



A quantitative study on the minimal pair of Korean phonemes: Focused on syllable-initial consonants

Jieun Jung*

Department of Korean Language and Literature, Hankuk University of Foreign Studies, Seoul, Korea

Abstract

The paper investigates the minimal pair of Korean phonemes quantitatively. To achieve this goal, I calculated the number of consonant minimal pairs in the syllable-initial position as both raw counts and relative counts, and analyzed the part of speech relations of the two words in the minimal pair. 『Urimalsaem』 was chosen as the object of this study because it was judged that the minimal pair analysis should be done through a dictionary and it is the largest among Korean dictionaries. The results of the study are summarized as follows. First, there were 153 types of minimal pairs out of 337,135 examples. The ranking of phoneme pairs from highest to lowest was ‘ㅅ-ㅈ, ㄱ-ㅅ, ㄱ-ㅈ, ㄱ-ㅊ, ㄱ-ㅎ, ..., ㅍ-ㅋ, ㅌ-ㅋ, ㅍ-ㅋ, ㄹ-ㅍ, ㅍ-ㅋ’. The phonemes that played a major role in the formation of the minimal pair were /ㄱ, ㅅ, ㅈ, ㅊ, ㅌ/, in that order, which showed a high proportion of palatals. The correlation between the raw count of minimal pairs and the relative count of minimal pairs was found to be quite high $r=0.937$. Second, 87.91% of the minimal pairs shared the part of speech (same syntactic category). The most frequently observed type has been ‘noun-noun’ pair (70.25%), and ‘vowel-vowel’ pair (14.77%) was the next ranking. It can be indicated that the minimal pair could be grouped into similar categories in terms of semantics. The results of this study can be useful for various research in Korean linguistics, speech-language pathology, language education, language acquisition, speech synthesis, and artificial intelligence-machine learning as basic data related to Korean phonemes.

Keywords: minimal pair, dictionary, syllable-initial, consonant, part of speech, functional load

1. 서론

이 연구의 목적은 한국어 음소의 최소대립쌍 출현 양상에 대해 계량언어학적으로 알아보는 것이다. 최소대립쌍(minimal pair)¹은 현대 음운론 연구에서 가장 기본적인 개념 중 하나로, 동일한 위치에 놓인 하나의 소리만 다르고 나머지 소리는 모두

같은데도 불구하고 서로 구별되는 두 개의 단어를 묶어 최소대립쌍이라고 부른다(Lee, 2010:119). 최소대립쌍은 한 언어에서 음소의 체계를 세우는 데 중요한 역할을 하고, 기능부담량(functional load)²의 측정에도 중요한 척도가 됨에도 불구하고 아직까지 한국어 음소의 최소대립쌍에 대해 전면적이고 다각적으로 살핀 연구는 찾아보기 어렵다. 최소대립쌍의 분석은 기본

* jieunkjung@gmail.com, Corresponding author

Received 2 February 2019; Revised 16 March 2019; Accepted 18 March 2019

© Copyright 2019 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.

적으로 사전을 통해서 이루어져야 한다(Jin, 1996:169; Kim & Kang, 1997:6; Kim et al., 2014:118; Lee et al., 1987:120). 하지만 사전의 분석은 방대한 작업이기 때문에 현실적인 어려움으로 인해 연구가 이루어지기 쉽지 않았다.

한국어 음소의 최소대립쌍을 분석한 종합적인 연구는 없지만 부분적으로 살펴본 논의로는 Jin(1996), Kim et al.(2014), Lee et al.(1987) 이 있다³. Lee et al.(1987)은 한국정신문화연구원에서 간행한 『한국어표준발음사전』에 등재된 92,691개 표제어의 발음을 대상으로 음소 및 음소열의 출현 빈도, 최소대립쌍을 분석했다. 이 연구는 최초로 한국어 음소의 최소대립쌍에 대한 분석을 시도한 점에서는 의의가 있으나 ‘아람’과 ‘사람’의 관계도 최소대립쌍으로 인정한 것에서 볼 수 있듯이 공백소와 음소도 하나의 쌍으로 보고 분석한 것, 최소대립쌍의 개수 결과에 대한 해석 없이 결과만 평면적으로 나열한 것이 한계라고 할 수 있다. Jin(1996)은 한국방송공사가 펴낸 『한국어발음대사전』의 65,291개 표제어의 발음을 대상으로 모음 최소대립쌍의 수를 세어 모음과 모음 간에 나타나는 대립의 정도를 통해 기능부담량을 밝히고, /ㄱ/와 /ㅋ/ 모음이 합류되는 것을 기능부담량의 관점에서 살폈다. 이 연구는 음운 변화의 양상을 최소대립쌍과 기능부담량의 관점에서 해석하고자 한 점에서는 의의가 있으나 이러한 연구 목적과 관련된 특정 음소의 대립쌍만을 분석하여 전반적인 모음 대립의 양상을 확인할 수 없다는 것이 한계라고 할 수 있다. Kim et al.(2014)은 음성학 관련 연구자를 위하여 최소대립쌍 검색 도구를 개발하여 제안한 연구이다. 이 연구에서 제안한 도구를 사용하면 Shin(2010)의 발음사전(『연세한국어사전』)에 수록된 표제어 47,401개에 대한 최소대립쌍 검색이 가능하다고 했다. 이 검색 도구는 단어와 음소 검색뿐만 아니라 어종, 품사에 대한 세부적인 검색도 가능하다고 했지만 실제로 찾아서 사용할 수가 없어서 그 효율성은 확인이 불가능했다.

이상의 선행 연구에서 확인할 수 있듯이 방대한 사전 자료를 중심으로 한국어 음소의 최소대립쌍을 계량적으로 분석한 사례가 거의 없다고 볼 수 있다. 따라서 이 연구에서는 『우리말샘』의 표제어를 대상으로 초성 위치 자음의 최소대립쌍과 관련된

정보를 계량적으로 추출해 보고자 한다.

최소대립쌍의 개수는 절대수치와 상대수치로 측정할 수 있고, 상대수치는 두 가지 방법으로 계산이 가능하다. 첫 번째 방법은 최소대립쌍의 절대수치를 전체 단어의 수로 나누는 것이고, 두 번째 방법은 최소대립쌍의 절대수치를 최소대립쌍을 이루는 두 음소를 포함하는 단어의 개수로 나누는 것이다. 코퍼스 내에서 빈도가 낮은 음소는 최소대립쌍의 개수도 적을 가능성이 높다. 그러나 두 번째 방법으로 산출한 상대수치는 음소의 빈도까지 고려하여 상대적인 비교가 가능하다는 이점이 있다⁴. 따라서 본 연구에서는 최소대립쌍의 개수를 절대수치와 두 번째 방법의 상대수치로 산출하고 그 두 수치 간의 상관관계도 살펴 보겠다.

Kim et al.(2014:120)은 최소대립쌍은 동일 위치에 놓인 하나의 음소만 다르고 나머지 음소는 모두 같기 때문에 두 단어는 인지적으로 비슷한 범주로 묶일 가능성이 높고, 따라서 동일한 품사 내에 최소대립쌍이 존재할 가능성이 품사 범주를 넘어서 존재할 가능성보다 더 높다고 했다. Brown(1988, 1991), Catford(1987), Derwing & Munro(2015), Levies & Cortes(2008) Munro & Derwing(2006) 등은 영어 발음 교육의 우선순위를 매길 때 최소대립쌍의 개수가 중요한 기준이 된다고 했다. 예를 들어 /e/-/æ/나 /p/-/b/의 경우 만들 수 있는 최소대립쌍의 개수가 많고 사용 빈도가 높기 때문에 우선적으로 교육해야 하고, 반대로 /u/-/ʊ/나 /ʃ/-/ʒ/와 같은 음소 쌍은 사용 빈도도 낮을 뿐만 아니라 원어 빈도 혼용해서 쓰는 경우가 있기 때문에 발음 교육에서 크게 중요하지 않다고 주장했다. 그리고 최소대립쌍을 이루는 두 단어의 품사가 같으면 기능부담량이 높고 다르면 기능부담량이 낮기 때문에 최소대립쌍을 기준으로 기능부담량을 설정하고 발음 교육에 적용할 때는 품사와의 관련성도 고려해야 한다고 했다⁵. 따라서 본 연구에서는 최소대립쌍과 품사와의 관계도 살펴 보겠다.

