



Phonological processes of vowels from orthographic to pronounced words in the Buckeye Corpus by sex and age groups*

Byunggon Yang**

Abstract

This paper investigated the phonological processes of monophthongs and diphthongs in the pronounced words present in the Buckeye Corpus and compared the frequency distribution of these processes by sex and age groups to provide a clearer understanding of spoken English to linguists and phoneticians. Both orthographic and pronounced words were extracted from the transcribed label scripts of the Buckeye Corpus using R. Next, the phonological processes of monophthongs and diphthongs in the orthographic and pronounced labels were tabulated using R scripts, and a frequency distribution by vowel process types, as well as sex and age groups, was created. The results revealed that 95% of the orthographic words contained the same number of syllables, whereas 5% had different numbers of vowels, thereby proving that speakers tend to preserve vowels in spontaneous speech. In addition, deletion processes were preferred in natural speech. Most vowel deletions occurred with an unstressed syllable. Chi-square tests were performed to calculate dependence in the distribution of phonological process types for male and female groups and young and old groups. The results showed a very strong correlation. This finding indicates that vowel processes occurred in approximately the same pattern in natural and spontaneous speech data regardless of sex and age, as well as whether or not the vowel processes were identical. Based on these results, the author concludes that an analysis of phonological processes in spontaneous speech corpora can greatly enhance practical understanding of spoken English.

Keywords: phonological processes, vowels, sex, age, orthographic, pronounced, Buckeye Corpus

1. Introduction

Speech varies according to the anatomical size of the vocal tract or vocal folds and speakers' linguistic and non-linguistic factors such as dialect and emotion, all of which may influence acoustic measurements (Yang, 1990, 1996). Moreover, informally recorded spontaneous speech data contain various productions of the same

words or phrases, a fact that English learners encounter when listening to or conversing with native or nonnative English speakers. Spontaneous speech data would be useful to train intermediate or higher level learners to achieve nativelike fluency, particularly with respect to learning phonological processes such as sound deletions or insertions during rapid speech without compromising the intended message. Controlled speech, by contrast, does not feature such

* This work was supported by 2-year Research Grant of Pusan National University.

** Pusan National University, bgyang@pusan.ac.kr

Received 30 April 2018; Revised 5 June 2018; Accepted 8 June 2018

© Copyright 2018 Korean Society of Speech Sciences. This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

processes. Most phonetic research has been based on carefully controlled speech data generated in a laboratory environment (Fosler-Lussier *et al.*, 2005). Researchers could control each factor separately and observe detailed interactions among several factors. The ensuing findings have contributed greatly to identifying which phonological processes occur in natural speech, but additional studies of natural phonetic processes are nevertheless warranted.

Recently, many corpora have been built to make studies of natural speech possible (see Durand *et al.*, 2017, for a review). One of the large corpora (76 speakers) frequently cited in phonetics research was developed by Peterson & Barney (1952). TIMIT started with 30-minute speech fragments in the initial stage, expanding later to 6,300 sentences from 630 speakers' production of ten target sentences (Garofolo *et al.*, 1993). The Buckeye Corpus collected 40 speakers' spontaneous speech, which was orthographically transcribed and phonetically labeled (Pitt *et al.*, 2007). Thirty-six hours of urban French corpora exist (Cappeau *et al.*, 2004). The Kiel Corpus of Spontaneous Speech consists of eight hours of read and spontaneous German (Kohler *et al.*, 2017). In Korea, the Seoul Corpus was created to match the Buckeye Corpus in the number of the participants and the sophistication of phonetic transcriptions (Yun *et al.*, 2015).

Yang (2012) investigated the Buckeye Corpus to determine the most frequent consonants and vowels from 2,638,882 phonetic symbols and compared the findings with previous results on language universals (Maddieson & Disner, 1984; Maddieson, 1997). He reported a nonlinear distribution of the segmental inventory and a general trend in the massive reduction from the dictionary transcriptions to actual productions. The orthographic dictionary symbols decreased 38.2%. His research focused on the frequency distribution of both orthographic and pronounced phonetic labels but did not endeavor to determine if any preferences for a certain set of vowels or consonants based on sex and age. Yang (2016b) investigated the phonological processes of monophthongs and diphthongs in the pronounced phrasal words of the Seoul Corpus by gender and age. The corpus contained 546,404 syllables in the pronounced labels. The results showed that 97% of the number of syllables in the orthographic and pronounced phrasal words were identical and that 90.5% of the peak vowels in the orthographic phrasal words were pronounced in the phrasal words. In the orthographic words, 65.8% of the vowels differed from the pronounced syllable structure. Statistical tests between gender and age groups revealed a significant dependence in the distribution of phonological process types and a very strong correlation. The study indicated a sociolinguistic trend that older females in their 40s produced the diphthong *yo* as *yv* at the end of the pronounced phrasal words more often than the male group did.

