



서울코퍼스의 성별·연령 집단별 말 어절 모음에 나타난 음운변동*

Phonological processes of vowels in pronounced phrasal words of the Seoul Corpus by gender and age groups

양 병 곤**

Yang, Byunggon

Abstract

This paper investigated the phonological processes of monophthongs and diphthongs in pronounced phrasal words of the Seoul Corpus by gender and age groups in order to provide linguists and phoneticians with a clearer understanding of the spoken Korean. Both orthographic and pronounced phrasal words were extracted from the transcribed label scripts of the Corpus using *Praat*. Then, phonological processes of monophthongs and diphthongs were tabulated using an *R* script after syllabifying the phrasal words into separate components. Results revealed that 97% of the number of syllables in the orthographic and pronounced phrasal words were the same while 65.8% showed difference in the syllable structure. 90.5% of the vowels in the orthographic phrasal words were realized in the pronounced phrasal words. A *Chi-square* test of independence was performed to obtain a significant dependence in the distribution of phonological process types of male and female groups along with a very strong correlation. Female group changed the diphthong *yɔ* into *yɥ* at the end of the pronounced phrasal words more often than the male group did. Age groups also showed a significant dependence in the distribution of phonological process types along with a very strong correlation. Females in the 40s produced the diphthong *yɥ* and made the vowel raising at the end of the pronounced phrasal words most often among the gender and age groups. From the results, this paper concludes that an analysis of phonological processes in light of syllable structure can contribute greatly to the understanding of the spoken Korean.

Keywords: phonological processes, vowels, gender, age group, syllable, orthographic, pronounced, Korean corpus

1. 서론

최근에 발표된 40여명의 서울화자가 자연스럽게 발화한 서울 코퍼스(Yun 외, 2015)는 한국어 연구에 매우 귀중한 자료로 평가된다. 서울코퍼스는 글 어절과 말 어절의 발음을 함께 전사하여 나타냄으로써 구어 언어 사용에 나타나는 음운변동에 대해서 구체적으로 살펴볼 수 있다. 또한 성별·연령별로 균형 있

게 피험자를 선정하여 자료를 제공하고 있어서, 이런 분류를 중심으로 언어습관이나 사회언어학적인 음운변동도 관찰할 수 있다. 하지만, 말 어절에서만 약 55만개에 달하는 음절이 포함된 대용량코퍼스를 처리하여 분석하려면 수작업으로는 불가능하고, 음운변동의 종류도 다양하여 아직까지 연구가 활발하지는 않다(강지은 & 공은정, 2016; 양병곤, 2016). 강지은 & 공은정(2016)은 서울코퍼스의 세대별 남녀 30여명의 발화자료

* 이 논문은 2017년도 부산대학교 인문사회연구기금의 지원을 받아 연구되었음

** 부산대학교, bgyang@pusan.ac.kr

Received 27 April 2017; Revised 26 May 2017; Accepted 13 June 2017

에서 모음의 단일지점 포먼트값과 궤적양상을 동시에 분석하여 살펴본 결과 모음 /ɔ, ɪ/ 의 포먼트가 기존의 연구(Han & Kang, 2013; Kang & Han, 2013)에서와 같이 /ɔ/의 F1이 하강하여 /ɪ/의 F1으로 겹쳐지면서 합류하는 양상이 아닌 모음체계 전반에 걸쳐 연쇄이동하는 것으로 해석했다. 기존의 연구에서는 모음의 합류 현상은 남성보다는 여성에게서 더 뚜렷하게 관찰되었다고 하는데, 강지은 & 공은정(2016)에서도 여성에게서 더욱 뚜렷하게 /ɪ/의 F2값이 상승하여 /—/쪽으로 전진조음을 하는 것을 보고했다. 양병곤(2016)은 서울코퍼스의 말 어절과 글 어절을 음절단위로 분석하여 전체적인 음소의 분포와 음운변동에 대해 살펴보았다. 그의 연구에 따르면 초성의 음소분포에서는 k0이, 중성에서는 단모음 aa가, 종성에서는 비음인 mn이 가장 많이 나타났고, 음절유형별로는 CV유형이 가장 큰 빈도수를 보였다고 한다. 또한 이중모음이 글 어절에서는 8.8%이었는데 말 어절에서는 15.3%로 2배나 증가되어 추가로 자세한 분석은 하지 않았지만 어말의 yo가 yu로 바뀐 것이 주된 이유로 보인다고 했고, w로 시작되는 이중모음 군이 단모음화 되는 현상도 나타났다고 한다. 서울코퍼스와 비슷한 규모와 수집 방식으로 구축된 백아이 코퍼스에서도 음운변동에 대한 연구는 매우 제한적이다(Johnson, 2003; Kim, 2015; Pitta *et al.*, 2005; Yang, 2012). Pitta *et al.*(2005: 90)에 따르면 일상생활에서 음운변동은 발화속도에 관계없이 자연스럽게 화자 내에서나 화자 간에 일어나며 예를 들어, went라는 단어가 [went], [wen], [wənt] 등으로 바뀌어 when이나 want 등으로 혼동될 만큼 적지 않게 왜곡되기도 한다고 지적하고 있다. 이러한 자연발화 코퍼스는 어떤 음운변동이 얼마의 빈도로 규칙적으로 나타나는지 확인해 볼 수 있다(Dalby, 1986; Jurafsky *et al.*, 2001; Labov, 1994). 또한 음운변동현상은 자동음성인식 시스템을 구축할 때 모델 설정에도 도움이 되고(Strik & Cucchiari, 1999), 세대별로 구축된 코퍼스를 살펴본다면 사회언어학적인 의미도 찾아 볼 수 있을 것으로 기대된다.