최소대립쌍 관련 정보를 다양한 측면에서 분석하는 것은 한국어 음소 체계의 특성을 밝히는 데 중요한 단서가 될 수 있을 것이고, 한국어학뿐만 아니라 한국어 교육학, 언어 병리학, 음성

1 이 연구에서는 두 단어의 묶음으로 이루어진 쌍만을 최소대립쌍으로 인정한다. 논의에 따라서는 가령 국어 장애음의 삼지적 상관속을 언급하면서 ‘달, 탈, 딸’과 같은 세 단어의 묶음을 최소대립쌍으로 설정하기도 하지만 엄밀하게 말하면 최소대립쌍은 두 개로 이루어진 단어 묶음만을 가리킬 뿐이다. 최소대립쌍과 비슷한 개념으로 최소대립군(minimal set)도 있다. 이것은 동일한 위치에 놓이는 소리의 차이에 따라 구분되는 단어들을 모두 묶는 개념으로 사용된다. 가령 ‘둘(二), 물(水), 불(火), 술(酒), ...’과 같이 ‘ㄹ’ 앞에 오는 초성의 차이에 따라 구별되는 단어들을 총괄하여 최소대립군이라고 부를 수 있다(Lee, 2010:19).

2 음운의 대립이 단어의 의미 구별에 관여하는 정도를 가리킨다. 어떤 음운의 기능부담량이 많다는 것은 그만큼 그 음운이 많은 단어의 구별에 쓰인다는 것이다. 기능부담량이 많은 음운은 음운 체계 내에서의 지위가 확고하다. 반면 기능부담량이 적은 음운은 그렇지 못하므로 다른 음운으로 바뀌거나 음운으로서의 지위를 잃어버릴 가능성이 있다. 가령 현대 국어의 장단은 기능부담량이 많지 않은 것으로 알려져 있는데 이것이 음운으로서의 지위가 흔들리는 데 중요하게 작용했다고 보기도 한다. 기능부담량은 대체로 최소대립쌍의 수를 세는 방식으로 이루어지는 것이 일반적이다. 물론 현실적으로 이런 과정을 거쳐 국어의 음운들이 각각 어느 정도의 기능부담량을 가지는지 제어 본 논의는 별로 없다. 그러나 기능부담량의 개념을 볼 때 최소대립쌍의 수가 기능부담량 측정에 중요한 척도가 된다는 것은 부인하기 어렵다(Lee, 2017:87).

3 Lee et al.(1987)은 ‘최소대립쌍’ 대신 ‘최소음소쌍’이라는 용어를 사용하고, Jin(1996)은 ‘최소대립쌍’ 대신 ‘최소대립어’라는 용어를 사용했다.

4 최소대립쌍의 상대수치 측정 방법은 Hall et al.(2017)을 참고하였다.

5 Derwing & Munro(2015:75-76)은 청자는 단어를 들을 때 명사를 다른 명사와 혼동할 확률이 높지 명사를 다른 품사인 형용사나 동사와 혼동할 확률이 높은 것이 아니기에 두 단어의 품사가 같을 때 기능부담량이 높다고 할 수 있다고 했다.

공학 분야에서 활용할 수 있는 기초 자료로 제공될 수 있을 것이다. 이에 이 연구에서는 다음과 같은 연구 문제를 중심으로 살펴보고자 한다.

- ① 음소 쌍별로 최소대립쌍의 절대/상대수치는 어떠한가?
- ② 최소대립쌍을 이루는 두 단어의 품사 정보는 어떠한가?

2. 연구 방법

이 연구의 분석 대상이 된 사전은 『우리말샘』이었다. 이 사전을 분석의 바탕 자료로 삼은 이유는 이 사전이 첫째, 어휘의 규모가 역대 최대인 사전이고, 둘째, 표제어의 표준 발음의 표기가 확대된 사전이며, 셋째, 저작권이 개방된 사전이기 때문이다⁶.

『우리말샘』의 표제어 수는 약 1,000,000여 개였는데 표제어 중 잘못된 형태, 지역어(방언), 북한어, 옛말, 자모형을 걸러낸 후, 품사 정보와 원어 정보가 없는 표제어도 제외했다. 품사가 두 개 이상인 표제어는 두 개의 표제어로 분리하고, 발음이 두 개 이상인 표제어는 표제어별로 한 개씩의 대표 발음을 정했다. 이러한 자료 구성 원칙에 따라 선정된 표제어의 수는 총 458,365 개였다⁷. 각 표제어별로 발음, 품사, 원어 정보를 부여한 후 이 네 정보가 모두 같은 동음이의어는 하나만 남겼다⁸. 이런 작업을 거쳐 발음사전을 완성했고 총 325,717개였다. 완성된 자료는 Python(파이썬)으로 분석 프로그램을 설계하여 최소대립쌍의 개수를 분석했다.

표 1은 최소대립쌍 분석을 위해 구축한 발음 사전의 예이다.

표 1. 발음 사전의 예시
Table 1. Example of pronunciation dictionary

ID	표제어	발음형	음절형	음소형	품사	원어
000001	가방	가방	가부방말	기초 ㅏ ㅈ ㅍ ㅁ 비초 ㅏ ㅈ ㅍ ㅁ	명사	고유어
000002	가방	가방	가부방말	기초 ㅏ ㅈ ㅍ ㅁ 비초 ㅏ ㅈ ㅍ ㅁ	명사	한자어
000003	가오리	가오리	가두오종리말	기초 ㅏ ㅈ ㅍ ㅁ C초 ㅏ ㅈ ㅍ ㅁ 리초 ㅏ ㅈ ㅍ ㅁ	명사	고유어
000004	가장	가장	가부장말	기초 ㅏ ㅈ ㅍ ㅁ 지초 ㅏ ㅈ ㅍ ㅁ	명사	한자어
000005	가장	가장	가부장말	기초 ㅏ ㅈ ㅍ ㅁ 지초 ㅏ ㅈ ㅍ ㅁ	부사	고유어
000006	갈락하다	갈라카다	갈두라중카중다말	기초 ㅏ ㅈ ㅍ ㅁ 리초 ㅏ ㅈ ㅍ ㅁ 카초 ㅏ ㅈ ㅍ ㅁ 다초 ㅏ ㅈ ㅍ ㅁ	형용사	혼종어

ID	표제어	발음형	음절형	음소형	품사	원어
000007	강가	강까	강두까말	기초 ㅏ ㅈ ㅍ ㅁ 기초 ㅏ ㅈ ㅍ ㅁ	명사	한자어
000008	꽁지	꽁지	꽁두지말	기초 ㅏ ㅈ ㅍ ㅁ 지초 ㅏ ㅈ ㅍ ㅁ	명사	고유어
...
325717	횡하다	헝하다	헝두하중다말	기초 ㅏ ㅈ ㅍ ㅁ 헝초 ㅏ ㅈ ㅍ ㅁ 다초 ㅏ ㅈ ㅍ ㅁ	형용사	고유어

먼저 표제어마다 아이디, 품사 정보, 원어 정보를 부여했다. 발음형은 음절 단위로 분리하여 개별 음절에서 어두, 어중, 어말 정보를 확인했고, 음절형을 음소 단위로 분리하여 초성, 중성, 종성 정보를 확인했다.

