Voice-conditioned allophones of MOUTH and PRICE in Bahamian Creole

Janina Kraus

Department of English Linguistics, Department of Phonetics and Speech Processing
Ludwig-Maximilian-University, Munich

Abstract

Allophonic height alternations in MOUTH and PRICE, conditioned by coda voicing, have been confirmed for a wide variety of English accents, especially in the North American context. Realisations may vary in degree but not in direction. Moreton and Thomas [1] suggest that universal processes may underly the varying productions and propose the Asymmetric Assimilation model to unify observations on the effect of voicing. So far, the hypothesis has not been tested on creolised varieties of English. This study aims to describe the spectral nature of voice-conditioned allophones of MOUTH and PRICE in two social varieties of Bahamian Creole. The results of an acoustic analysis of 15 creole speakers from Nassau, Bahamas, support the claim that automatic phonetic processes may underly the initial direction of the allophony, which is then available for sociolinguistic differentiation.

Index Terms: Bahamian Creole, Caribbean, acoustic analysis, social variation, Canadian Raising

1. Introduction

1.1. Asymmetric Assimilation

Pre-voiceless nuclear raising is generally described as a rule-governed alternation in which the diphthongs in PRICE and/or MOUTH are raised to mid height before voiceless consonants. Often referred to as ‘Canadian Raising’ [2], this alternation is by no means restricted to Canada, but has been observed in many North American English accents from Canada to coastal South Carolina and Georgia (cf. [1]) and in the British Fens [3]. In a number of other dialects, for example in varieties of African American Vernacular English [4, 5], PRICE is reported to undergo glide weakening or monophthongisation before voiced consonants. Moreton and Thomas [1] propose that these phenomena are actually two sides of the same coin, arising from universal phonetic processes which protect the glide against undershoot in pre-voiceless contexts. Voiceless codas cause peripheralisation of diphthong glides in formant space, which can be accessed by speakers as a perceptual cue for lexical distinction. This has been confirmed for speakers who do not exhibit phonologised Canadian Raising [6] as well as for closing diphthongs other than PRICE and MOUTH [7]. Furthermore, pre-voiceless shortening in diphthongs appears to affect primarily the beginning of the transition, which leads to shorter nucleus durations and possibly loss of nuclear steady states [6]. A shortened nucleus is consequently more exposed to coarticulatory pressures from neighbouring sounds. These processes create the condition for asymmetric assimilation, whereby pre-voiceless contexts promote the assimilation of the nucleus to the glide, while the glide is dominated by the nucleus in pre-voiced contexts. Moreton and Thomas [1] suggest that in the initial stages, pre-voiceless raising and glide weakening are just a matter of subtle phonetic differences between glides in pre-voiced and pre-voiceless context. As assimilatory pressures continue to persist, the spectral difference may become phonologised and the allophones may change even further. In the case of glide weakening, the difference between nucleus and glide continues to wane in pre-voiced contexts as the glide becomes more centralised. If, however, pre-voiceless raising spreads to the nucleus, a new case of Canadian Raising emerges.

1.2. MOUTH and PRICE in Bahamian varieties

Bahamian Creole (BahC) is spoken by about 250000 speakers in the Commonwealth of the Bahamas. It is the mother tongue of the majority of Afro-Bahamians, who account for about 90% of the total population [8], while the nation’s official language is English. The roles that standard English and BahC play in Bahamian society must be seen in the context of the development of other anglophone countries in the Caribbean from British colonial societies to independent nations and a culture strongly influenced by North America. BahC differs from other Caribbean varieties, however, in the fact that it did not evolve on Bahamian soil, but originated from an 18th century North American import, an earlier form of the creole language Gullah [9, 10]. Today, the Bahamas is a place of great linguistic diversity, internally as well as with regard to competing external influences. With continued migration to the cities, New Providence, the site of the nations capital Nassau, is now home to about 70% of the total population [8]. The urban variety of BahC can therefore be regarded as most representative of the creole at large and will be the focus of the present study.