The current study investigated the phonological processes of monophthongs and diphthongs in pronounced words of the Buckeye Corpus by sex and age to provide linguists and phoneticians with a clearer understanding of spoken English. Specifically, the author extracted phonetic symbols from the Buckeye Corpus files using computer scripts and examined the phonological processes of English vowels. In addition, statistical tests were conducted to determine the frequency distribution of vowel processes by sex and age.

The results of this study can help English educators and teachers develop essential teaching materials on natural phonological processes of English vowels. Textbooks based on real speech could help English learners become more proficient as they learn the principles of phonological processes and apply them to listening and

speaking activities. Finally, the results can be compared with those of equivalent speech corpora cross-linguistically. Eventually, mosaic results of many languages may reveal universal patterns.

2. Method

2.1. The Buckeye Corpus

The Buckeye Corpus was built by researchers (Pitt *et al.*, 2007) at Ohio University. It consisted of individual recordings by forty speakers in Columbus, Ohio, which occurred through a spontaneous conversation with one of the two interviewers. The speakers were ten participants representing each sex and age group so that patterns specific to a group could be examined. The corpus offers both audio and text files of the speech as well as orthographically transcribed and phonetically labeled productions.

2.2. Data collection

Orthographic and pronounced phonetic symbols of spontaneous speech were collected to facilitate a comparison and analysis of the phonological processes between them. The procedure was as follows. First, 255 phonetically transcribed text files of the Buckeye Corpus were copied manually into four group folders by sex and age group (i.e., young males, old males, young females, and old females). Then, an R (R Core Team, 2018) script was written to create an integrated text file to which all the files within each folder were appended as character strings using the table writing function of the R library "reshape." The first eleven lines of each text file and the final line contained file header and file end information. Hence, these lines were excluded in the appending procedure. After checking the integrated file in Microsoft Excel, the author corrected file format errors and unrelated labels (for example, B_TRANS, SIL, VOCNOISE, UNKNOWN) after repeatedly sorting the data by each column. Nasalized vowels with the phonetic symbol *Vn* were included after removing the nasal consonant. For example, the symbols *aan* and *ihn* were changed to *aa* and *ih*, respectively. All syllabic *el*'s, *em*'s, and *en*'s were discarded for purposes of analysis. The total number of words analyzed was 283,522.

2.3. Data analysis

The corrected Excel file was imported onto R Studio, and another R script was created to trace the phonological processes from the orthographic to the pronounced phonetic symbols and to extract factors for statistical analysis. The procedure was as follows. First, the script initialized the list of 33 vowels and created a temporary matrix of eighteen columns. Phonetic strings of the orthographic and pronounced words were extracted along with the number and position of vowels in the words. The vowel processes were divided into the same or different orthographic and pronounced vowels and placed in corresponding columns to determine their frequency distribution. If the number of vowels in a given orthographic word precisely matched that of the corresponding pronounced word, then each position of the vowel was assigned to the matrix column in the same order. The entire word was also assigned to one of those columns to facilitate later visual checks. The same number of vowels in both orthographic and pronounced words accounted for most vowel processes. The sex and age variables were also added to the matrix columns. After extracting all pertinent information from each word, the matrix was appended to a final output file. If the

numbers were unequal, then all information except the pronounced symbols were saved onto the matrix for manual verification. Instances of syllables of some frequently used words were significantly reduced in the pronounced forms of the Buckeye Corpus. For example, the word *definitely* had an orthographic label of *d eh f ih n ih t l iy* but was pronounced as *d eh f ih en l iy*, *d eh f ah l iy* and *d eh fl iy*. The four vowels were reduced to two or three vowels. Here, the stressed first vowel retained its vowel qualities in the pronounced labels. It might be possible to create a script to match the vowels case by case if a script were meticulously organized. Instead, the author checked incongruent words manually because the size was manageable and because in so doing, general patterns could be observed. The steps included finding the corresponding syllables between orthographic and pronounced words and assigning them to appropriate columns in the Excel file. To test dependence between male and female groups and between young and old groups, Pearson's Chi-square value and correlation statistics were generated in R using the frequency data.