그림 1은 최소대립쌍을 분석하는 방법에 대한 절차를 단계별로 정리하여 보인 것이다.

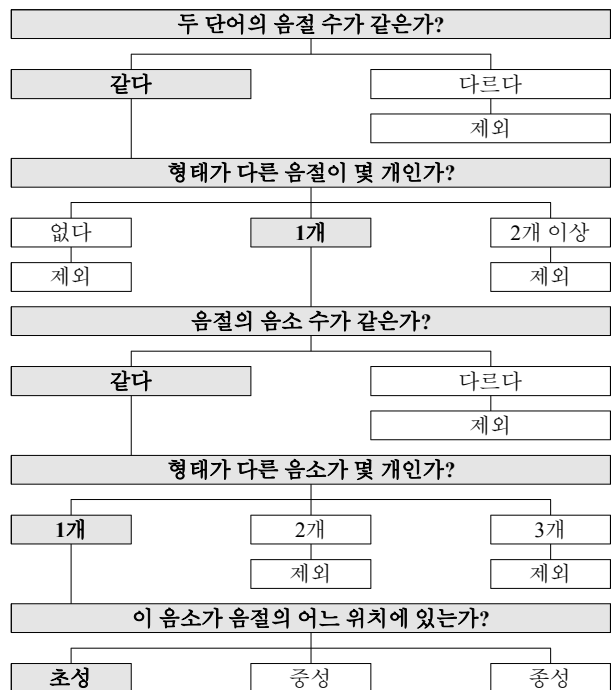


그림 1. 최소대립쌍 분석 절차
Figure 1. Minimal pair analysis procedure

하나의 표제어에 대해서 위와 같은 방법으로 최소대립쌍을 찾았다. 예를 들어 기준 단어가 아이디 000001 ‘가방’인 경우이다. ‘가방’과 ‘가방’ 이외의 325,716개의 단어를 각각 비교했다. 음절 수가 같은지, 음절 수가 같다면 형태가 다른 음절이 몇 개

⁶ opendict.korean.go.kr

⁷ 구체적인 자료 구성 원칙과 자료 처리 방법, 발음형 설정 기준은 Jung(2019)에 상세히 기술했다. 발음형 설정 기준 중 자음의 발음과 관련 있는 것은 아래와 같다.

- /L/을 첨가하여 발음하기도 하지만, 표기대로 /L/ 첨가 없이 발음하기도 하는 단어는 /L/을 첨가한 발음을 발음형으로 한다. 예) 검열[검녕], 야금야금[야금냐금]
- 표준국어대사전에서 정보가 수정된 단어는 수정된 내용을 발음형으로 한다. 예) 김밥[김:뽕]

⁸ 본 연구에서는 다루지 않았지만 추후 최소대립쌍과 원어 정보의 관계에 대한 분석을 하기 위하여 각 단어에 고유어, 한자어, 혼종어와 같은 원어 정보를 부여했다.

인지, 형태가 다른 음절이 1개라면 음절을 구성하는 음소의 수가 같은지, 음소의 수가 같다면 형태가 다른 음소가 몇 개인지, 형태가 다른 음소가 1개라면 이 단어가 ‘가방’의 최소대립쌍이 되는 것이다. 이에 해당하는 단어의 예로는 ‘사방, 고방, 가상, 가봉, 가반’ 등이 있을 수 있는데 ‘가방’과 비교하여 형태가 다른 음소가 음절 내 초성, 중성, 종성 중 어느 위치에 있는지를 확인했고, 음소 쌍별 최소대립쌍의 개수(절대수치)를 산출했다.

표 2는 분석 대상 자료를 품사별로 나누어 빈도와 비율을 보인 것이다.

표 2. 품사별 구성
Table 2. Distribution in the classified the part of speech

	빈도	비율(%)
명사	249,109	76.48
대명사	334	0.10
수사	169	0.05
동사	49,210	15.11
형용사	12,084	3.71
관형사	3,237	0.99
부사	10,801	3.32
조사	151	0.05
감탄사	622	0.19
계	325,717	100.0

가장 높은 비율을 차지한 품사는 76.48%를 보인 명사였고, 가장 낮은 비율을 차지한 품사는 0.05%를 보인 조사였다. 동사가 형용사에 비해 4.07배 높은 빈도를 보인다는 것과 부사가 관형사에 비하여 3.34배 높은 빈도를 보인다는 것이 특징적이다. 9개의 품사 중 5개인 대명사, 수사, 관형사, 조사, 감탄사의 비율은 1% 미만이고, 명사와 동사 두 품사의 비율의 합은 91.59%를 차지하고 있었다.

3. 연구 결과

초성의 위치에서 대립 가능한 음소의 쌍은 산술적으로 $153(=18 \times 17 \times \frac{1}{2})$ 개이다. 분석 결과 모든 음소 쌍에서 최소대립쌍이 존재했고 그 개수는 총 337,135개였다.

3.1. 최소대립쌍의 개수 분석

표 3은 초성 위치의 최소대립쌍의 절대수치를 나타낸 분포 표이고, 표 4는 상대수치를 나타낸 분포표이다. 본 절에서의 상대수치는 1장에서 살펴본 상대수치 계산의 두 방법 중 두 번째 방법으로 산출한 수치로, 최소대립쌍의 절대수치를 최소대립쌍을 구성하는 개별 음소가 하나라도 포함된 단어들의 수로 나누어 산출한 값이다.

표 3. 최소대립쌍의 절대수치
Table 3. Raw count of minimal pairs

	ㄱ	ㄲ	ㄴ	ㄷ	ㄸ	ㄹ	ㅁ	ㅂ	ㅃ	ㅅ	ㅆ	ㅈ	ㅉ	ㅊ	ㅋ	ㆁ	ㅇ	
ㄱ																		
ㄲ	1,524																	
ㄴ	6,052	364																
ㄷ	5,777	403	3,552															
ㄸ	673	576	305	919														
ㄹ	3,082	318	1,652	930	411													
ㅁ	8,140	499	4,401	3,804	498	1,861												
ㅂ	9,064	547	3,961	4,518	583	1,851	6,917											
ㅃ	383	626	218	250	422	114	350	1,381										
ㅅ	10,023	671	4,588	5,415	561	1,504	6,685	7,466	434									
ㅆ	928	871	253	360	849	525	595	796	736	993								
ㅈ	9,978	659	4,209	5,328	724	1,557	5,792	7,068	384	12,931	613							
ㅉ	760	889	243	505	951	508	589	649	654	611	1,714	1,171						
ㅊ	7,084	744	3,054	3,581	935	1,291	4,283	5,289	600	8,294	1,281	8,632	1,476					
ㅋ	1,225	343	475	362	121	586	539	603	89	505	176	497	121	423				
ㆁ	3,558	367	2,173	2,863	747	640	2,149	2,563	277	2,987	450	3,025	527	2,641	311			
ㅇ	4,228	558	2,027	1,990	482	1,036	3,175	2,934	459	2,911	524	2,820	528	2,275	640	1,307		
합	8,826	562	4,034	3,760	419	1,543	4,632	4,810	290	5,530	403	5,374	397	3,774	2,043	2,471	2,820	
계	81,305	10,521	41,561	44,317	10,176	19,409	54,909	61,000	7,667	72,109	12,067	70,762	12,293	55,657	9,059	29,056	30,714	51,688

