) have been working on a more comprehensive study of regional variation in American English. The Telephone Survey (TELSUR) project at the University of Pennsylvania includes telephone interviews with 700 talkers representing all major urban areas in the United States. The recordings have been analyzed acoustically and Labov (1998) defined three major dialects of American English based on the vowel systems of his 700 talkers: Northern, Southern, and the “third dialect”. This third dialect includes Eastern and Western New England, Western Pennsylvania (centered on Pittsburgh), the Midland, and the West. Labov (1998) described the Mid-Atlantic metropolitan areas from New York City to Washington, DC as “exceptions” to the three-dialect division because speakers from this region do not exhibit the characteristic properties of any of the three major dialects. The Florida penin-

sula is also treated as a unique region because of the high level of dialect mixing that occurs there as a result of migration from other states. Fig. 1 is a map of the United States showing these regions, based on the work by Labov and his colleagues (forthcoming). No data are available for the gray areas on the map because these regions are sparsely populated and the TELSUR project focused on 145 urban areas with an average population of 1.7 million people, ranging from 88,000 in Aberdeen, South Dakota to 17.6 million in New York City (Ash, n.d.).

The vowel system of the Northern dialect of American English is characterized by the Northern Cities Chain Shift (Labov, 1998). The Northern Cities Chain Shift is a clockwise shift of the low vowels that includes the fronting and raising of /æ/, the fronting of /ɑ/, the lowering of /ɔ/, and the backing of /Λ/ and /ε/. /ɪ/ is also reported to be backed in the Northern dialect as a parallel shift to /ε/ backing. Fig. 2 depicts the major features of the Northern Cities Chain Shift.

The Southern dialect of American English is characterized by the Southern Vowel Shift (Labov, 1998). The primary feature of this shift is the fronting of the back vowels /u/ and /o/. In addition, the front lax vowels /ɪ/ and /ε/ are raised in Southern American English and the front tense vowels /i/ and /e/ are lowered. The Southern Vowel Shift is shown in Fig. 3. The Southern dialect is also characterized by the monophthongization of the diphthongs /ɑy/ and /oy/ (Thomas, 2001).

The common feature of the “third dialect” of American English is the merger of the low back

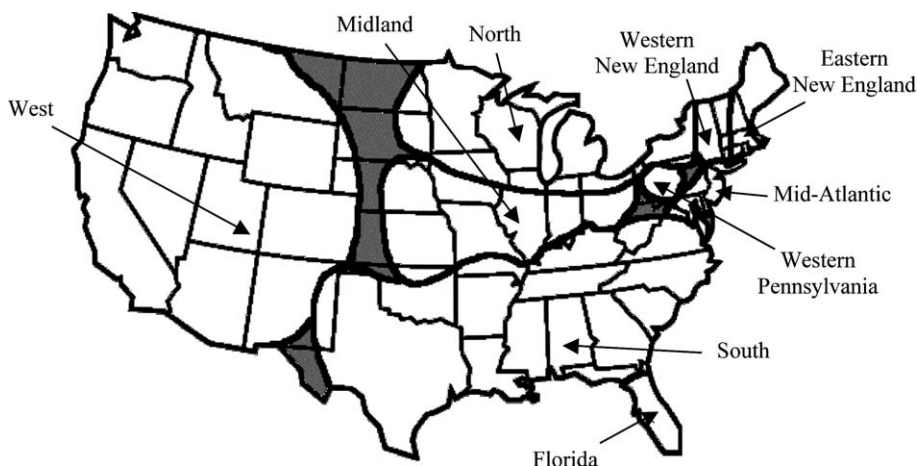


Fig. 1. The major dialects of American English, based on Labov et al. (forthcoming).

Download English Version:

<https://daneshyari.com/en/article/567773>

Download Persian Version:

<https://daneshyari.com/article/567773>

[Daneshyari.com](https://daneshyari.com)