이 연구에서는 양병곤(2016)의 연구에서 지적한 국어이중모음의 음운변동 특징이 성별, 세대별로 어떻게 나타나는지 살펴 보기 위해 시작하였고, 분석하는 과정에서 서울코퍼스의 말 어절에 나타난 세대별 국어 단모음과 이중모음의 음운변동의 분포를 다음과 같은 연구문제를 중심으로 살펴보려고 한다.

1. 서울코퍼스의 말 어절에서 단모음과 이중모음은 어떤 음운변동 유형과 빈도수 분포를 보이는가?
2. 성별과 연령별로 나누었을 때 말 어절의 단모음과 이중모음은 어떤 음운변동 유형과 빈도수 분포를 보이는가?

이 연구의 결과는 한국어를 연구하는 언어학자나 음성학자들에게 국어 모음의 자연 발화에 나타난 음운변동 현상에 대한 기초자료를 제공하여 국어코퍼스 연구에 기여할 것으로 기대된다.

2. 자료 분석

자료 분석은 서울코퍼스의 글 어절과 말 어절을 분리하는 양병곤(2016)이 사용한 Praat(Boersma & Weenink, 2017) 스크립트를 보완하여 텍스트 파일을 만들었다. 성별·연령별로 나누어 음운변동을 살펴보기 위해서 이 연구에서는 각 파일을 읽어서 마침표 이하는 삭제한 newname\$란 문자열을 구해서 이름을 붙이는 다음의 스크립트를 추가하였다.

```
name$=Get string... i
dotpoint=rindex(name$,".TextGrid")-1
newname$=left$(name$,'dotpoint')
appendFileLine: newname63$, newname$," ", ortho$,
    " ", pronos$
```

이렇게 처리한 결과파일(out63all.txt)의 첫 부분은 다음과 같다.

```
Filename Dict Real
s01m16f1 nnee nnee
s01m16f1 c0ee c0ee
s01m16f1 ii-llxxmm-xxnn ii-llxx-mmxxnn
s01m16f1 ii-k0oo-yo ii-k0uu-yv
s01m16f1 aa aa
s01m16f1 yvll-yv-s0vvs0 yvll-llyv-s0vv
s01m16f1 s0aall ssaall
```

이어서 R을 이용해서 546,405줄로 된 결과파일(out63all.txt)을 불러와서 처리했다. 전체적인 처리과정을 설명하면, 먼저 자음과 모음의 목록을 컴퓨터 메모리에 등록하고, 10열로 된 빈 행렬을 만들어 말 어절과 글 어절을 분석한 결과를 입력할 준비를 했다. 이어서 글 어절과 말 어절을 음절별로 분할하여 처리한 뒤 만약 두 어절의 음절수가 같으면 총 음절수만큼의 행으로 된 행렬을 만들고, 글 어절을 먼저 초성, 중성, 종성의 음절 구성요소로 분석하고, 이어서 이에 대응하는 말 어절을 음절구성요소로 분석하여 각각의 구성요소들이 같은지 다른지를 판단하여 빈 행렬에 총 음절수와 해당 음절의 위치를 숫자로 기록했다. 예를 들어, 위의 결과파일의 다섯 째 줄에서 글 어절 ii-k0oo-yo와 말 어절 ii-k0uu-yv는 음절수가 각각 3개로 같기 때문에 총 음절수 3을 기록하고, 먼저 두 어절이 같은지 다른지를 분석하여 결과를 입력하고 이어서 각 음절의 구성요소를 비교해서 같은지 다른지를 판단했고, 해당 음절의 위치를 숫자로 기록하여 나중에 총 음절수에서 해당음절을 빼서 0이 되면 맨 끝음절에서 yo를 yv로 바꾸어 발음했다고 판단하는데 활용했다. 만약 두 어절의 음절수가 다르다면, 어절의 특정 음절이 삭제되거나 삽입 또는 변동이 일어나서 매우 다양하게 나타나기 때문에 우선 총 음절수와 말 어절의 음절만 음절구성요소로 분리하여 음절 위치를 숫자로 기록하고 엑셀에서 수작업으

로 음운변동을 처리했다.