표 4. 최소대립쌍의 상대수치
Table 4. Relative count of minimal pairs

	ㄱ	ㄲ	ㄴ	ㄷ	ㄸ	ㄹ	ㅁ	ㅂ	ㅃ	ㅅ	ㅆ	ㅈ	ㅊ	ㅌ	ㅍ	ㅑ	ㅓ	ㅕ
ㄱ																		
ㄲ	0.012																	
ㄴ	0.042	0.005																
ㄷ	0.032	0.004	0.027															
ㄸ	0.006	0.014	0.005	0.009														
ㄹ	0.022	0.005	0.019	0.007	0.007													
ㅁ	0.054	0.006	0.046	0.027	0.007	0.020												
ㅂ	0.058	0.007	0.039	0.032	0.008	0.018	0.062											
ㅃ	0.003	0.016	0.004	0.002	0.014	0.002	0.005	0.019										
ㅅ	0.058	0.006	0.036	0.033	0.005	0.012	0.050	0.054	0.004									
ㅆ	0.007	0.018	0.004	0.003	0.021	0.008	0.008	0.010	0.020	0.009								
ㅈ	0.057	0.006	0.033	0.032	0.007	0.012	0.042	0.050	0.004	0.080	0.005							
ㅊ	0.006	0.018	0.004	0.004	0.024	0.008	0.008	0.008	0.018	0.006	0.036	0.010						
ㅌ	0.049	0.011	0.034	0.027	0.015	0.014	0.044	0.051	0.010	0.065	0.019	0.066	0.021					
ㅍ	0.010	0.009	0.008	0.004	0.004	0.010	0.008	0.008	0.003	0.005	0.005	0.005	0.003	0.007				
ㅑ	0.029	0.008	0.034	0.026	0.021	0.010	0.029	0.033	0.009	0.028	0.011	0.028	0.012	0.040	0.010			
ㅓ	0.033	0.011	0.029	0.017	0.011	0.015	0.040	0.034	0.012	0.026	0.011	0.025	0.011	0.032	0.016	0.029		
ㅕ	0.052	0.005	0.034	0.027	0.004	0.013	0.036	0.037	0.003	0.036	0.004	0.034	0.004	0.031	0.022	0.026	0.027	

그림 2는 초성 최소대립쌍의 절대수치와 상대수치의 상관관계를 나타낸 것이다.

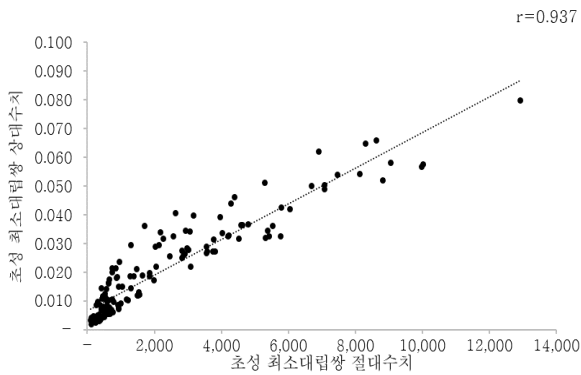


그림 2. 초성 최소대립쌍의 절대수치와 상대수치의 상관관계
Figure 2. Correlation between raw counts and relative counts

초성 최소대립쌍의 절대수치와 상대수치의 상관관계는 0.937로 높은 편이었다. 절대수치와 상대수치에 대한 잔차를 구한 결과 잔차의 값이 큰 음소 쌍은 /ㅃ-ㅆ/, /ㅌ-ㅑ/, /ㅑ-ㅓ/, /ㄴ-ㅑ/, /ㄷ-ㅓ/, /ㄷ-ㅕ/, /ㄱ-ㅓ/, /ㄱ-ㅕ/, /ㄱ-ㅌ/, /ㄱ-ㅍ/, /ㄱ-ㅑ/, /ㄱ-ㅓ/였다.

표 5는 절대수치가 높은 쌍과 낮은 쌍을 15개씩(상하위 10%) 정리한 것이다. 전체 음소 쌍별 초성 최소대립쌍의 개수는 부록 1에 제시하였다.

표 5. 고빈도/저빈도 최소대립쌍(상·하위 10%)
Table 5. Most and least frequently minimal pairs(top 10% and bottom 10%)

순위	초성 쌍	개수	비율(%)	누적비율(%)
1	ㅅ-ㅈ	12,930	3.84	3.84
2	ㄱ-ㅅ	10,023	2.97	6.81
3	ㄱ-ㅈ	9,978	2.96	9.77
4	ㄱ-ㅂ	9,063	2.69	12.46
5	ㄱ-ㅕ	8,825	2.62	15.08
6	ㅈ-ㅌ	8,631	2.56	17.64
7	ㅅ-ㅌ	8,293	2.46	20.10
8	ㄱ-ㅁ	8,138	2.41	22.51
9	ㅂ-ㅅ	7,464	2.21	24.72
10	ㄱ-ㅌ	7,084	2.10	26.83
11	ㅂ-ㅈ	7,066	2.10	28.92
12	ㅁ-ㅂ	6,915	2.05	30.97
13	ㅁ-ㅅ	6,683	1.98	32.95
14	ㄱ-ㄴ	6,053	1.80	34.75
15	ㅁ-ㅈ	5,791	1.72	36.47
:	:	:	:	:
139	ㄲ-ㅋ	343	0.10	99.08
140	ㄲ-ㄹ	318	0.09	99.18
141	ㅋ-ㅌ	311	0.09	99.27
142	ㄴ-ㄸ	305	0.09	99.36
143	ㅃ-ㅕ	290	0.09	99.45
144	ㅃ-ㅑ	277	0.08	99.53
145	ㄴ-ㅃ	254	0.08	99.60
146	ㄷ-ㅃ	250	0.07	99.68
147	ㄴ-ㅆ	243	0.07	99.75
148	ㄴ-ㅃ	218	0.06	99.82
149	ㅃ-ㅋ	176	0.05	99.87
150	ㄸ-ㅋ	121	0.04	99.90
151	ㅆ-ㅋ	121	0.04	99.94
152	ㄹ-ㅃ	114	0.03	99.97
153	ㅃ-ㅋ	89	0.03	100.0
계		337,135	100.0	100.0

초성에서 최소대립쌍을 가장 많이 이룬 음소 쌍은 /ㅅ-ㅈ/로 그 개수와 비율은 12,930개와 3.84%였고, 이어서 /ㄱ-ㅅ/ 쌍이

10,023개(2.97%)로 그 뒤를 따라 많았다. 반면에 가장 적은 음소 쌍은 /ㅁ-ㄷ/였는데, 그 개수와 비율은 89개와 0.03%로 /ㅁ/와 /ㄷ/가 초성 위치에서 대립하여 최소대립쌍을 이루는 경우는 매우 적었다.

상위 15개의 쌍은 전체 출현 최소대립쌍의 36.47%를 차지한 반면 하위 15개의 쌍은 1%도 되지 않아 출현 비율의 차이가 매우 컸다. 상위 15개 최소대립쌍은 ‘장애음과 장애음’의 쌍이 10개, ‘비음과 장애음’의 쌍이 5개였다. /ㄱ/는 15개 중 7개, /ㅅ/와 /ㅈ/는 15개 중 5개의 최소대립쌍에 나타났는데 상위 음소 쌍은 평음의 비중이 높았다. /ㅈ/도 3개의 최소대립쌍에 나타났는데 상위 최소대립쌍에 나타난 격음은 /ㅈ/가 유일했다. 한편 하위 최소대립쌍은 ‘장애음과 장애음’의 쌍이 9개, ‘비음/유음과 경음’의 쌍이 6개였다. /ㅁ/와 /ㄷ/는 15개 중 6개의 최소대립쌍에 나타났는데 하위 음소 쌍은 경음의 비중이 높았다. 그 뒤를 이어 /ㄴ/도 4개의 최소대립쌍에 나타났다.

