2002

[Introduction to] Race and the Rise of Standard American

Thomas Paul Bonfiglio
University of Richmond, tbonfigl@richmond.edu

Race and the Rise of Standard American

by

Thomas Paul Bonfiglio

Mouton de Gruyter
Berlin · New York 2002

Introduction

In the first half of the twentieth century, Americans began to view the accent of the midwest and west as a "general American accent" that represented a standard for pronunciation. In the second half of the twentieth century, American linguists began to reject the rubrics of *midwestern* and *general American* and to problematize the status of a standard American speech in itself. This had little or no effect upon the popular consciousness; folkish notions of a standard American and (mid)western accent continued throughout the century and were extended to include network broadcast speech, as well. Indeed, Americans came to recognize the pronunciation of network announcers as a (mid)western norm. The general features of this accent are readily identifiable; the phoneme /r/ is pronounced both before and after vowels, there is no intrusive /r/, as in "I 'sawr' her standing there," diphthongs like /ay/ and /aw/ are not monophthongized, and the phoneme /æ/ is used in words like rather, bath, and calf. Americans came to recognize obvious deviations from these sounds as nonstandard and regional, such as the dropping of /r/ after vowels in New York and Boston, the Bostonian pronunciation of "rather" so that it rhymes with "father," and the southern pronunciation of "right" as /ra:t/.

The question as to why and how this (mid)western accent rose to be perceived as the standard has neither been satisfactorily answered nor engaged in a systematic way. The discourses of popular social science and popular opinion have been content with tangential and impressionistic explanations for the evolution of standard American pronunciation. The discipline of sociolinguistics has not fared much better in this regard. It has either avoided the issue, offered its own insufficient explanations, or made some late inroads, most notably in the research done in the recently emerged field of perceptual dialectology (folk linguistics).

Explanations for the etiology of standard American pronunciation have been riddled with misprisions from the onset. Some of the major ones are:

–American English pronounced the /r/ after vowels in order to differentiate from British speech, not from other forms of American speech.
–Because America is a democracy, the speech of the average person was taken as a standard. Two-thirds of the country pronounced the /r/ after vowels in the 1920's; the standard was simply the pronunciation of the majority.
–The standard that arose was simply the pattern of speech that was most pleasing to the greatest spectrum of radio listeners.
–The early radio announcers were from the midwest. This caused the mid-western pronunciation to become imitated and standardized.
–American English has no real standard pronunciation. There are many speech areas and differing pronunciations within any given speech area.
–There is no such thing as "general American" or even "mid-western" pronunciation.

This study progressively engages and deconstructs these myths in the process of developing its thesis.

My curiosity on this subject was stimulated by the observation that the process of standardization in the United States occurred in a fashion quite dissimilar from standardization in other countries, especially as regards phenomena of economic, social, and cultural power. Economic power is an important determinant of the status of a kind of speech and generally marks the difference between a language and a dialect. There are some jokes in linguistics that demonstrate this; one is that a dialect becomes a language when the dialect speakers get rich; another is that a language is a dialect with an army. In general, the standard language of a nation will derive from the speech area that is also the center of economic and cultural power in that nation. Examples of this are the British "received pronunciation," which derives from upper-class London speech; simi-

larly, Parisian is the hegemonic standard for French, and the standard for German is generally associated with the northern industrial centers. It is highly uncommon that standard pronunciation should be taken from rural or agrarian areas. It would be strange to imagine British emulating the speech of Yorkshire or German emulating the Alpine dialects. Yet, this is basically what happened in the standardization of American English. The pronunciation of the economic and cultural centers of power was not taken as a model. Instead, the pronunciation of a largely rural area, the midwest and west, was preferred.

New York was clearly the American center of economic power at the turn of the twentieth century. It had a metropolitan population of nearly four million, at a time when there were only two other American cities with populations over a million, and was the cultural center of the country as well. Along with Boston, it centralized the power of the northeast, which was clearly the most influential part of the country at that time. The combined population of New York and the New England states comprised one-sixth of the national population in 1900 and had comprised one-fourth of the national population in 1850. The most distinctive phonetic feature of this area was the marked dropping of postvocalic /r/. Why then did this feature not develop into the national standard? Some massive cultural counterforce must have been at work here that was strong enough to override the power of the patrician pronunciations of New York and New England, which remained the determinants of American stage pronunciation in the first half of the twentieth century. This stage pronunciation generally replaces postvocalic /r/ with a schwa. The diction of Katharine Hepburn is a prime example of this type of speech, and one would have well expected it to rise to the status of a national standard, especially in view of the cultural power of such figures as Hepburn and of the New York milieu with which they are associated. Even though this was also the pronunciation for radio plays, it eventually yielded to the (mid)western pronunciation for radio broadcast speech.