엑셀에서의 수작업 과정을 간략하게 서술하면 다음과 같다. 먼저 R로 처리된 결과파일(outanalyzer)을 엑셀에 불러와서 순서를 나타내는 열과 파일이름을 텍스트나누기를 이용해서 참가자, 성별, 나이, 면담자의 열로 분할하고, 음절수가 같은 것과 다른 것들로 정렬한 뒤 음절수가 다른 것들은 어절별로 하나씩 수작업으로 분류를 하였다. 엑셀에서 수작업으로 처리하는데 도움이 되도록 말 어절의 특정 모음(oo)이 글 어절(ii-k0oo-yo)에 나타나 있는 위치를 확인하는 ISNUMBER(SEARCH(k0oo, ii-k0oo-yo))라는 두 개의 엑셀함수를 이용하여 결과를 TRUE, FALSE로 처리했다. 이 때 모음만으로는 검색이 되지 않기 때문에 말 어절에 나타난 초성과 결합한 음절과 종성과 결합한 음절을 각각 글 어절에서 찾아서 처리하게 하고 한쪽이라도 TRUE가 있으면 글 어절의 모음이 변동 없이 말 어절에 실현된 것으로 보고, 글 어절에 나타나지 않았으면 삭제되거나 변형된 것으로 보고, 글 어절의 어떤 모음이 말 어절에서는 어떻게 변했는지 하나하나 눈으로 확인해보고 수작업으로 입력해 나갔다.

마지막으로 이렇게 만든 음운변동 목록이 들어간 모음분석 결과파일을 R에 불러와 attach하여 열 이름을 변수로 이용할 수 있게 만든 다음, R에 들어있는 table함수를 이용해 조건별로 묶어서 음운변동의 종류별 빈도수를 구해 파일로 저장했다. 다른 연구자들이 참고할 수 있도록 구체적인 예를 들어보면, table(Peak[(30<age)&(age<39)&(sex=="m")&(syfinal==0)])은 나이가 30대인 남성화자들의 음절말의 음운변동의 종류와 빈도수를 요약해준다.

3. 분석 결과 및 논의

3.1. 전체 모음의 음운변동 분포

서울코퍼스의 말 어절의 모든 음절수는 546,404개였다. 양병곤(2016, 표 2)에서는 글 어절과 말 어절을 한꺼번에 대비하여 나타내었는데 총 555,690개의 글 어절의 음절 가운데 9,286개가 말 어절에서는 실현되지 않았기 때문에 실제 이 값을 빼면 이 연구에서 구한 546,404개와 일치한다. <표 1>은 말 어절의 음운변동 유형별로 살펴본 서울코퍼스의 음절수와 어절, 모음이 서로 같은 지, 다른 지의 빈도 분포를 보여준다.

<표 1>에 따르면 글 어절과 말 어절의 음절수가 같은 경우는 530,037개로 나타났으며, 다른 경우는 16,367개로 드러났다. 이것을 비율로 계산해보면 말 어절의 약 97%가 글 어절의 음절수를 그대로 유지했고 단지 3%만이 발음에서 다른 수의 음절로 변동되었다고 할 수 있다. 참고로 말 어절 각각의 총 음절수 분포를 분석한 결과 3음절로 된 것이 178,185개로 가장 많은 32.6%를 차지했으며, 다음으로 2음절 162,362개로 29.7%, 4음절 94,580개로 17.3%, 1음절 57,225개로 10.5%를 차지했다.

표 1. 서울코퍼스의 음절수, 어절, 중성별로 비교해 본 음운변동 빈도수 분포

Table 1. Frequency distribution of phonological processes compared by syllable number, phrases, or syllable peaks of the Seoul Corpus

Division	Same (%)	Different (%)
Syl. no.	530,037 (97.0)	16,367 (3.0)
Phrases	186,784 (34.2)	359,620 (65.8)
Syl. peaks	494,351 (90.5)	52,053 (9.5)