Prior accounts of Bahamian vowels are scarce and present a rather heterogeneous picture, as authors, for the most part, do not distinguish consistently between the different regional, social and ethnic varieties. The most comprehensive descriptions of Bahamian vowels can be found in Wells [11], Holm and Schilling [12] and Childs and Wolfram [13]. Other more cursory accounts of individual vowel sounds are for example [14], [15] and [16]. The only acoustic study available on Bahamian vowels was conducted by Childs, Reaser and Wolfram [17], who analysed accommodation patterns in white and black speakers from Abaco, a small island to the northeast of New Providence.

The vowels in words like PRICE and MOUTH are generally described as fairly standard, high-front [aː - aː] and high-back [aʊ - əʊ, əʊ - əʊ] gliding diphthongs, respectively. Some authors have noted monophthongal productions of PRICE in pre-voiced contexts [13], exhibiting a pattern similar to that found in African American Vernacular English, or of both PRICE and MOUTH in more basilectal speech [14]. Holm [15], however, argues that this phenomenon is confined to cer-
tains islands and otherwise as rare in the Bahamas as it is in the rest of the Caribbean. At least for non-creolised Bahamian varieties, it has also been suggested that the allophones of MOUTH and PRICE follow the rules of Canadian Raising [18]. While in some Caribbean varieties and in Gullah, the nucleus in MOUTH may be raised and/or backed to [-a · o] irrespective of the phonetic context [4, 19, 20], no evidence of pre-voiceless raising has been observed in any of the Atlantic creole languages. Childs, Reaser and Wolfram [17] did not find raising in MOUTH for Bahamian speakers from Abaco island, voice-conditioned or otherwise.

It is the general aim of this study to analyse acoustically productions of MOUTH and PRICE by urban BahC speakers and, thus, confirm or reject prior impressionistic accounts as to their typical realisations. Moreton and Thomas’ notion of Asymmetric Assimilation [1], though essentially a diachronically-based model, can be extrapolated to formulate hypotheses that apply to socially conditioned synchronic variation within a speech community, where prestigious standard productions are associated with (nearly) identical diphthongs in pre-voiced and pre-voiceless contexts: Voice-conditioned spectral differences will be found in the production of all speakers. If glide weakening or pre-voiceless raising is exploited as a social marker in BahC, non-standard productions associated with the speech of lower social classes will be affected more than productions by speakers of higher social classes closer to the standard. If pre-voiceless raising is an established social marker in BahC, social variation will be more pronounced in the spectral pattern of the perceptually prominent nucleus than in that of the glide.

2. Methods

2.1. Speakers and data

The speakers analysed are a subset drawn from the tape-recordings of free conversations with 20 Nassauvians that Stephanie Hackert conducted in 1997/98 as part of her research on past temporal reference in urban Bahamian Creole [21]. In order to assign participants to different social classes, a classification scheme based on occupation was adopted, originally devised for the Jamaican context [22]. The result is a broad three-way distinction: the ‘middle strata’, the ‘petite bourgeoisie’ and the working class. The selection of speakers for the present study was mainly driven by the availability and quality of recordings. The final sample contained ten working-class speakers and six middle-strata or petite-bourgeoise speakers (all female), who were combined to form the category of ‘higher-class’ speakers. All participants were black and had lived in Nassau for most of their lives. The age of the participants ranged from 25 to 70 at the time of recording and the distribution of age groups across social classes was approximately equal (mean_{higher}=51, mean_{working}=42).

2.2. Data analysis procedure

The recordings were digitised at 20 kHz. While the length of the original recordings varied considerably, a section of 60 to 90 minutes per speaker was arbitrarily selected for analysis. Monosyllabic or disyllabic word tokens of MOUTH and PRICE in VC contexts were manually marked and extracted. Where possible, at least 10 tokens were collected for each lexical set, with no more than two tokens of the same lexical item. Pre-nasal and pre-pausal contexts were avoided as well as tokens following /t/ or semivowels. Tokens followed by /t,d/ in a potential flapping context, where intervocalic /t,d/ precede an unstressed vowel in the same word, were excluded regardless of whether the consonant was actually flapped or not. This procedure yielded a total of 382 tokens, 118 for MOUTH and 264 for PRICE.

