

A variationist perspective on the pronunciation of *the* in English: A TV corpus analysis
Kym Taylor and Walcir Cardoso, Concordia University

Language variation and change are dynamic, positive forces that constitute evidence of linguistic vitality (Labov, 1972). Therefore, documentable changes in pronunciation – especially when driven by certain social groups – are of special interest to variationist linguistics and related disciplines. Consider the pronunciation of the English determiner ‘the’, which is often described as pronounced [ði] in pre-vocalic contexts (e.g., the apple) and [ðə] elsewhere. There is some evidence (Taylor, 2016) that in actual use, the [ði] + vowel rule is being replaced by [ðə] + vowel in certain contexts by native English speakers in the United States. This paper considers the implications of such a shift and how it can be captured within a variationist perspective.

The general question that guided this study was: Which linguistic and extralinguistic factors influence the variable production of the article ‘the’ in pre-vocalic position in English? To answer this question, we stratified data among a set of potentially influencing factors. The linguistic factors included: the following (post-the) phonological environment (vowel height: high, mid, low; vowel backness: front, mid, back; vowel quantity: lax, tense), the stress pattern of the following syllable (stressed or unstressed), and the location of the target article within the sentence (initial, medial or final). Influenced by research on the effects of social factors on language variation (e.g., Eckert, 1997; Hannah & Murachver, 1999; Tagliamonte, 2012), the extralinguistic factors included: age (divided in 5-year intervals), gender (male or female), occupation (e.g., teacher, luxury real estate agent), area of residence (Northeast, Northwest, Southeast, Southwest), and cognitive planning (planned or unplanned speech). The latter factor was motivated by research on the effects of cognitive planning on non-standard pronunciation (e.g., Bortfeld et al., 2001).

Speech samples were collected from fifteen American television programs: ten minimally scripted/unscripted (e.g., home improvement, cooking competition), and five traditionally scripted (e.g., situation comedy, crime drama), distributed among the relevant social classes considered for the study (i.e., gender, age, occupation and area of residence). The coding and analysis of the data followed standard sociolinguistic/variationist protocols and were subjected to multivariate analysis using Goldvarb X (Sankoff, Tagliamonte, & Smith, 2005).

The statistical results indicate that the alternation observed in ‘the’ production is being neutralized towards [ðə], as hypothesized by the authors, and that the extralinguistic factors age and gender play an important role in favouring the non-standard [ðə] variant: this form is more frequently found in the speech of young women, an indication of language change in progress (see Cameron (2003) for a similar claim). In addition, [ðə] is favored in the speech of those in blue collar and middle class occupations (e.g., building contractor and police officer respectively), and when the oral production is unscripted or minimally scripted (i.e., unmonitored), serving as further evidence for a potential pattern of sound change: “sound change is pre-social, originating in the most unconscious and systematic reaches of the speaker’s linguistic system, and emerging in the speaker’s most unmonitored speech” (Eckert, to appear). Finally, the [ðə] variant was favored when followed by a lax vowel.