이어서 글 어절과 말 어절이 같은 지 다른 지를 분석해 본 결과, 186,784개가 같게 나타나 약 34.2%를 차지했고, 음운변동이 적용되어 달라진 경우는 359,620개로 65.8%나 된다. 또한 단모음과 이중모음으로 된 중성이 글 어절에서 말 어절로 그대로 실현된 경우는 모두 494,351개가 되고, 다르게 실현된 경우는 52,053개로 나타났다. 이 분포를 비율로 환산하면, 모음의 약 90.5%가 음운변동 없이 발화되었고, 약 9.5%가 음운변동을 일으킨 것으로 볼 수 있다. 결국 음운변동은 음절의 중심인 모음보다는 자음의 변화가 주된 역할을 했음을 알 수 있다. 앞으로 자음의 복잡한 음운변동에 대해서도 연구해 볼 필요가 있다. 이 연구에서는 좀 더 자세한 모음의 음운변동을 살펴보기 위해 중성인 모음을 단모음과 이중모음, 두 개의 다른 모음의 조합으로 분류하였다. 일부 두 개의 다른 단모음인 ii-aa나 uu-aa가 합쳐져서 이중모음 ya, wa 등으로 변동된 경우도 일부 있었다. 글 어절에서 말 어절로 동일한 유형의 단모음으로 실현된 것은 478,333개로 87.5%나 된다. 이러한 결과는 다른 나라 언어에 비해 국어의 단모음이 적절한 지각적인 거리를 유지하고 있어서 (Yang, 1996), 다소 안정적으로 말 어절에서 그대로 실현된 것으로 여겨진다. 앞으로 다양한 외국어 자연발화코퍼스에서 모음의 갯수에 따른 음운변동의 빈도수를 측정하여 언어간의 비교를 통해서 연구해 볼 필요가 있다. 한편, 이중모음은 동일한 유형으로 실현된 경우는 47,501개로 8.7%를 차지했고, 이중모음이 단모음으로 바뀐 경우는 19,515개로 3.6%, 반대로 단모음이 이중모음으로 바뀐 경우는 536개로 0.1%를 차지했다. 이 세 가지 유형이 전체의 99.9%나 차지하였으며, 나머지 0.1%는 단모음이나 이중모음 삽입, 두 개의 단모음이 이중모음으로 또는 그 반대로 실현되거나 모음삽입과 같은 현상이 나타났다. 이중모음을 단모음으로 변동시킨 경우가 3배 이상이나 많은 것을 보면 실제 대화에서 문맥에 따라 쉽게 의미가 전달되기 때문에 의미의 차이를 크게 가져오지 않은 경우에는 발음하기 편한 단모음으로 발화하는 경우가 더 많았던 것으로 여겨진다.

다음으로는 모든 글 어절에 있는 모음이 말 어절에서 어떻게 실현되었는지를 유형별 목록을 만들어 살펴보았다. 모두 327가지로 나타났는데 이 논문에서는 출현 빈도수가 높은 자료를 중심으로 단모음 이중모음 등의 유형별로 살펴보기로 한다. 먼저 글 어절에서 말 어절로 동일하게 실현된 모음은 aa로서 112,712개를 차지했다. 참고로 글 어절 모음 aa의 모든 개수는 114,205개인데 약 98.7%가 변동 없이 발음되었다고 할 수 있다. 양병곤(2016, 표 2)에 제시된 aa의 글 어절과 말 어절이 일치된

총 수는 114,548개로 들어가 있는데, 이 연구에서는 음운변동을 자세히 살펴보기 위해 말 어절을 기준으로 글 어절과 다른 경우에 수작업으로 확인하였고, 앞에서 언급한 것처럼 두 개의 단모음이 이중모음으로 바뀐(예, “노느라고”를 “노냐고”로 발음, ii-aa>ya, *기호 >는 변동을 나타냄) 경우에는 aa의 단일 목록에 포함시키지 않았기 때문에 약간의 수치상의 차이가 있음을 밝힌다. 말 어절의 aa에는 글 어절의 “영화를”을 “영아를”과 같이 말 어절에서 이중모음 wa를 단모음 aa로 바꾼 음운변동이 1,967개로 나타났다. 이러한 이중모음의 단모음화는 ya>aa가 53개, yv>vv가 685개, yo>oo가 287개, yu>는 20개로 나타났고, wE>EE가 3,149개, WE>EE가 740개, wi>ii가 538개, wv>vv가 3,161개, xi>ii가 1,505개로 드러나 비록 비율은 다르지만 전반적으로 글 어절의 이중모음을 단모음으로 실현한 사례가 제법 나타나고 있음을 보여준다. 다시 단모음으로 돌아가서 변동 없이 발화된 모음을 살펴보면, xx와 ii가 각각 75,159개와 74,885개로 나타났고, vv는 66,351개, oo는 40,491개, ee와 EE는 합쳐서 63,246개, uu>uu로 실현된 것이 22,898개이고 “가지고”를 “가지구”로 발음한 예처럼 oo>uu로 모음상승을 보인 것은 12,203개를 보였는데, 좀 더 자세한 음운변동을 일으킨 환경에 대한 관찰이 필요할 것으로 여겨진다. 이러한 변동은 앞서 살펴본 Han & Kang(2013)과 Kang & Han(2013)의 연구에서 /ㄴ/의 F2값이 내려와 /ㄷ/로 합류한다는 보고와도 관련이 있을 수 있다. 다시 말해서, 서울코퍼스의 연구자들이 실제 발화를 듣고 전사하는 과정에서 /ㄴ/를 /ㄷ/로 적었을 수도 있을 것으로 추정되는데, 앞으로 서울코퍼스의 음성자료 가운데 이 부분을 뽑아 지각실험을 통해 살펴볼 필요가 있을 것으로 여겨진다.