상하위 음소 쌍을 통해 최소대립쌍을 이루는 데 가장 많이 사용되는 음소는 /ㄱ/고, 반대로 가장 덜 사용되는 음소는 /ㅁ/와 /ㄷ/일 것이라고 예상할 수 있다. 정확한 확인을 위하여 음소별 최소대립쌍 빈도를 분석한 결과는 표 6과 같다.

표 6. 음소별 최소대립쌍 빈도 분포
Table 6. Frequency distribution of minimal pairs by phoneme

순위	음소	빈도	비율(%)	누적비율(%)
1	ㄱ	81,301	12.06	12.06
2	ㅅ	72,103	10.69	22.75
3	ㅈ	70,755	10.49	33.25
4	ㅂ	60,988	9.05	42.29
5	ㅊ	55,650	8.25	50.55
6	ㅇ	54,906	8.14	58.69
7	ㅎ	51,683	7.67	66.36
8	ㄷ	44,315	6.57	72.93
9	ㄴ	41,561	6.16	79.09
10	ㅍ	30,713	4.56	83.65
11	ㅌ	29,046	4.31	87.96
12	ㄹ	19,411	2.88	90.84
13	ㅍ	12,294	1.82	92.66
14	ㅍ	12,066	1.79	94.45
15	ㄱ	10,524	1.56	96.01
16	ㄷ	10,171	1.51	97.52
17	ㅋ	9,059	1.34	98.86
18	ㅁ	7,668	1.14	100.0
	계	674,270	100.0	100.0

가장 높은 빈도를 보인 음소는 /ㄱ/로 12.06%를 나타냈고, 이어서 /ㅅ, ㅈ, ㅂ, ㅊ/의 순서로 높게 나타났다. 반면에 가장 낮은 빈도를 보인 음소는 /ㅁ/로 1.14%에 불과했다. 표 5의 상위 15개에 포함된 음소 쌍의 음소가 표 6에서도 높은 빈도를 보였고, 표 5의 하위 15개에 포함된 음소 쌍의 음소는 표 6에서도 낮은

빈도를 보였다. 1위부터 9위까지는 /ㅈ/와 /ㅎ/를 제외하면 모두 평음과 비음이며, 10위부터 18위까지는 /ㄷ/를 제외하면 모두 경음과 격음이었다. 장애음 중에는 경음과 격음보다 평음이, 공명음 중에는 유음보다 비음이 최소대립쌍을 많이 만든다는 것을 알 수 있는데, /ㅈ/는 격음임에도 불구하고 8.25%로 높은 것이 특징적이다. /ㅁ/도 경음 중에서는 가장 높은 빈도를 나타냈다.

초성마다 최소대립쌍을 가장 많이 만드는 음소도 다르게 나타났다는데 결과는 표 7에 제시하였다.

표 7. 음소별 최고빈도 최소대립쌍
Table 7. Most frequently minimal pairs by phoneme

구분	음소 쌍	개수	비율(%)	
평음	ㄱ-ㅅ	10,023	12.33	
	ㄷ-ㄱ	5,777	13.04	
	ㅂ-ㄱ	9,063	14.86	
	ㅅ-ㅈ	12,930	17.93	
	ㅈ-ㅅ	12,930	18.27	
	장애음	ㄱ-ㄱ	1,524	14.48
		ㄷ-ㅁ	951	9.35
		ㅁ-ㅁ	1,381	18.01
		ㅁ-ㅁ	1,714	14.21
		ㅁ-ㅁ	1,714	13.94
격음		ㅈ-ㅈ	8,631	15.51
		ㅋ-ㅎ	2,043	22.55
	ㅌ-ㄱ	3,558	12.25	
	ㅍ-ㄱ	4,228	13.77	
후음	ㅎ-ㄱ	8,825	17.08	
	공명음	ㄴ-ㄱ	6,053	14.56
		ㄹ-ㄱ	3,081	15.87
ㅇ-ㄱ		8,138	14.82	

평음은 최소대립쌍이 가장 많은 음소는 모두 조음 위치 또는 조음 방법에는 차이가 있으나 발성 유형이 같은 평음이었다. 경음의 경우 /ㄱ/와 /ㅁ/는 조음 위치와 방법은 동일하나 발성 유형이 다른 평음과 가장 많은 대립 쌍을 만들었고, /ㄷ/, /ㅁ/, /ㅁ/는 조음 위치와 방법은 다르나 발성 유형이 같은 경음과 가장 많은 대립 쌍을 만들었다. 격음은 일정하지가 않았는데 /ㅈ/는 조음 위치와 조음 방법을 공유하는 /ㅈ/와, /ㅋ/는 유기성 자질을 공유하는 /ㅎ/와, /ㅌ/는 /ㄱ/와 가장 많은 대립 쌍을 만들었다. 그렇지만 대체로 조음 위치나 발성 방법을 공유하는 음소와 최소대립쌍을 만드는 빈도가 높은 것으로 보인다. 반면에 공명음은 모두 /ㄱ/와 가장 많은 최소대립쌍을 만들었다.

장애음의 최소대립쌍만 따로 조금 더 구체적으로 살펴보겠다. 표 8은 삼지적 상관속을 이루는 장애음 내의 최소대립쌍을 정리하여 나타낸 것이다.

9 Jung(2019)에서 분석한 초성의 출현 빈도는 /ㄱ, ㄷ, ㅈ, ㅅ, ㅎ, ㅂ, ㅁ, ㄹ, ㅊ, ㄴ, ㅍ, ㄱ, ㅁ, ㅁ/ 순으로 높게 나타났는데 이와는 다소 차이가 있다.

표 8. 양순음, 치경음, 경구개음, 연구개음의 최소대립쌍
Table 8. Minimal pairs of bilabials, alveolars, palatals, and velars

구분	음소 쌍	개수	비율(%)	계
양순음	ㅂ-ㅃ	1,381	28.93	4,774
	ㅂ-ㅍ	2,934	61.46	
	ㅃ-ㅍ	459	9.61	
치경음	ㄷ-ㄸ	919	20.29	4,529
	ㄷ-ㄹ	2,863	63.21	
	ㄸ-ㄹ	747	16.49	
경구개음	ㅈ-ㅉ	993	100.0	993
	ㅉ-ㅊ	1,171	10.38	
	ㅉ-ㅌ	8,632	76.53	
연구개음	ㅊ-ㅌ	1,476	13.09	11,279
	ㄱ-ㅋ	1,524	49.29	
	ㄱ-ㆁ	1,225	39.62	
연구개음	ㄱ-ㆁ	343	11.09	3,092
	ㄱ-ㆁ	343	11.09	

그림 3부터 그림 6까지는 삼지적 상관속을 이루는 장애음의 최소대립쌍 관계를 그림으로 나타낸 것이다¹⁰.