Segmentation and acoustic measurements were carried out using the Praat software [23]. The onset of a vowel was determined from the waveform and set at the first regular pitch pulse. The offset was marked at the last regular pitch pulse or at the point at which the complex wave smoothed. In ambiguous cases, the spectrogram was consulted to determine the point of diminishing F2. Measurements were made in a semi-automatic procedure using linear predictive coding (LPC). Parameters were initially set to the Praat default, though the LPC order was adjusted in some cases to improve formant readings. F1 and F2 measurements were taken at 10% intervals through the vowel and confirmed or corrected by visual inspection. Formant estimates were then saved to file and further processed in the R environment [24].

In order to avoid the problematic inter-speaker comparison of acoustic data in raw Hertz values, formant frequencies were converted to Bark and the effect of variable postvocalic voicing was represented as proportional change in a procedure adapted from Moreton and Thomas [6, 7, 1]. For each speaker and formant, log ratios were calculated at nucleus and glide between the mean of all pre-voiceless tokens and the mean of all pre-voiced tokens. Nucleus and glide were defined as the point of maximum F1 between 20% and 40% into the vowel and as the point of maximum/minimum F2 between 80% and 90%, respectively. To allow for more immediate comparisons of diphthongality, the extent of gliding movement was quantified as the Euclidean distance between nucleus and glide for each token and subjected to the same ratio calculations as the other measures: For each speaker, a log ratio was calculated between the mean Euclidean distance of all pre-voiceless tokens and the mean of all pre-voiced tokens. Vowel normalisation proper was applied only for illustrative purposes (see figure 1), using a formant intrinsic centroid-based technique developed by Watt and Fabriani [25, 26].

3. Results

Figure 1 displays all measurement points, averaged across higher- and working-class participants. Superscript zero in P0 and M0 indicates pre-voiceless contexts of PRICE and MOUTH, respectively. Peripheralisation of the glide in pre-voiceless MOUTH and PRICE appears to be characteristic of both social groups, while raising of the nucleus is more pronounced among working-class speakers.

Figure 1: Vowels plots: Median values of measurement points for higher- and working-class speakers.
The pattern of contextual and social variation observed in figure 1 is reflected in the results of the log ratio calculations of formant frequencies at nucleus and glide (figure 2). Negative values indicate raising (F1) or backing (F2) and positive values lowering (F1) or fronting (F2) in pre-voiceless relative to pre-voiced contexts. The closer the log ratio is to zero, the smaller is the effect of voicing on the spectral pattern of the preceding diphthong. For higher-class speakers, log ratios of F1 and F2 in nuclear position tend to be close to zero. A median log ratio of −0.067 for F2 in MOUTH, which roughly corresponds to a 7% decrease in F2 in pre-voiceless contexts, represents the maximum deviation from zero for this group and position. With increased proximity to the following consonant, the effect of voicing becomes more prominent at the glide, where median log ratios range from 0.065 (F2 in PRICE) to −0.164 (F1 in PRICE).

Working-class speakers display the reverse pattern. While median log ratios between 0.049 (F2 in PRICE) and −0.195 (F1 in PRICE) at the glide indicate a similar degree of pre-voiceless peripheralisation than that exhibited by higher-class speakers, the spectral difference becomes even more pronounced in nuclear position. With a median log ratio of −0.211, F1 at the nucleus of MOUTH is the measure most affected by the presence or absence of voicing of the following consonant. However, all measures of MOUTH exhibit considerable inter- as well as intra-speaker variation, coinciding with an overall scarcity of tokens. For this reason, only log ratios for PRICE were subjected to statistical testing.

As can be observed in figure 1, the total extent of formant movement appears to be equal in pre-voiced and pre-voiceless contexts for higher-class speakers, while working-class speakers produce shorter diphthong trajectories in pre-voiceless tokens. If, however, the Euclidean distance between nucleus and glide is determined for each token separately and only afterwards averaged across speaker and class (see table 2), a surprising pattern emerges, which was masked in figure 1. For higher-class speakers, PRICE is considerably more diphthongal in pre-voiceless contexts. This is indicated by a statistically significant mean log ratio of 0.239, which corresponds to an increase of approximately 27%. Working-class speakers exhibit the expected pattern of less formant movement in pre-voiceless MOUTH and PRICE, which, however, failed to reach significance in the t-test performed on PRICE due to extensive inter-speaker variation.