양병곤(2016)에서는 이중모음이 글 어절에서 8.8%를 차지했고, 말 어절에서는 15.3%로 거의 2배나 증가했다고 지적했는데, 그 원인으로 젊은이들의 마지막 발화에서 yo를 yv로 바꾸어 말했을 것이라고 추정했다. 이 연구에서는 yv음절의 글이나 말 어절에서의 위치를 성별 연령별로 나누어 자세히 분석하겠지만, 변동이 전혀 없이 말 어절로 실현된 yv>yv가 16,138개를 보였고 이어서 yo>yv가 9,599개를 기록했고, 그 외에도 vv>yv가 153개, wi-vv>yv로 된 경우가 66개, ii-vv>yv는 21개, 나머지는 20개 이하의 다양한 변화를 보였다. 다른 이중모음에서도 이렇게 두 개의 단모음이 합쳐져서 실현된 것이 있는데, ii-aa>ya로 12개, ii-vv>yv가 21개, ii-ee>ye가 3개, uu-aa>wa가 3개, uu-ii>wi가 1개 등으로 생각보다 적은 수가 말 어절에서 이중모음으로 변동을 보였다. 그런 소수의 이중모음 변동에서 국어보다는 “이탈리아”나 “다이어트” 등의 외래어의 두 개의 모음이 이중모음으로 축약되어 발음된 것들이 포함되어 있었다. 이 연구에서는 모음에만 집중하여 살펴보고 있지만, 이러한 음운변동을 좀 더 자세히 관찰하려면 이문규(2015)나 신승용(2009)이 지적했듯이 앞뒤의 음절의 자음과 모음의 구성요소까지도 함께 분석해야만 음운변동규칙을 제대로 설명할 수 있을 것으로 여겨진다.

요약해보면 글 어절의 단모음의 높은 비율이 말 어절에서 그대로 실현되었고, 글 어절의 이중모음은 대체로 말 어절에서는

단모음화하는 경향을 보였는데, 그 반대의 경우는 상대적으로 낮게 나타났다고 할 수 있다.

3.2. 성별로 본 모음의 음운변동 분포

이번에는 서울코퍼스의 모음을 남녀 집단별로 나누어서 살펴보기로 한다. 남성의 말 어절에서의 모음 수는 모두 297,851개이고, 음운변동 유형수는 모두 280가지로 나타났고, 여성의 모음 수는 248,553개로 음운변동 유형수는 229가지다. 여성의 발화에서 모음의 수는 남성에 비해 49,298개가 적고 음운변동 유형도 51개가 적다. 모음의 음운변동 유형이 너무 많기 때문에 이 논문에서는 말 어절에서 음운변동 유형 가운데 <표 2>와 같이 출현 빈도수가 가장 높은 15개를 중심으로 남녀로 대응시켜 살펴보기로 한다.

표 2. 남녀 집단별 음운변동 유형별 빈도수 분포
Table 2. Frequency distribution of phonological process types by male and female groups

Phonological process types	Frequency	
	Males	Females
aa>aa	61331	51381
ii>ii	41875	33010
xx>xx	41201	33958
vv>vv	34619	31742
oo>oo	23480	17011
ee>ee	20910	18578
EE>EE	13332	11426
uu>uu	12812	10086
yv>yv	9197	6941
oo>uu	5717	6486
yo>yv	4248	5351
yo>yo	3229	1954
ya>ya	2372	2231
wv>vv	1748	1413
wa>wa	1736	1177