3. Results and Discussion

3.1. Vowel processes by syllable number and position

The total number of 283,522 words in the Buckeye Corpus had 365,002 vowels in the orthographic labels and 354,635 vowels in the pronounced ones. Of the orthographic words, 269,291 (95.0%) had the same number of vowels as did the pronounced words, with only 14,231 (5.0%) words differing between the two sets of words. This result proves that vowels tend to retain their quality in spontaneous speech. Among the many vowel processes, deletion occurred 12,204 times (4.3%, See Johnson, 2003 for discussions on massive deletion), nearly six times as often as insertion, which occurred 2,027 times (0.7%). Thus, we can conclude that deletion processes are preferred in natural speech. The total number of vowels indicates that changes are not one-to-one. In other words, each count of the pronounced words for a given syllable number may involve both deletion and insertion processes. For example, the word *elementary* with five vowels was realized as *eh l ah m eh n t r iy*, and *eh l m eh n t r iy*, including the full pronunciation with five vowels. Thus, one vowel or two vowels of the word can be deleted, resulting in a three- or four-vowel word in spontaneous speech. Of course, the insertion of a vowel is involved with the word *job*, for instance, as *jh aa aa*, or *jh aa b ah*. Deletion mostly occurred in words containing more than two syllables; thus, the number of one-syllable words increased in spontaneous speech. Although the absolute number of deletions tends to decrease by vowel position, the proportion of total orthographic labels increased. For example, 586 of the fourth vowels were deleted, accounting for 17.3% of 3,389 orthographic labels. Additionally, 35.5% of the fifth vowels were not pronounced. The reason for this finding might be related to the fact that compared with insertion processes, deletion processes would be more effective in decreasing a speaker's efforts and accelerating their rate of speech. For example, the word *especially* was labeled *ih s p eh sh ah l iy* in the orthographic form but was produced as *s p eh sh* in the pronounced label. Nearly three of its four vowels were deleted. Another example is the three-syllable word *anything*, which was pronounced as one syllable, i.e., *th ih ng*. The two-syllable word *because* was produced mostly with *k z* and sometimes, with one consonant, *k*. We will return to the process in

the following section to generalize deletion in the case of unstressed syllables.

Additional analysis found that 165,078 (58.2%) orthographic words were pronounced with the same vowels in the orthographic labels, whereas 118,444 (41.8%) orthographic words were pronounced with different vowels.

Table 1. The number of vowels (V No) in orthographic (Ortho) and pronounced (Prono) words and the difference (Diff) between them

| V No | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|-------|---------|--------|--------|-------|------|-----|----|----|
| Ortho | 221,200 | 45,424 | 11,567 | 3,389 | 810 | 89 | 15 | 1 |
| Prono | 222,038 | 44,338 | 9,890 | 2,803 | 522 | 54 | 15 | |
| Diff | 838 | -1,086 | -1,677 | -586 | -288 | -35 | 0 | -1 |

In the table, most words from the spontaneous speech data consist of one-syllable words, and the number of vowels in the words decreases as the syllable becomes longer. Specifically, one-syllable words account for 78.3% of the data, followed by two-syllable words, which represent 15.6% of the data, and finally, three-syllable words, accounting for 3.5% of the pronounced labels. For example, one-syllable words such as *like*, *know*, *my*, and *one* have the same phonetic labels in both orthographic and pronounced words. This result might occur because speakers attempt to convey the meaning clearly to the interviewers by keeping the words intact without deleting or inserting vowels to facilitate their pronunciation. The reason that one syllable word remained intact may be related to the fact that its meaning may not be relayed clearly without it. Thus, we can infer that speakers tend to retain only one vowel in spontaneous speech. However, certain function words were reduced to one consonant (i.e., the word *has* was pronounced *s* or *z*), which might be easily identified based on the context. For another example, the function word *was* was pronounced as a reduced form *s* or as a two-syllable word, *uw ah z*. The number of vowels was reduced mostly in words with more than two syllables, as shown in <Table 1>. The longest eight-syllable word was *socioeconomically*. The seven-syllable word *homosexuality* occurred nine times.

In the case of the Seoul Corpus, Yang (2016b) reported that the most frequent syllable type in Korean was 3-syllable phrases (178,185), which accounted for 32.6% of the phrases, followed by two-, four-, and one-syllable phrases.

The Korean syllable phrases included particles indicating syntactical information such as subjects, objects, and others so that the number of syllables in the original word might decrease by at least one syllable. In addition, the Carnegie Mellon University Pronouncing Dictionary listed two-syllable words as the most frequent, followed by three- and one-syllable words (Yang, 2016a). Yang assumed that two-syllable words would better express different names for a greater number of objects and actions that are used daily, without overloading human memory space.

3.2. Orthographic to pronounced vowel processes

The contingency table of the orthographic symbols to pronounced vowels is given in <Table 2>. To ensure legibility of the crowded columns in the table, five cases of *ao>ao* for the example word *bought* in both the orthographic and pronounced forms were deleted. The rare vowel might be related to either the speakers' dialect or inappropriate transcriptions. One *ey* and two *ow*'s were realized as *ye* or *oa*, respectively, in the phonetic labeling. These cases were