Table 1: F1 and F2 log ratios (vl/vd) at nucleus and glide in PRICE. Summary data and results of t-tests (mean≠0)

<table>
<thead>
<tr>
<th>Class</th>
<th>Position:Fn</th>
<th>mean</th>
<th>sd</th>
<th>t(df)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Higher</td>
<td>Nucleus:F1</td>
<td>−0.03</td>
<td>0.02</td>
<td>−3.4(4)</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td></td>
<td>Nucleus:F2</td>
<td>0.04</td>
<td>0.03</td>
<td>2.7(4)</td>
<td>&gt;0.05</td>
</tr>
<tr>
<td></td>
<td>Glide:F1</td>
<td>−0.15</td>
<td>0.04</td>
<td>−8.5(4)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>Glide:F2</td>
<td>0.06</td>
<td>0.02</td>
<td>5.9(4)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Working</td>
<td>Nucleus:F1</td>
<td>−0.15</td>
<td>0.07</td>
<td>−7.3(9)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Nucleus:F2</td>
<td>0.08</td>
<td>0.04</td>
<td>6.7(9)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Glide:F1</td>
<td>−0.19</td>
<td>0.07</td>
<td>−8.3(9)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Glide:F2</td>
<td>0.05</td>
<td>0.02</td>
<td>7.9(9)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Table 1 shows summary data and results of two-sided t-tests computed for each formant, position and social class to assess whether the effect of voicing as reflected in the log ratios is statistically significant. Although the results are more robust for working-class speakers, the t-tests confirm that both social groups display voice-conditioned spectral differences in the glide as well as the nucleus of PRICE. All measures showed to be significant, except for nuclear F2 produced by higher-class speakers, which reflects group-internal variation across zero.

One participant out of five revealed a negative mean log ratio of −0.0002.

In order to determine whether the distribution and extent of spectral differences varies significantly across social class, two separate repeated-measures ANOVA [27] were performed on the log ratios of F1 and F2 with position (nucleus, glide) as within-speaker and social class (higher, working) as between-speaker variable. For F1, the effects of both class (F(1,13)=10.2, p<0.01) and position (F(1,13)=10.3, p<0.01) reached significance. For F2, the test revealed a significant interaction between social class and position (F(1,13)=5.1, p<0.05). While post-hoc t-tests traced back the class difference found for F1 to the nucleus (t(11.6)=−5.4, p=0.001) rather than the glide, they did not show any significant individual differences for F2.

Table 2: Log ratios (vl/vd) of Euclidean distance between nucleus and glide: Summary data and results of t-tests (mean≠0)

<table>
<thead>
<tr>
<th>Class</th>
<th>Lexical set</th>
<th>mean</th>
<th>sd</th>
<th>t(df)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Higher</td>
<td>PRICE</td>
<td>0.24</td>
<td>0.05</td>
<td>10.9(4)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>MOUTH</td>
<td>0.02</td>
<td>0.25</td>
<td>−</td>
<td></td>
</tr>
<tr>
<td>Working</td>
<td>PRICE</td>
<td>−0.12</td>
<td>0.27</td>
<td>−1.4(9)</td>
<td>&gt;0.05</td>
</tr>
<tr>
<td></td>
<td>MOUTH</td>
<td>−0.05</td>
<td>0.57</td>
<td>−</td>
<td></td>
</tr>
</tbody>
</table>