먼저 집단별 상호 연관성을 검정해보기 위해 카이스퀘어 검정을 실시해본 결과 $\chi^2 = 1377.9$, $df = 14$, $p < .05$ 의 결과가 나와서 남녀 집단의 음운변동이 서로 유의미하게 연관이 되어 있음을 알 수 있다. 실제 두 집단 간에 피어슨의 상관분석을 실시해본 결과 $r = 0.996$ ($p < .05$)로 나타나서 매우 강한 상관을 보이고 있었다. 이러한 결과는 남녀 집단별 음운변동의 절대적인 빈도수에서는 다소 차이가 있지만, 유형별로 비슷한 음운변동의 분포를 보였음을 증명해준다고 할 수 있다. 카이스퀘어 검정 후 피어슨 잔차(Pearson residuals)를 구해서 성별에 따른 연관성 정도를 살펴본 결과 긍정적인 영향은 남성화자들에서는

oo>oo(9.8)와 yo>yo(7.7), ii>ii(5.6)순으로 나타났고, 여성화자들에서는 yo>yv(14.7), oo>uu(12.4), vv>vv(8.6) 순으로 나타났다. 반대로 집단의 특성과 관련이 없다는 부정적인 연관성 값의 분포는 남성화자들에서는 yo>yv(-13.5.), oo>uu(-11.3) vv>vv(-7.8) 순으로 나타났고, 여성화자들에서는 oo>oo(-10.7), yo>yo(-8.4), ii>ii(-6.1)순으로 나타나서 yo>yv가 남녀집단의 발화행위를 구분하는데 중요한 요인임을 알 수 있다. 양병곤(2016)에서 지적한 yo를 yv로 변형하여 발음한 빈도수는 남성의 4,248개에 비해 전체모음 수가 적은 여성이 상대적으로 많은 5,351개를 보였다. 이러한 음운변동은 음절 말을 포함해서 음절의 다른 위치에서도 일어난 현상을 포함하고 있기 때문에, 앞 절의 연구 방법에서 기술한대로 각 음절의 끝자락에서 음운변동을 보이는 것만 따로 추출하여 조사해보았다. 그 결과 남성 집단에서는 4,230개를 보였고, 여성 집단에서는 전체 말 어절의 모음의 수가 남성보다 적은데도 불구하고 5,330개나 보여서, 여성이 음절말의 “오”를 “어”로 발음하는 경향이 더 강하다고 할 수 있다. 한 가지 더 언급할 것은 후설모음을 상승시키는 oo>uu변동인데 어말에서는 남성이 4,833개를 보였고, 여성은 5,683개를 보여서 yo>yv와 비슷한 경향을 보이고 있는데, 비록 말 어절의 전체 모음 수에서 oo>uu로 변동된 비율은 남성 화자에서는 1.6%에 불과하고 여성에서는 2.3%에 불과하지만 사회언어학적으로 남녀 간의 언어쓰임새에 어떤 영향을 미치는지, 또 그런 음운변동을 적용한 음성을 들려줬을 때 집단별로 어떤 반응을 보이는지 등은 추가로 연구할 필요가 있을 것으로 여겨진다.

3.3. 연령별로 본 모음의 음운변동 분포

연령별로 모음의 음운변동을 살펴본 결과 각 집단별 말 어절에 나타난 총 음운변동 유형수(총 모음 개수)는 10대가 197가지(106,712개), 20대가 211가지(133,027개), 30대가 190가지(153,797개), 40대가 203가지(152,868개)이다. 또한 yo>yv의 30대 음운변동이 2,712번으로 전체 모음에서 차지하는 비율은 1.8%에 불과하다. 이러한 분포가 사회언어학적인 의미를 가지는지는 앞의 성별 구분과 함께 추가 연구가 필요할 것으로 여겨진다.

<표 3>은 연령별로 구분한 음운변동 분포의 일부를 보여준다. 먼저 집단별 연관성을 살펴보기 위해 카이스퀘어 검정을 실시해본 결과 $\chi^2 = 1558.8$, $df = 42$, $p < .05$ 의 결과가 나와서 연령별 집단의 음운변동이 서로 유의미하게 연관이 되어 있음을 알 수 있다. <표 4>는 네 연령 집단 간에 피어슨의 상관분석을 실시하여 얻은 상관계수를 보여준다. 모두 0.994이상의 상당히 강한 상관관계를 보여주고 있으며 통계적으로 유의미한 결과를 나타내고 있다. 앞 절에서 살펴보았던 yo>yv의 음운변동은 30대가 2,712개로 가장 많고, 이어서 40대가 2,516개, 20대가 2,379개, 마지막으로 10대가 1,992개의 순서로 이어져 있다. 모음상승현상(oo>uu)은 <표 3>에 따르면 10대에서 40대로 점점 많아지고 있음을 알 수 있다. 따라서 이러한 수치 값은 절대적인 값으로 비교하기 보다는 전체에서 차지하는 비율을 고려해

볼 필요가 있을 것이다.