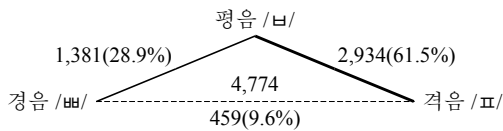


그림 3. /ㅂ-ㅃ-ㅍ/ 최소대립쌍의 비교
Figure 3. Comparison of the minimal pairs of /ㅂ-ㅃ-ㅍ/

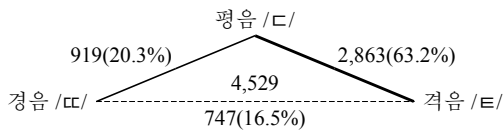


그림 4. /ㄷ-ㄸ-ㄹ/ 최소대립쌍의 비교
Figure 4. Comparison of the minimal pairs of /ㄷ-ㄸ-ㄹ/

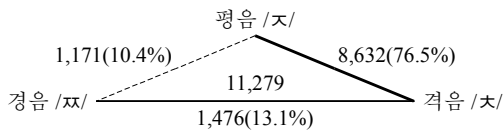


그림 5. /ㅈ-ㅉ-ㅊ/ 최소대립쌍의 비교
Figure 5. Comparison of the minimal pairs of /ㅈ-ㅉ-ㅊ/

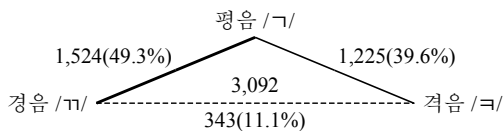


그림 6. /ㄱ-ㅋ-ㆁ/ 최소대립쌍의 비교
Figure 6. Comparison of the minimal pairs of /ㄱ-ㅋ-ㆁ/

삼지적 상관속을 이루는 장애음의 최소대립쌍은 조음 위치와 조음 방법에 따라 차이를 보이고 있다. 연구개음 /ㄱ-ㅋ-ㆁ/은 ‘평음-경음’ 쌍의 비율이 가장 높고, ‘경음-격음’ 쌍의 비율이 가장 낮았다. 경구개음 /ㅈ-ㅉ-ㅊ/은 ‘평음-격음’ 쌍의 비율이 가장 높고, ‘평음-경음’ 쌍의 비율이 가장 낮았는데 연구개음의 양상과 매우 다르게 나타났다. 치경음 /ㄷ-ㄸ-ㄹ/과 양순음 /ㅂ-ㅃ-ㅍ/은 모두 ‘평음-격음’ 쌍의 비율이 가장 높고, ‘경음-격음’ 쌍의 비율이 가장 낮았다.

많은 연구에서 한국어를 배우는 외국인 학습자들은 ‘평음-경음-격음’의 구분이 있는 파열음, 파찰음의 발음 습득에 어려움이 있다고 한다. 모국어의 영향에 따라 차이가 있을 수 있겠지만 /ㄱ, ㄷ, ㅂ, ㅈ/의 ‘평음-경음-격음’ 쌍에 대한 습득 난이도는 같지 않을 것이다. 음소 쌍별로 만들 수 있는 최소대립쌍의 수는 차이를 보이므로 음소 쌍마다 기능부담량이 다를 것이라고 할 수 있다. 최소대립쌍의 개수를 기반으로 기능부담량을 설정하여 발음 교육의 우선순위를 매긴다면 표 9와 같을 것이다.

표 9. 최소대립쌍 개수 기반 기능부담량
Table 9. Functional load based on the raw count of minimal pairs

	음소 쌍	최소대립쌍 개수	
High	/ㅈ-ㅉ-ㅊ/	8,631	
	/ㅂ-ㅃ-ㅍ/	2,934	
	/ㄷ-ㄸ-ㄹ/	2,863	
	/ㄱ-ㅋ-ㆁ/	1,524	
	/ㅈ-ㅉ-ㅊ/	1,475	
	/ㅂ-ㅃ-ㅍ/	1,381	
	↓	/ㄱ-ㅋ-ㆁ/	1,225
		/ㅈ-ㅉ-ㅊ/	1,171
		/ㅂ-ㅃ-ㅍ/	993
		/ㄷ-ㄸ-ㄹ/	919
/ㄷ-ㄸ-ㄹ/		747	
Low	/ㅂ-ㅃ-ㅍ/	459	
	/ㄱ-ㅋ-ㆁ/	343	

한국어 장애음의 경우 /ㅈ-ㅉ-ㅊ/ 쌍은 최소대립쌍의 수가 가장 많으므로 우선적으로 교육해야 하고, 반대로 /ㄱ-ㅋ-ㆁ/ 쌍은 최소대립쌍의 수가 적으므로 발음을 교육할 때 크게 중요하지 않다고 주장할 수도 있을 것이다.

3.2. 최소대립쌍의 품사 분석

이 절에서는 초성 위치에서 자음이 대립하는 최소대립쌍의 품사에 대해 살펴본다. 표 10은 두 단어의 품사 관계를 나타낸 분포표이다.

10 굵은 실선(—), 가는 실선(—), 점선(- -)의 순으로 최소대립쌍의 개수가 많음을 의미한다.

표 10. 최소대립쌍 단어의 품사 관계
Table 10. Syntactic category of words in the pairs

	명사	대명사	수사	동사	형용사	관형사	부사	조사	감탄사
명사	236,851 (70.25%)								
대명사	2,446 (0.73%)	46 (0.01%)							
수사	751 (0.22%)	26 (0.01%)	12 (0.00%)						
동사	1,221 (0.36%)	8 (0.00%)	1 (0.00%)	49,778 (14.77%)					
형용사	358 (0.11%)	-	1 (0.00%)	17,399 (5.16%)	4,621 (1.37%)				
관형사	2,760 (0.82%)	62 (0.02%)	62 (0.02%)	2 (0.00%)	-	532 (0.16%)			
부사	11,999 (3.56%)	162 (0.05%)	96 (0.03%)	90 (0.03%)	30 (0.01%)	271 (0.08%)	4,429 (1.31%)		
조사	808 (0.24%)	40 (0.01%)	25 (0.01%)	36 (0.01%)	6 (0.00%)	58 (0.02%)	77 (0.02%)	23 (0.01%)	
감탄사	1,545 (0.46%)	63 (0.02%)	31 (0.01%)	14 (0.00%)	3 (0.00%)	88 (0.03%)	170 (0.05%)	43 (0.01%)	91 (0.03%)

초성 최소대립쌍 단어의 품사 관계를 보면 ‘명사-명사’ 쌍이 70.25%로 가장 많고, 그다음으로 ‘동사-동사’ 쌍이 14.77%로 나타나 이 두 품사 쌍이 전체의 85% 이상을 차지했다. 전체 45개의 품사 쌍 중 39개는 1%도 되지 않았는데 그 중 ‘대명사-형용사’와 ‘형용사-관형사’의 품사 관계를 보이는 최소대립쌍은 한 쌍도 없었다. 용언인 동사와 형용사는 용언 내에서 최소대립쌍을 이루는 비율이 높고, 용언을 제외한 다른 모든 품사는 명사 또는 부사와의 쌍이 많은 것이 특징적이다. 즉, 용언은 특정 품사와 최소대립쌍을 이루는 비율이 높은 반면에 부사는 모든 품사와 자유롭게 최소대립쌍을 이룬다고 할 수 있다.

표 11은 음소 쌍별 최소대립쌍 중 품사가 같은 단어 쌍의 개수를 조사하여 두 단어의 품사 일치 정도를 계산한 후, 일치율이 높은 상위 10%와 일치율이 낮은 하위 10%에 해당하는 음소 쌍을 제시한 것이다. 전체 초성 최소대립쌍의 품사 일치율은 부록 2에 제시하였다. 초성 최소대립쌍 337,135개 중 296,356개가 두 단어의 품사가 같았고 그 결과 품사 일치율은 87.91%로 상당히 높은 편이었다. 품사 일치율이 가장 높은 음소 쌍은 /ㄱ-ㄱ/(93.88%)고, 가장 낮은 음소 쌍은 /ㅉ-ㄱ/(68.60%)였다. 품사 일치율이 높은 음소 쌍은 대체로 경음과 격음으로 구성되어 있었는데 최소대립쌍의 품사 일치율은 음소 쌍의 음성적 유사성과 관련이 있다고 볼 수 있겠다.