The period of standardization of American pronunciation coincided with the growth of radio, and these developments also occurred during and in the aftermath of the passing of 12 million im-

migrants through Ellis Island, New York (1892–1924). Most of this immigration was from southern and eastern Europe. In 1907, 75% of immigration was from those regions. By 1910, 75% of the population of New York and Boston was comprised of immigrants or the children of immigrants, and 25% of the population of New York consisted in Russians and Italians. 1907 was also the year that the American congress started looking into the restriction of immigration. This culminated in the Immigration Quota Acts of 1921 and 1924, which reduced the average southern and eastern European immigration from an average of 783,000 per year to a maximum of 155,000 in 1921 and 25,000 in 1924 (Chermayeff 1991: 70, 17). The cultural and economic national capitals of New York and Boston came to be seen as sources of contamination of the "purity" of America. This was especially true of New York, which saw the immigration of 2.3 million eastern European Jews, and which became the focus of extreme antisemitism. This aversion to the large cities may be compared to similar phenomena in the prefascist movements in Germany at the turn of the century that idealized the rural German as an unspoiled, uncitified, and unsemiticized noble man of the soil. For similar reasons, Americans began to emulate the (mid)westerner; he was the Nordic man, be he of native Anglo-Saxon or immigrant northern European "blood."

Major shifts in the cultural values of a nation will be reflected in the language of that nation. This brings me to my thesis: the adoption of western speech patterns as the preferred norm was influenced by the xenophobic and antisemitic movements of the early twentieth century. Thus Americans gravitated toward the pronunciation associated with a "purer" region of the country, and they did so in a largely non-conscious manner. Consequently, this study gradually moves toward the reintroduction of the regional terms *western* and *midwestern,* which linguistics discarded after 1945 as overgeneralizations. This study shows that the ideological construction of the categories *western* and *midwestern* was a prime agent in the process of the standardization of American pronunciation.

Thus this study coordinates a dialogue between the waxing xenophobia of the early twentieth century and the discussion of American pronunciation, linking the two via the common discourses of

empowerment, disempowerment, and the articulation of identity. The dynamics of pronunciation that I am trying to illuminate by using models of ethnocentrism are largely unconscious. While the antisemitic and racial statements themselves were clearly conscious, if not shameless, the evolution of pronunciation itself was not one that was consciously mapped out, nor was it the product of a conscious, unified decision. It is analogous to the phenomenon of the post-war "white flight" to the suburbs, which was a process of gradual and incremental gravitation, the ethnocentrism of which can generally be read only on the level of submerged or coded discourse.

While this study is clearly indebted to the work done by William Labov on the changes in the speech patterns of New York City, it reviews that work, however, within an alternative methodology. Labov's findings, produced by an inquiry that is categorically linguistic in nature, are rearticulated here in a broader sociohistorical and sociocultural context, which enables this study to arrive at different causal explanations than those offered by Labov. While Labov speculated that the shift in the pronunciation of postvocalic /r/ could be coordinated with the role of the United States in World War II, this study demonstrates that the determining factors for the change were already operative well before the decade of the forties and corresponded to radically different social phenomena.

It is not the purpose of this inquiry to offer a detailed description of the phonetics of American English. Indeed, such an endeavor would be an impediment to the objective at hand. I am concerned instead with the larger cultural causes for the popular perception and valorization of regional accents and with describing the cultural milieu that gave rise to positive and negative value judgments. For this study will seek to demonstrate that it was the prejudices of nonlinguists that created the idea of standard American pronunciation. In his work on perceptual dialectology, Preston (1999) has pointed out that it is imperative to study "the triggering mechanisms of language regard among the folk and through such study the potential influence of such regard on the more general process of variation and change" (xxxviii). In his studies of the perceptions of standard United States English, he emphasizes that "research puts the weight of describing SUSE precisely where it belongs–in the mind, out of

the mouths, and from the word processors of nonlinguists" (29). And this evidence can answer the questions as to how and why American English pronunciation standardized as "network standard" or, informally, "midwestern" in the twentieth century.

It should be emphasized, however, that the phenomenon of a standard language cannot be reduced to pronunciation alone, which is but a subset thereof; nor can it be claimed that postvocalic /r/ constitutes a whole variety in itself. This study views pronunciation, especially that of postvocalic /r/, rather as a reduction, as a symptomatic and metonymic indication of a preferential shift in prestige discourse, and not as constituting prestige discourse in itself.