표 3. 연령별 음운변동 유형별 빈도수 분포

Table 3. Frequency distribution of phonological processes by age groups

Age Groups	10s	20s	30s	40s
aa>aa	22652	27860	30499	31701
ii>ii	14132	18181	21265	21307
xx>xx	13926	17502	22064	21667
vv>vv	13464	16510	19112	17275
ee>ee	8125	9826	11639	9898
oo>oo	7854	10208	11970	10459
EE>EE	5204	6189	6314	7051
uu>uu	4185	5926	6056	6731
yv>yv	3189	3860	4404	4685
oo>uu	2043	2518	3440	4202
yo>yv	1992	2379	2712	2516
yo>yo	1438	1243	1115	1387
ya>ya	1065	967	1346	1225
wE>EE	750	902	794	703
wv>vv	713	736	852	860

표 4. 연령별 집단 간의 음운변동 빈도수의 상관계수

Table 4. Correlation of frequency distribution of phonological processes among age groups

Age Groups	10s	20s	30s
20s	0.999*		
30s	0.995*	0.997*	
40s	0.994*	0.996*	0.995*

*p<.05

마지막으로 연령별로 말 어절 끝에서 yo>yv와 oo>uu로 실현된 분포를 조사하고 <표 5>로 나타내어 보았다.

<표 5>에 따르면 두 가지 음운변동의 총 빈도수는 20,076개로 yo>yv의 실현은 주로 40대 여성에서 가장 많이 나타났고 40대 남성이 가장 적으며 10대 남녀도 상대적으로 적게 사용하고 있음을 알 수 있다. oo>uu로의 음운변동은 40대 여성 집단이 가장 높고 이어서 40대 남성 집단과 30대 남성 집단이 따라오고 10대 남성 집단이 가장 낮은 빈도수를 보이고 있다. 이러한 분포는 절대적인 발화 수에 따라 달라지긴 하지만 대체로 연령별 언어습관을 반영하는 것으로 여겨진다. 앞으로 더 많은 자연발화 자료가 나오면 집단별 발화경향의 사회언어학적 의미를 찾을 수 있을 것으로 기대된다.

표 5. 연령별 음운변동 yo>yv와 oo>uu 유형별 빈도수 분포
 Table 5. Frequency distribution of phonological processes(yo>yv, oo>uu) by age groups

Types Age Groups	yo>yv		oo>uu		Total
	Males	Females	Males	Females	
10s	983	1003	556	1309	3851
20s	1142	1235	984	1148	4509
30s	1342	1358	1639	1268	5607
40s	763	1734	1654	1958	6109
Sum	4230	5330	4833	5683	20076

4. 요약 및 결론

이 논문에서는 자연발화를 수집한 서울코퍼스의 글 어절과 말 어절을 각 음절 요소별로 분리하고, 이를 성별과 연령별로 나누어 단모음과 이중모음의 음운변동을 살펴보았다. 연구방법으로는 Praat의 층위표기자료와 참여자의 정보를 읽어 들어 어절단위로 나눈 다음, R의 table 함수를 이용해서 음운변동 유형별로 빈도수를 분석하고, 이 분포에 대해 카이스퀘어 검정과 상관분석을 통해 통계적으로 분석해 보았다. 연구결과를 요약하면 다음과 같다.

첫째, 서울코퍼스의 음운변동을 분석해본 결과 글 어절과 말 어절의 음절수가 같은 경우는 약 97%로 나타났고 3%만이 다른 수의 음절로 발화되었다. 글 어절과 말 어절이 다른 경우는 65.8%나 되었는데, 이런 음운변동의 주된 원인은 자음의 변화이고, 모음은 약 90.5%가 변동 없이 말 어절에서 실현되었고, 이 가운데 단모음은 87.5%가 그대로 실현되었다. 이러한 경향은 국어에서 모음이 음운변화에 매우 강인함을 증명한다. 모음 가운데서도 aa가 가장 많이 쓰이고 있으며 글 어절에 나타난 aa의 98.7%가 말 어절에서 그대로 실현되었다. 이중모음은 말 어절에서 단모음으로 변한 경우가 그 반대쪽의 변동보다는 많았다.

둘째, 성별로 나누어 살펴본 결과 남성의 모음 발화수가 여성에 비해 많았고 음운변동 유형수도 많았다. 음운변동 유형의 출현 빈도수가 가장 높은 15가지를 남녀집단별로 구해 카이스퀘어 검정을 실시해본 결과 유의하게 연관성을 보였고, 피어슨의 상관분석에서도 0.996의 매우 높은 상관계수를 보여서 절대적인 빈도수의 차이는 있지만, 유형별로 비슷한 통계적인 분포를 보였다고 할 수 있다. 음절말의 이중모음 “요”를 “여”로 발음하는 경향은 여성이 남성보다 더 강하였다.