표 11. 음소 쌍별 최소대립쌍 품사 일치 비율(상하위 10%)
Table 11. Percentage of same part of speech(top 10% and bottom 10%)

순위	초성 쌍	품사 일치율	순위	초성 쌍	품사 일치율
1	ㄱ-ㄱ	93.88	139	ㄷ-ㄱ	77.35
2	ㄱ-ㅁ	93.77	140	ㅉ-ㄱ	76.70
3	ㄱ-ㄴ	93.48	141	ㄹ-ㅁ	76.57
4	ㅁ-ㄴ	93.43	142	ㅁ-ㄱ	76.40
5	ㅁ-ㄴ	92.93	143	ㄷ-ㅇ	76.13
6	ㄱ-ㄴ	92.65	144	ㅉ-ㄱ	76.06
7	ㄱ-ㅇ	92.11	145	ㄱ-ㅇ	75.92
8	ㄷ-ㅇ	92.00	146	ㅁ-ㄴ	75.81
9	ㄷ-ㅇ	91.97	147	ㄱ-ㅁ	75.46
10	ㅁ-ㄷ	91.49	148	ㄷ-ㄴ	75.25
11	ㅁ-ㄴ	91.48	149	ㄷ-ㅁ	74.80
12	ㅁ-ㅇ	91.33	150	ㄴ-ㅁ	74.77
13	ㅇ-ㅇ	91.28	151	ㄴ-ㅁ	73.68
14	ㄴ-ㄷ	91.03	152	ㅁ-ㄷ	72.14
15	ㄱ-ㅇ	90.74	153	ㅉ-ㄱ	68.60

4. 결론

이 연구에서는 『우리말샘』의 표제어 발음을 대상으로 초성 위치에서 한국어 자음의 최소대립쌍 출현 빈도를 조사하고, 최소대립쌍을 이루는 두 단어의 품사 관계에 대해서 분석했다. 연구 결과를 요약하면 다음과 같다.

첫째, 분석 대상이 된 325,717개의 표제어에서 관찰된 초성 자음 대립의 최소대립쌍은 총 337,135개였다. 음소 쌍별 최소대립쌍의 개수는 ‘ㅉ-ㅉ, ㄱ-ㅉ, ㄱ-ㄷ, ㄱ-ㅁ, ㄱ-ㅇ, ..., ㅁ-ㄱ, ㄴ-ㅁ, ㅉ-ㄱ, ㄷ-ㄱ, ㅉ-ㄱ’ 순으로 나타났다. 고빈도 음소 쌍은 평음의 비중이 높고, 저빈도 음소 쌍은 경음의 비중이 높았다. 음소별로 최소대립쌍을 많이 이루는 빈도를 분석한 결과 ‘ㄱ, ㅉ, ㄷ, ㅁ, ㄴ, ㅎ, ㄷ, ㄴ, ㅉ, ㅉ, ㄴ, ㅉ, ㄴ, ㄱ, ㅁ’ 순으로 나타났다. 장애음 중에는 경음과 격음보다 평음이, 공명음 중에는 유음보다 비음이 최소대립쌍을 많이 만든다는 것을 확인

할 수 있는데, 경구개음의 비율이 높게 나타난 것도 특징적이었다.

최소대립쌍의 절대수치와 상대수치의 상관관계를 살펴본 결과 $r=0.937$ 의 높은 상관계수를 나타냈다. 절대수치와 상대수치에 대한 잔차를 구한 결과 잔차의 값이 큰 음소 쌍은 /ㅅ-ㅆ/, /ㅈ-ㅊ/, /ㄷ-ㅌ/, /ㄱ-ㅋ/, /ㄴ-ㄹ/, /ㄴ-ㄷ/, /ㄴ-ㄹ/였는데 이 쌍들은 음소 빈도에 비해 최소대립쌍의 개수가 상대적으로 많다고 할 수 있다. 반대로 잔차의 값이 작은 음소 쌍은 /ㄷ-ㅈ/, /ㄱ-ㅎ/, /ㄱ-ㄷ/, /ㄱ-ㅅ/, /ㄱ-ㅈ/였다.

한편, 삼지적 상관속을 이루는 장애음의 최소대립쌍 관계에도 차이가 나타났다. 경구개 파열음은 ‘평음(ㄱ)-경음(ㄲ)’ 쌍의 비율이 가장 높고, ‘경음(ㄲ)-격음(ㅋ)’ 쌍의 비율이 가장 낮았다. 반면에 파찰음은 ‘평음(ㅈ)-격음(ㅊ)’ 쌍의 비율이 가장 높고, ‘평음(ㅈ)-경음(ㅉ)’ 쌍의 비율이 가장 낮았다. 한편, 치경 파열음과 양순 파열음은 모두 ‘평음(ㄷ/ㅌ)-격음(ㅌ/ㄷ)’ 쌍의 비율이 가장 높고, ‘경음(ㄷ/ㅌ)-격음(ㅌ/ㄷ)’ 쌍의 비율이 가장 낮았다.

둘째, 최소대립쌍을 이루는 두 단어 쌍의 품사 정보 분포를 살펴본 결과 두 단어의 품사가 같은 비율은 87.91%로 나타나 최소대립쌍은 인지적으로 비슷한 범주로 묶일 가능성이 높다는 것을 확인할 수 있었다. 품사 일치율이 가장 높은 음소 쌍은 /ㄱ-ㅋ/, 가장 낮은 음소 쌍은 /ㅅ-ㅆ/ 쌍이었다. 이 결과는 이후 발음 교육을 위한 한국어 음소의 기능부담량 설정에도 활용할 수 있을 것이다.

이 연구는 사전 표제어의 발음형을 대상으로 기존 연구에서 다루지 못한 최소대립쌍의 양상을 밝혔다는 의의가 있다. 다만 모음과 종성 위치의 자음의 최소대립쌍을 다루지 못했고, 단어 구조를 반영한 어두, 어중, 어말 위치에서의 대립 관계도 다루지 못했다. 추후 이에 대해서도 연구를 진행하여 한국어 음소의 전체적인 대립 체계를 종합적으로 정리해 나가고자 한다. 이 연구의 결과는 한국어 음소와 관련된 기초 자료로서 발음 교육, 언어 병리학, 음성 공학 등의 다양한 응용 분야에서 유용하게 활용될 수 있을 것이다.

References

Brown, A. (1988). Functional load and the teaching of pronunciation. *TESOL Quarterly*, 22(4), 593-606.

Brown, A. (1991). *Pronunciation models*. Singapore: Singapore University Press.

Catford, J. C. (1987). Phonetics and the teaching of pronunciation: A systemic description of English phonology. In J. Morley (Ed.), *Current perspectives on pronunciation: Practices anchored in theory* (pp. 87-100). Washington, DC: Teachers of English to Speakers of Other Languages.

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. Amsterdam: John Benjamins.

Hall, K. C., Allen, B., Fry, M., Johnson, K., Lo, R., Mackie, S., & McAuliffe, M. (2017). *Phonological corpus tools (Version 1.3)* [Computer program]. Retrieved from <https://corpustools.readthedocs.io/>

Jin, N. (1996, October). The function load of the vowel opposition and phonological change in Korean. *Proceedings of the 1st Conference of the Seoul International Conference of Phonetic Sciences (SICOPS 1996)* (pp. 169-175). Seoul, Korea.