Table 2. A contingency table of orthographic labels (Ortho) (columns) and pronounced phonetic symbols (Prono) (rows) of the Buckeye Corpus

| Prono\Ortho | aa | ae | ah | ao | aw | ay | eh | er | ey | ih | iy | ow | oy | uh | uw | null | Total |
|-------------|--------|--------|--------|-------|-------|--------|--------|--------|--------|--------|--------|--------|-----|-------|--------|-------|---------|
| aa | 10,920 | 47 | 1,197 | 4,821 | 50 | 35 | 108 | 487 | 5 | 67 | 12 | 113 | 12 | 20 | 4 | 48 | 17,946 |
| ae | 204 | 18,037 | 2,660 | 6 | 17 | 34 | 8,804 | 24 | 157 | 3,307 | 25 | 10 | 1 | 94 | 13 | 311 | 33,704 |
| ah | 475 | 90 | 38,258 | 283 | 16 | 76 | 2,234 | 1,091 | 88 | 8,150 | 79 | 113 | 5 | 963 | 136 | 3,726 | 55,783 |
| aw | 699 | 150 | 145 | 16 | 4,826 | 8 | 13 | | 1 | 3 | | 37 | 99 | | | | 5,997 |
| ay | 2,011 | 352 | 2,602 | 23 | 25 | 23,019 | 757 | 3 | 112 | 167 | 119 | 3 | 5 | 8 | 1 | 63 | 29,270 |
| eh | 31 | 67 | 2,487 | 11 | 4 | 22 | 19,532 | 834 | 176 | 2,052 | 40 | 19 | 2 | 66 | 17 | 250 | 25,610 |
| er | 9 | | 1,152 | 11 | | | 125 | 10,766 | 11 | 275 | 33 | 9 | | 50 | 6 | 248 | 12,695 |
| ey | 142 | 32 | 3,947 | 4 | | 42 | 1,630 | 53 | 13,883 | 1,224 | 143 | 17 | 8 | 63 | 14 | 199 | 21,401 |
| ih | 17 | 24 | 4,488 | 7 | 1 | 81 | 2,144 | 289 | 282 | 36,685 | 3,257 | 7 | 2 | 484 | 63 | 63 | 47,894 |
| iy | 15 | 17 | 4,889 | 4 | 1 | 76 | 436 | 300 | 221 | 4,825 | 29,070 | 20 | 1 | 172 | 87 | 602 | 40,736 |
| ow | 260 | 16 | 3,101 | 3,246 | 47 | 5 | 165 | 2,692 | 11 | 235 | 13 | 17,808 | 34 | 207 | 74 | | 27,914 |
| oy | 1 | | 6 | 1 | | | | | | | | 5 | 480 | | | | 493 |
| uh | 2 | | 274 | 3 | | 2 | 12 | 24 | 1 | 409 | 21 | 9 | | 3,165 | 133 | 72 | 4,127 |
| uw | 16 | 16 | 4,014 | 16 | 3 | 12 | 475 | 840 | 18 | 5,697 | 702 | 44 | | 625 | 10,838 | 907 | 24,223 |
| Sum | 14,802 | 18,848 | 69,220 | 8,452 | 4,990 | 23,412 | 36,435 | 17,403 | 14,966 | 63,096 | 33,514 | 18,214 | 649 | 5,917 | 11,386 | 6,489 | 347,793 |

deleted because of incorrect transcriptions, not in the list of the aligner's phonetic alphabet (See Table 2, Kiesling *et al.*, 2018). In addition, 1,062 *er>r*'s are not in the table, whose changed consonants will be examined in a future study. The syllables *el*, *em*, *en*, and *eng* and certain orthographic vowels were produced as the semivowels *w* and *y*; these cases were also removed from the list. If a nasalized vowel combined with a full phonetic vowel (i.e., *aan* or *eyn*), then those vowels were included in the frequency count after removing the nasal consonant. The number of removed symbols was 15,235. Nine monophthongs were chosen to display the frequency distribution of both the same and different vowels in <Figure 1>.

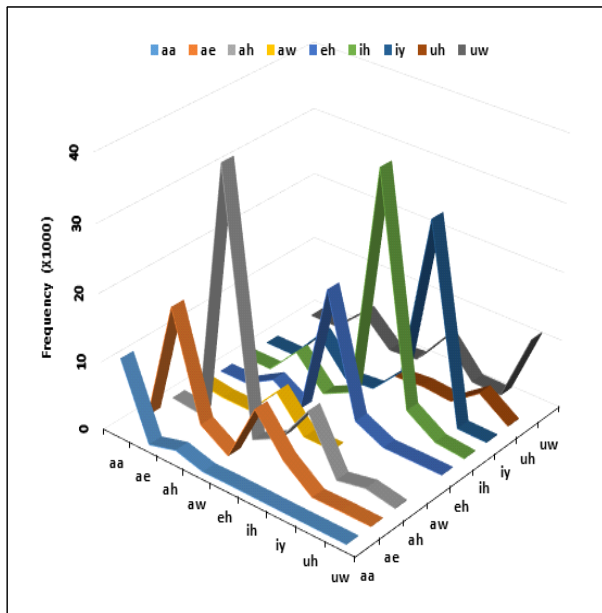


Figure 1. Frequency distribution of vowel processes in 3D format. The y-axis indicates the frequency of each vowel occurrence.