셋째, 연령별로 음운변동 유형 분포를 살펴본 결과 유의하게 연관을 보였으며, 피어슨의 상관분석에서도 0.99의 매우 강한 상관관계를 보였다. 말 어절의 이중모음 “요”를 “여”로 발음한 빈도수는 30대가 가장 많고 이어서 40대, 20대, 10대의 순서로 나타났다. 이런 음운변동이 말 어절 끝에서 실현된 수는 40대 여성에서 가장 많고 이에 반해 40대 남성이 가장 적었다. 이

러한 경향은 40대 집단의 어말에서 모음 “오”를 “우”로 바꾸어 발음한 음운변동에서도 같은 경향을 보였다.

이러한 결과를 보면 자연발화 코퍼스를 이용해서 국어 음운변동의 분포를 음절구조로 분리하여 성별·연령별로 분류하여 보면 다양한 언어의 쓰임새를 파악할 수 있다고 결론지을 수 있다.

앞으로 영어의 벽아이코퍼스나 타 언어로 된 자연발화코퍼스의 모음의 음운변동을 분석한 결과가 나온다면 매우 흥미로운 언어 간의 비교를 해 볼 수 있을 것으로 기대된다.

참고문헌

- Boersma, P., & Weenink, D. (2017). Praat: Doing phonetics by computer. Retrieved from <http://www.fon.hum.uva.nl/praat/> on April 21, 2017.
- Dalby, J. M. (1986). *Phonetic structure of fast speech in American English*. Ph.D. Dissertation, Indiana University.
- Han, J., & Kang, H. (2013). Cross-generational change of /o/ and /u/ in Seoul Korean II: Spectral interactions in normalized vowel space. *Phonetics and Speech Sciences*, 5(2), 33-41.
- 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.
- Jurafsky, D., Bell, A., Gregory, M. L., & Raymond, W. D. (2001). Probabilistic relations between words: evidence from reduction in lexical production. In J. Bybee & H. Paul (Eds.), *Frequency and the Emergence of Linguistic Structure* (pp. 229-254). John Benjamins; Amsterdam.
- Kang, H., & Han, J. (2013). Cross-generational Change of /o/ and /u/ in Seoul Korean I: Proximity in Vowel Space. *Phonetics and Speech Sciences*, 5(2), 25-31.
- Kang, J., & Kong, E. (2016). Static and dynamic spectral properties of the monophthong vowels in Seoul Korean: Implication on sound change. *Phonetics and Speech Sciences*, 8(4), 39-47. (강지은·공은정 (2016). 서울 방언 단모음의 소리 변화와 음향 단서 연구: 단일지점 포먼트와 궤적 양상. *말소리와 음성과학*, 8(4), 39-47.)
- Kiesling, S., Dille, L., & Raymond, W. D. (2006). The Variation in Conversation (ViC) project: Creation of the Buckeye Corpus of Conversational Speech. Retrieved from <http://buckeyecorpus.osu.edu/BuckeyeCorpusmanual.pdf> on April 27, 2017.
- Kim, J. (2015). The effect of word frequency on the reduction of English CVCC syllables in spontaneous speech. *Phonetics and Speech Sciences*, 7(3), 45-53.
- Labov, W. (1994). *Principles of linguistic change*. Cambridge, MA: Blackwell.

- Lee, M. (2015). *Modern Korean phonology for Korean education*. Seoul: Hankookmunhwasa. (이문규 (2015). *국어교육을 위한 현대 국어 음운론*. 서울: 한국문화사.)
- Pitta, M. A., Johnson, K., Hume, E., Kiesling, S., & Raymond, W. (2005). The Buckeye corpus of conversational speech: Labeling conventions and a test of transcriber reliability. *Speech Communication*, 45(1), 89-95.
- R Core Team (2017). R: A language and environment for statistical computing. Retrieved from <https://www.r-project.org/> [R Foundation for Statistical Computing, Vienna, Austria] on April 27, 2017.
- Shin, S. (2009). *Korean syllable phonology*. Seoul: Bakijung. (신승용 (2009). *국어 음절 음운론*. 서울: 박이정.)
- Strik, H., & Cucchiari, C. (1999). Modeling pronunciation variation for ASR: a survey of the literature. *Speech Communication*, 29, 225-246.
- Yang, B. (1996). A comparative study of English and Korean monophthongs produced by male and female speakers. *Journal of Phonetics*, 24, 245-261.
- Yang, B. (2012). Reduction and frequency analyses of vowels and consonants in the Buckeye Speech Corpus. *Phonetics and Speech Sciences*, 4(3), 75-83.
- Yang, B. (2016). 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(3), 1-9. (양병곤 (2016). 음절구조로 본 서울코퍼스의 글어절과 말어절의 음소분포와 음운변동. *말소리와 음성과학*, 8(3), 1-9.)
- 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.

• **양병곤 (Yang, Byunggon)**

부산대학교 영어교육과

부산시 금정구 장전동 30

Tel: 051-510-2619

Email: bgyang@pusan.ac.kr

Homepage: <http://fonetiks.info/bgyang>