In order to illuminate the cultural milieu that generated the popular perception and evaluation of regional accents, this study focuses on the linguistic, racial, and ethnic ideologies of influential figures in the United States, among them statesmen, writers, philologists, speech trainers, and historians. It also investigates the perception and reception of the accents of major American actors, announcers, and political figures. The ideologies and receptions of such influential figures are not only symptoms, but also determiners of the national consciousness of pronunciation as it relates to race, class, and power. With that in mind, the study discusses the findings of linguistic experiments on attitudes toward various American accents, for explicating the influence of the kinds of American figures mentioned above can help reveal the larger socio-cultural background that determined the results of those experiments and place the data in a larger interpretive context.

Consequently, the investigation will concern itself with phonemes that have high cultural visibility and that can be focused upon as diagnostic markers of the migration and legitimation of accent. The most central and pivotal of these phonemes is the characteristic American /r/; it was a principal marker of the difference between British and American, as well as between inland and coastal American speech. This phoneme became a major point of contention in pronunciation debates, invested with the ideologies of the first half of the twentieth century, and supercharged with linguistic capital and cultural significance.

The standard American postvocalic /r/ is referred to in this study variously as continuant, constricted, alveolar, retained, and rhotic. (The category of retroflex is reserved for the /r/ of the inland, i.e. non-coastal south, which includes the southern mountain, south midland, and Texas areas.) All of these designations refer to the same phoneme; it is the unmistakable sound of /r/ heard in the diction of standard network broadcasters from Lowell Thomas to Walter Cronkite and Dan Rather. It is peculiar to the United States and Canada. It is contrasted with the coastal postvocalic /r/, which is referred to here as dropped and non-rhotic. Among the other strong phonemes discussed are the more constricted retroflex /r/ of the inland south, the phoneme /oy/ if the New York and Tidewater areas, the back vowel /ɑ/ of the northeastern coast, found in the Boston pronunciation of *dance* as /dɑnts/, and the inland standard low front vowel /æ/. This last vowel is also a very strong marker of the characteristic American pronunciation.

It will be emphasized throughout this study that the phonemes in question have no essential value in themselves. The history of post-Saussurean linguistics has firmly held that there is no natural or ontological connection between a sign and its referent. This means that signs in themselves do not possess any particular intrinsic value or meaning; their value is gotten by virtue of their relationship to other signs. Thus value is culturally constructed by an associative network of signs. Sounds will gain value in the same fashion. A certain sound becomes associated with a certain positive or negative sign or image. Then, the relationship becomes reciprocal, not unlike a conditioned response, with the sign evoking the sound image and the sound image evoking the sign. Finally, the relationship becomes iconic, and the sound image is held to convey the value in itself. Network standard speech, which arose by the power of its association with midwestern and western speech, came to evoke positive personality images, i.e. to "sound better." Thus the characteristic phonemes of that speech came to indicate these positive personality values. To say that these phonemes in themselves already had *a priori* the requisite positive connotations would be untenable and would contravene the progress of linguistics in the twentieth century.

Chapter one of this study develops a social theoretical construct for analyzing the legitimation of accent, reviews the recent literature on language standardization, and develops a working concept of standard American English, especially in the context of power, race, and class. It also accounts for the differing regional pronunciations of postvocalic /r/ and the origins of those differences. Chapter two focuses on the relationship between pronunciation and ideology in the eighteenth, nineteenth, and early twentieth centuries and demonstrates that the prescriptive discussion of proper American pronunciation does not exist in a vacuum, but is instead buttressed and rationalized by ideological interests of morality, class, race, and ethnicity. It also shows how fundamental ideologies of race and immigration were instrumental in determining the modes of the broadcast voice. In order to illustrate the socio-cultural context that generated prescriptions on pronunciation, the methodology of chapter two departs from the realm of the purely linguistic. These excursions, however, are always intended to be viewed for their sociolinguistic implications, for the purpose of this study is to demonstrate that there is, in the United States, a long historical tradition of confounding the linguistic and the extra-linguistic and of configuring pronunciation within a matrix of race and class. Chapter three examines the relationship between immigration to the eastern seaboard and migration to the western regions and correlates this relationship with a phonemic shift away from New England and New York toward western and midwestern prestige patterns. It also shows how this shift precipitated a reversal in the speech patterns of New York City itself.

The (mid)western accent was constructed and desired by forces external to the area itself that projected a preferred ethnicity upon that region and defined it within a power dynamic of difference, i.e. it was precisely *not* the speech of the ethnically contaminated areas of the northeast metropolis and the south. Prior to a discussion of the social, cultural, and historical contexts of the discourses of race, ethnicity, and standardization in the United States as they existed in the popular sphere, it is necessary to develop an operative model of pronunciation as a strategic social phenomenon that is determined by factors of economy, prestige, status, and power.