Since the data are unbalanced with respect to the participants’ gender, all higher-class speakers being female, the ques-
tion arises whether the differences observed so far may rest in part on this imbalance. If only working-class speakers are considered, female speakers show a somewhat greater effect of voicing on MOUTH tokens than males, while the reverse pattern emerges for PRICE (see figure 3). However, the results of repeated-measures ANOVA performed for PRICE on the F1 and F2 log ratios with position (nucleus, glide) as within-speaker and gender (male, female) as between-speaker variable did not show significant differences. As for the total extent of gliding movement, quantified as the Euclidean distance between nucleus and glide, male and female working-class participants revealed very similar patterns which clearly set them apart from the productions of the higher-class speakers in the sample. With mean log ratios close to zero for MOUTH (mean Fem = -0.07, mean Males = -0.03), voicing of the following consonant does not appear to have a strong effect on the diphthongal quality of the vowel. Negative log ratios for PRICE in both genders (mean Fem = -0.12, mean Males = -0.12) indicate that, in contrast to the glide weakening produced by higher-class speakers, realisations were actually more diphthongal in pre-voiced than in pre-voiceless tokens.

4. Discussion

The analysis showed that all participants produced peripheralised offglides in pre-voiceless relative to pre-voiced contexts. This finding is important because it reflects the pattern observed for speakers of non-creolised varieties of English [6, 7], lending support to the claim that relative peripheralisation just before the VC transition may be universally correlated with the voice contrast. For higher-class speakers, voicing had a weaker effect, if any, on the nucleus, which is consistent with the view that the primary effect of voicing is on the glide while spectral differences in the nucleus arise indirectly via assimilation. In contrast, working-class speakers displayed a significant voice-conditioned spectral difference in diphthong nuclei, with the nuclear clear effect equal to or even exceeding the offglide effect. The social variation that the speakers exhibited in the extent to which nuclei were affected as opposed to glides was robust enough to be tested significant.

The overall extent of gliding movement also differed between the two social groups. Working-class speakers produced on average shorter trajectories for both diphthongs in pre-voiceless tokens. Combined with the evidence on nuclear spectral differences, these speakers exhibited a regular, voice-conditioned alternation that conforms to the rules of Canadian Raising. Higher-class speakers produced tokens of MOUTH that were on average equally diphthongal, while the gliding movement in PRICE was considerably shorter before voiced codas, even though none of the speakers actually produced monophthongal tokens.

From a necessarily synchronic perspective, it is difficult to explain the varying pronunciation norms in the urban BahC community. Canadian Raising could have been the preferred pattern in earlier, basilectal forms of BahC. With increased access to and pressure from standard English, upwardly mobile speakers would gradually adapt their pronunciation to reflect standard norms. Alternatively, accepting the fact that there is little evidence that BahC ever was a homogeneous language variety, Canadian Raising could have been a social marker of working-class speakers for some time and any spectral differences exhibited by higher-class speakers which cannot be explained by universal phonetic processes might then be attributed to the effect of long-term accommodation. Finally, pre-voiceless raising in MOUTH could be the result of standard English influences on earlier, more basilectal BahC. Non-fronted, raised productions of MOUTH with nuclei in the range of [aː] are common among Caribbean creoles and are also found in Gullah [11, 19, 20]. Based on the acoustic evidence of a Gullah speaker born in 1844, Bailey and Thomas [5] show that this was also an accepted realisation of MOUTH 150 years ago. It seems plausible, then, that when earlier Gullah was transported to the Bahamas about 200 years ago, Gullah speakers, who would eventually be referred to as BahC speakers, produced MOUTH as backed and raised [əʊʉ] in all contexts. With increased inter-racial contact and access to standard and other non-creolised forms of English in the 20th century, an allophonic distribution could have emerged by reallocation as the result of dialect contact [18, 3, 28]. All of the described scenarios are, of course, merely based on speculation. In the absence of diachronic data, no definite answer can be given at present as to why and how the alternation took root in the Bahamian context.

5. Conclusions

Whatever the origin of Canadian Raising in BahC may be, the fact is that voice-conditioned alternation best accounts for the observed pattern and that it is exploited as a social marker. Working-class speakers of both genders exhibited considerable raising and fronting/backing of the nucleus in pre-voiceless PRICE and MOUTH, whereas higher-class speakers did not. The variation found in the data for the present study is not linguistically contrastive, but the significant differences between social classes show that it cannot be considered merely a consequence of universal phonetic processes.

6. References