The total number of vowels in the table is 347,793, which represents 95.3% of all the vowels in the orthographic labels (see <Table 1>). The most frequently pronounced vowels were *ah* (69,220), followed by *ih* (63,096), *eh* (36,435), *iy* (33,514), and *ay* (23,412). The diphthong *ay* ranks the sixth most frequent (12,384) vowel for the pronoun *I*. The pronoun was produced mostly as *ay*, along with a few variations such as *ah*, *ao*, and *aw*. The back vowel

ah occurred slightly more often (68.6%) than did the back vowel *aa* (60.8%). The least frequently pronounced vowels were *oy* (649), *aw* (4,990), *uh* (5,917) and *ao* (8,452). The high back vowel *uw* (11,386) was the sixth least pronounced vowel.

The most common same vowel processes were *oy* (480 occurrences, which account for 97.4%), *er* (10,766, 84.7%), *aw* (4,826, 80.5%), *ay* (23,019, 78.6%), and *uh* (3,165, 76.7%), all of which are back vowels. The back vowel *uw* (10,838) had 44.8% of the same vowels. The average percentage of all back vowels with the same vowel processes is 72.9%. When the lowest *uw* is excluded, the average goes up to 76.4%. The average percentage for the five front vowels (*iy*, *ih*, *ey*, *eh*, and *ae*) is 68.5%. The lower rate of the front vowel processes might be related to the economy of production in spontaneous speech that relies on a smaller jaw movement. Wright (2003) reported that the vowels of words that people speak more frequently in their daily lives tend to be produced with reduced articulatory gestures; thus, vowel spaces decrease. Peripheral high vowels (*iy*, 76.6% and *uh*, 76.7%) show a higher percentage rate compared with peripheral low vowels (*ae*, 53.5% and *aa*, 60.5%). The speakers might have wanted to enunciate more clearly for the listeners by expanding their jaw opening. Interestingly, 44.8%, less than half the percentage of the high back rounded vowel *uw*, retained its vowel quality. This result might be explained by the fact that the speakers cannot further expand their movements upward to pronounce the higher vowel. In addition, the peripheral tense vowel *uh* showed a considerably lower rate of occurrence (3,165, 44.8%) than that of its lax counterpart *uw* (10,838, 76.7%). Some tense vowels were realized as *ih* (23.5%) and *ah* (16.6%). Despite the high percentage rate, lax vowels (11,386) occur nearly twice as often as tense vowels (5,917) do. In this case, the frequency must be considered along with the percentage in the description.

This trend may be applicable to the other tense and lax vowel pairs. The tense and lax vowel pair *iy* and *ih* exhibit an alternation of nearly double the ratio as the lax to tense ratio. Of the orthographic *iy*'s, 6.8% were produced as the lax counterpart *ih*, while 11.8% of *ih*'s became tense. Thus, one could surmise that in spontaneous speech, tense vowels tended to be pronounced as their lax counterparts.

Although the number of occurrences was low, the diphthong *oy* (493, 97.4%) was realized as the same one, followed by *ay* (23,019, 78.6%), *ey* (13,883, 64.9%) and *ow* (17,808, 63.8%). In addition, 6,489 vowels (1.9% of the total vowels) of the orthographic words

were deleted in the pronounced forms (Null). A negligible number of vowels were added to the pronounced words. The number of null vowels and one-syllable words increased to 2,835 and 838, respectively, in the pronounced forms. Most null vowel words were interjections such as *mm* as *em* or *hum-hum* as *hh em em*. As described in the previous section, the syllabic *em* was not classified as a full vowel; thus, it was excluded from the table.

From detailed observations of the final data, we found that the vowels in a stressed syllable typically remained in the pronounced form. For example, the second vowels of the words *police* and *remember* were always pronounced in spontaneous speech. The function word *about* has many different forms such as *ah*, *b aa*, *b aa dx*, *b ah*, *b ah dx*, *b ah t*, *b aw*, *b aw dx*, *b aw t*, and *b aw tq*. Here, all the pronounced forms have the stressed second vowel, possibly because this stressed syllable is crucial to identifying the intended word. We also found that a suffix in the final syllable position in a word with more than three syllables tends to be preserved to facilitate clear communication. For example, the words *definitely* and *especially* were pronounced as *d eh f l iy* and *s p eh sh l iy*. Deleting the final suffixes in these cases may cause listeners to wait for a noun to follow. Overall, we can generalize that speakers tend to preserve those vowels that are important to convey their thoughts but to change or delete vowels that are relatively unimportant to communicate their intended message. Additional perceptual studies on spontaneous speech would be desirable, specifically studies involving a detailed analysis of the role of each vowel in a word.

Other cases account for a small proportion of observations, but nevertheless merit attention. The orthographic vowel *ey* (3,947, 18.4%) was pronounced as *ah*. The word *butt* is transcribed as *b ah t*. Many speakers produced the indefinite article *a* (*ey* in the orthographic label) as *ah* including the contracted form *they'll* as *ah l*, and *dh ah*. The article is generally produced as *ey* when the speaker is emphasizing the single count of a particular type. Otherwise, it is produced as *schwa*. The word *schwa* in Hebrew means nothing or no vowel, referring to its short and indistinct vowel quality (Bauman-Waengler, 2009). The aligner's phonetic alphabet does not contain *schwa*. *Schwa* is the commonest vowel in unstressed syllables (Davenport & Hannahs, 1998). The only *schwa* label is used with the consonant *r*, i.e., *er*. That broad transcription might have limited the phonetic labels of the Buckeye Corpus. This limitation might owe to the high frequency (69,220) of *ah* in <Table 2>. A similar limitation is apparent in the orthographic vowel *aa*. Many *aa*'s (4,821, 26.9%) were pronounced as *ao*. Two hundred nineteen cases of the word *call* were transcribed as *k aa l* in the orthographic form and as either *k ao* or *k ao l* in the pronounced form. Moreover, the orthographic vowel *uw* (5,697, 23.5%) was pronounced as *ih*. The reason why this vowel has a higher frequency may be because many speakers produced the frequent word *you* (*y uw* in the orthographic labels). The vowel was pronounced as *ih* mostly in such combinations as *you've*, *you're*, and *you'd*. Thirty-four *yourself*'s were represented as *y ow r s eh l f*, whereas six *yours*'s were represented as *y ow r z*. We also observed cases of inappropriate vowel labels in such frequently occurring words as *eleven*, *election*, *expect*, and *experience*, whose first vowel was labeled *eh* in the orthographic form but *ih* in the pronounced form. Thus, the proportion of vowel processes may require further comparison with additional varieties and rigorous checks before a general trend can be discerned.

3.3. Sex- and age-group comparison

The selected distribution of the vowel processes is provided in <Table 3>. Instead of listing all the cases, twenty-six topmost vowel processes in descending frequency list (thirteen vowel processes each from the same and different groups) were chosen to test the dependence of sex and age to facilitate a balanced comparison. The thirteen vowel processes account for 298,729 vowel processes, or 81.8% of the total vowel processes (365,002), while vowel processes from the young and old groups account for 298,738 vowel processes, or 81.9%. Notably, the sex group difference in the same vowel processes was nearly negligible, accounting for 6,374 vowel processes, 1.8%. A difference in the same vowel processes owing to age was observed in approximately 19,573 instances, or 5.4% of the vowel processes. A difference in those vowel processes could be attributed to sex in 2,869 instances, or 0.8%, and to age in 4,951 instances, or 1.4%.

Table 3. Major frequency distribution of vowel process types according to the same or different orthographic and pronounced forms

| Types | Processes | Male | Female | Old | Young |
|------------------|-----------------|---------|---------|---------|---------|
| Same vowels | <i>aa>aa</i> | 5,700 | 5,220 | 5,767 | 5,153 |
| | <i>ae>ae</i> | 8,839 | 9,198 | 9,504 | 8,533 |
| | <i>ah>ah</i> | 20,831 | 17,427 | 21,293 | 16,965 |
| | <i>aw>aw</i> | 2,689 | 2,137 | 2,650 | 2,176 |
| | <i>ay>ay</i> | 11,297 | 11,722 | 11,390 | 11,629 |
| | <i>eh>eh</i> | 10,193 | 9,339 | 10,882 | 8,650 |
| | <i>er>er</i> | 5,822 | 4,944 | 6,127 | 4,639 |
| | <i>ey>ey</i> | 7,205 | 6,678 | 7,853 | 6,030 |
| | <i>ih>ih</i> | 18,634 | 18,051 | 19,450 | 17,235 |
| | <i>iy>iy</i> | 14,770 | 14,300 | 15,621 | 13,449 |
| | <i>ow>ow</i> | 8,491 | 9,317 | 9,896 | 7,912 |
| | <i>uh>uh</i> | 1,473 | 1,692 | 1,757 | 1,408 |
| | <i>uw>uw</i> | 5,642 | 5,187 | 6,000 | 4,838 |
| Total | | 121,586 | 115,212 | 128,190 | 108,617 |
| Different vowels | <i>aa>ao</i> | 2,666 | 2,155 | 2,805 | 2,016 |
| | <i>ae>eh</i> | 5,009 | 3,795 | 4,722 | 4,082 |
| | <i>ah>ih</i> | 3,573 | 4,577 | 3,935 | 4,215 |
| | <i>ey>ah</i> | 2,248 | 1,699 | 2,196 | 1,751 |
| | <i>ih>ah</i> | 2,555 | 1,933 | 2,432 | 2,056 |
| | <i>ih>iy</i> | 1,654 | 1,603 | 1,667 | 1,590 |
| | <i>iy>ah</i> | 2,738 | 2,151 | 2,803 | 2,086 |
| | <i>iy>ih</i> | 2,466 | 2,359 | 2,481 | 2,344 |
| | <i>ow>ah</i> | 1,747 | 1,354 | 1,508 | 1,593 |
| | <i>ow>ao</i> | 1,819 | 1,427 | 1,784 | 1,462 |
| | <i>ow>er</i> | 1,414 | 1,278 | 1,455 | 1,237 |
| | <i>uw>ah</i> | 2,048 | 1,966 | 2,311 | 1,703 |
| | <i>uw>ih</i> | 2,463 | 3,234 | 3,342 | 2,355 |
| Total | | 32,400 | 29,531 | 33,441 | 28,490 |

To ascertain statistical dependence from the table, Pearson's Chi-square tests were conducted on the frequency distribution of various sex and age groups. First, the statistical result on the sex group from both the same and different vowel processes was significant at $p < .05$ level ($\chi^2 = 1,127.1$, $df = 25$). This result indicates that the male and female groups were strongly dependent in the frequency distribution. Nearly the same significant result was obtained between young and old groups at $p < .05$ level ($\chi^2 = 689.9$, $df = 25$), which also showed a strong dependence in the frequency

distribution. The paired correlation between the two sex groups was significantly strong with $r=0.989$, as was that of the two age groups, whose $r=0.992$. Second, further detailed analyses separating the sex and age factors from the factor of same or different vowel processes were conducted to examine dependence. The statistical results for the same vowel processes were significant at $p<.05$ with respect to sex ($\chi^2=449.2$, $df=12$), $r=0.987$, and with respect to age ($\chi^2=390.5$, $df=12$), $r=0.988$. The statistical results for the different vowel processes were also significant at $p<.05$ with respect to sex ($\chi^2=660.0$, $df=12$), $r=0.822$, and age ($\chi^2=298.9$, $df=12$), $r=0.926$. The correlation coefficient of the sex group was slightly lower for different vowel processes than for other vowel processes, which might be caused partly by the relatively wider deviation of the male and female frequencies. The table indicates that $ah>ah$ showed a visible difference between the sex groups. In the different vowel processes, the female processes $ae>eh$ and $ah>ih$ had an opposite distribution. The age group comparison showed a moderate difference in the distribution of the same vowel process $ah>ah$ and the different vowel process $uw>ih$. Yang (2017) reported that the Korean female group in their forties changed the diphthong yo to yv at the end of the pronounced phrasal words more often than the male group did in the Seoul Corpus. If we pursue the sociological trait further after applying narrower transcriptions and dividing the four groups of the Buckeye Corpus by age, we may be able to identify a sociological trend in terms of either sex or age. Currently, however, the low frequency and transcription errors prevent further analysis. Thus, we tentatively offer the general conclusion that regardless of sex and age or whether the vowel processes are identical or different, the vowel processes occurred in nearly the same pattern in the natural spontaneous speech data. The results appear comparable because the corpus contained sufficient data. With less biased data, the vowel processes might have been mostly not significantly dependent.

4. Summary and Conclusion

The current study examined the phonological processes of monophthongs and diphthongs in pronounced words of the Buckeye Corpus and compared the frequency distribution of same and different vowel processes by sex and age groups based on their orthographic labels. Two hundred and fifty five files of spontaneous speech were processed to trace the phonological processes of monophthongs and diphthongs using R scripts, and general trends in the processes were described. Chi-square tests were conducted to calculate statistics for the vowels according to sex and age groups. The results revealed that most of the number of syllables in the orthographic and pronounced phrasal words were identical, while a fraction differed in the number of vowels. This finding indicates that speakers tend to preserve vowels even in spontaneous speech. More deletion processes than insertion ones were observed. Most vowel deletions occurred with an unstressed syllable. We interpreted that the speakers might have preserved those vowels that were important for conveying their thoughts to listeners but that they might have changed or deleted vowels that played relatively marginal or redundant roles in communication. The results of the statistical tests revealed significant dependence in the distribution of phonological process types of male and female or young and old groups and indicated a very strong correlation. Regardless of sex and age or whether the vowel processes were identical or different, the vowel

processes occurred according to approximately the same pattern in the natural spontaneous speech data. Given these results, this paper concludes that an analysis of phonological processes in spontaneous speech corpora can greatly enhance the practical understanding of spoken English.

During analysis, inappropriate and inconsistent transcriptions were observed for both orthographic and pronounced labels, thus requiring additional checks and corrections. Further studies would be desirable for comparing data from other European languages to discover universal patterns in phonological processes.

References

- Bauman-Waengler, J. (2009). *Introduction to phonetics and phonology: From concepts to transcription*. Boston: Pearson.
- Cappeau, P., Sabio, F., Bilger, M., & Chanet, C. (2004). *Autour du Corpus de référence du français parlé*. Aix-en-Provence: Publications de l'Université de Provence.
- Davenport, M., & Hannahs, S. J. (1998). *Introducing phonetics and phonology*. London: Arnold.
- Durand, J., Gut, U., & Kristoffersen, G. (2017). *The Oxford handbook of corpus phonology*. Oxford: Oxford University Press.
- Fosler-Lussier, E., Dilley, L., Tyson, N., & Pitt, M. (2005). The Buckeye corpus of speech: Updates and enhancements. Retrieved from http://www.academia.edu/14460741/The_buckeye_corpus_of_speech_updates_and_enhancements on March 1, 2018.
- Garofolo, J. S., Lamel, L. F., Fisher, W. M., Fiscus, J. G., Pallett, D. S., Dahlgren, N. L., & Zue, V. (1993). TIMIT. Acoustic-phonetic continuous speech corpus. Retrieved from <https://catalog.ldc.upenn.edu/Ldc93s1> on March 1, 2018.
- Johnson, K. (2003). Massive reduction in conversational American English. In K. Yoneyama & K. Maekawa (Eds.), *Spontaneous speech: Data and analysis, Proceedings of the 1st Session of the 10th International Symposium* (pp. 29-54). Tokyo, Japan.
- Kiesling, S., Dilley, L. & Raymond, W. D. (2018). The Variation in Conversation (ViC) project: Creation of the Buckeye Corpus of conversational speech. Retrieved from <https://buckeyecorpus.osu.edu/BuckeyeCorpusmanual.pdf> on March 1, 2018.
- Kohler, K. J., Peters, B., & Scheffers, M. (2017). The Kiel Corpus of spoken German: Read and spontaneous speech. Retrieved from <https://www.isfas.uni-kiel.de/de/linguistik/forschung/kiel-corpus/the-kiel-corpus-of-spoken-german-read-and-spontaneous-speech> on March 1, 2018.
- Maddieson, I., & Disner, S. F. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.
- Maddieson, I. (1997). Phonetic universals. In W. J. Hardcastle, J. Laver, & Gibbon, F. E. (Eds.), *The handbook of phonetic sciences* (pp. 619-639). Oxford: Blackwell.
- Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America* 24(2), 175-184.
- Pitt, M. A., Dilley, L., Johnson, K., Kiesling, S., Raymond, W., Hume, E., & Fosler-Lussier, E. (2007). *Buckeye corpus of conversational speech* (2nd release). Columbus, OH: Department of Psychology, Ohio State University. Retrieved from <https://buckeyecorpus.osu.edu/> on March 1, 2018.
- R Core Team. (2018). R: A language and environment for statistical computing. Retrieved from <https://www.r-project.org/> on March 1, 2018.

- Wright, R. (2003). Factors of lexical competition in vowel articulation. In J. Local, R. Ogden, R. Temple, M. E. Beckman, & J. Kingston (Eds.), *Phonetic interpretation: Papers in laboratory phonology VI* (pp. 75-87). Cambridge: Cambridge University Press.
- Yang, B. (1990). *Development of vowel normalization procedures: English and Korean*. Unpublished doctoral dissertation. The University of Texas at Austin.
- Yang, B. (1996). A comparative study of English and Korean monophthongs produced by male and female speakers. *Journal of Phonetics*, 24(2), 245-261.
- Yang, B. (2012). Reduction and frequency analyses of vowels and consonants in the Buckeye Corpus. *Phonetics and Speech Sciences*, 4(3), 75-83.
- Yang, B. (2016a). Phoneme distribution and syllable structure of entry words in the CMU English Pronouncing Dictionary. *Phonetics and Speech Sciences*, 8(2), 11-16.
- Yang, B. (2016b). Phoneme distribution and phonological processes of orthographic and pronounced phrasal words in light of syllable structure in the Seoul Corpus. *Phonetics and Speech Sciences*, 8(4), 1-9.
- Yang, B. (2017). Phonological processes of vowels in pronounced phrasal words of the Seoul Corpus by gender and age groups. *Phonetics and Speech Sciences*, 9(2), 23-29.
- Yun, W., Yoon, K., Park, S., Lee, J., Cho, S., Kang, D., Byun, K., Hahn, H., & Kim, J. (2015). The Korean corpus of spontaneous speech. *Phonetics and Speech Sciences*, 7(2), 103-109.

• **Byunggon Yang**

English Education Department
 Pusan National University
 63Bungil 2 Pusantaehak-ro, Keumjung-gu
 Pusan 46241, Korea
 Tel: 051-510-2619
 Email: bgyang@pusan.ac.kr
 Homepage: <http://fonetiks.info/bgyang>
 Fields of interest: Phonetics, Phonology, English Education