Jung, J. (2019). A study of phoneme frequencies based on the pronunciation in the entries of Korean dictionary. *Language and Linguistics*, 83, 179-217.

Kim, K., & Kang, B. (1997). *Frequency analysis of Hangul usage*. Seoul: Research Institute of Korean Studies, Korea University.

Kim, T. H., Lee, J. H., & Chang, M. S. (2014). A minimal pair searching tool based on dictionary. *Journal of Korean Institute of Intelligent Systems*, 24(2), 117-122.

Lee, J. H. (2010). On the selection of minimal pair in Korean. *The Korean Language and Literature*, 107, 119-137.

Lee, J. (2017). *Dictionary of Korean phonology terms*. Seoul: Youkrack.

Lee, Y. J., Kim, K. T., Jo, C. W., & Lee, S. G. (1987). Statistical properties of phonemes in Korean phonetic dictionary. *Proceedings of the 1987 Institute of Electronics Engineer of Korea, Conference* (pp. 117-121). Seoul, Korea.

Levis, J., & Cortes, V. (2008). Minimal pairs in spoken corpora: Implications for pronunciation assessment and teaching. *Towards adaptive CALL: Natural Language Processing for Diagnostic Language Assessment* (pp. 197-208).

Munro, M. J., & Derwing, T. M. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System*, 34(4), 520-531.

Shin, J. (2010). Phoneme and syllable frequencies based on the analysis of entries in the Korean dictionary. *Communication Sciences and Disorders*, 15(1), 94-106.

• 정지은 (Jieun Jung) 교신저자
한국외국어대학교 국어국문학과 박사과정 수료
서울시 동대문구 이문로 107
Tel: 02-2173-2345
Email: jieunkjung@gmail.com
관심분야: 외국어 음성 습득, 한국어 발음 교육

부록 1. 초성 최소대립쌍 개수(개수 내림차순 정렬)

순위	초성 쌍	개수	비율 (%)	누적비율 (%)	순위	초성 쌍	개수	비율 (%)	누적비율 (%)	순위	초성 쌍	개수	비율 (%)	누적비율 (%)
1	ㅅ-ㅈ	12,931	3.84	3.84	52	ㄴ-ㄷ	2,173	0.64	78.06	103	ㄷ-ㅂ	583	0.17	94.18
2	ㄱ-ㅅ	10,023	2.97	6.81	53	ㅇ-ㅈ	2,149	0.64	78.69	104	ㄱ-ㄷ	576	0.17	94.35
3	ㄱ-ㅈ	9,978	2.96	9.77	54	ㅋ-ㅈ	2,043	0.61	79.30	105	ㄱ-ㅅ	562	0.17	94.51
4	ㄱ-ㅂ	9,064	2.69	12.46	55	ㄴ-ㅈ	2,027	0.60	79.90	106	ㄷ-ㅅ	561	0.17	94.68
5	ㄱ-ㅅ	8,826	2.62	15.07	56	ㄷ-ㅈ	1,990	0.59	80.49	107	ㄱ-ㅈ	558	0.17	94.85
6	ㅈ-ㅂ	8,632	2.56	17.64	57	ㄷ-ㅇ	1,861	0.55	81.04	108	ㄱ-ㅂ	547	0.16	95.01
7	ㅅ-ㅂ	8,294	2.46	20.10	58	ㄷ-ㅂ	1,851	0.55	81.59	109	ㅇ-ㅋ	539	0.16	95.17
8	ㄱ-ㅇ	8,140	2.41	22.51	59	ㅅ-ㅈ	1,714	0.51	82.10	110	ㅈ-ㅈ	528	0.16	95.32
9	ㅂ-ㅅ	7,466	2.21	24.72	60	ㄴ-ㄷ	1,652	0.49	82.59	111	ㅈ-ㅈ	527	0.16	95.48
10	ㄱ-ㅂ	7,084	2.10	26.83	61	ㄷ-ㅂ	1,557	0.46	83.05	112	ㄷ-ㅂ	525	0.16	95.64
11	ㅂ-ㅂ	7,068	2.10	28.92	62	ㄷ-ㅂ	1,543	0.46	83.51	113	ㅂ-ㅂ	524	0.16	95.79
12	ㅇ-ㅂ	6,917	2.05	30.97	63	ㄱ-ㄱ	1,524	0.45	83.96	114	ㄷ-ㅂ	508	0.15	95.94
13	ㅇ-ㅅ	6,685	1.98	32.96	64	ㄷ-ㅂ	1,504	0.45	84.41	115	ㄷ-ㅂ	505	0.15	96.09
14	ㄱ-ㄴ	6,052	1.80	34.75	65	ㅂ-ㅂ	1,476	0.44	84.85		ㅅ-ㅂ	505	0.15	96.24
15	ㅇ-ㅂ	5,792	1.72	36.47	66	ㅂ-ㅂ	1,381	0.41	85.25	117	ㄱ-ㅇ	499	0.15	96.39
16	ㄱ-ㄷ	5,777	1.71	38.18	67	ㅈ-ㅂ	1,307	0.39	85.64	118	ㄷ-ㅇ	498	0.15	96.54
17	ㅅ-ㅂ	5,530	1.64	39.82	68	ㄷ-ㅂ	1,291	0.38	86.03	119	ㅂ-ㅂ	497	0.15	96.68
18	ㄷ-ㅂ	5,415	1.61	41.43	69	ㅂ-ㅂ	1,281	0.38	86.41	120	ㄷ-ㅂ	482	0.14	96.83
19	ㅂ-ㅂ	5,374	1.59	43.02	70	ㄱ-ㅂ	1,225	0.36	86.77	121	ㄴ-ㅂ	475	0.14	96.97
20	ㄷ-ㅂ	5,328	1.58	44.60	71	ㅂ-ㅂ	1,171	0.35	87.12	122	ㅂ-ㅂ	459	0.14	97.10
21	ㅂ-ㅂ	5,289	1.57	46.17	72	ㄷ-ㅂ	1,036	0.31	87.42	123	ㅂ-ㅂ	450	0.13	97.24
22	ㅂ-ㅂ	4,810	1.43	47.60	73	ㅅ-ㅂ	993	0.29	87.72	124	ㅂ-ㅂ	434	0.13	97.37
23	ㅇ-ㅂ	4,632	1.37	48.97	74	ㄷ-ㅂ	951	0.28	88.00	125	ㅂ-ㅂ	423	0.13	97.49
24	ㄴ-ㅂ	4,588	1.36	50.33	75	ㄷ-ㅂ	935	0.28	88.28	126	ㄷ-ㅂ	422	0.13	97.62
25	ㄷ-ㅂ	4,518	1.34	51.67	76	ㄷ-ㅂ	930	0.28	88.55	127	ㄷ-ㅂ	419	0.12	97.74
26	ㄴ-ㅇ	4,401	1.31	52.98	77	ㄱ-ㅂ	928	0.28	88.83	128	ㄷ-ㅂ	411	0.12	97.86
27	ㅇ-ㅂ	4,283	1.27	54.25	78	ㄷ-ㅂ	919	0.27	89.10	129	ㄱ-ㅂ	403	0.12	97.98
28	ㄱ-ㅂ	4,228	1.25	55.50	79	ㄱ-ㅂ	889	0.26	89.36		ㅂ-ㅂ	403	0.12	98.10
29	ㄴ-ㅂ	4,209	1.25	56.75	80	ㄱ-ㅂ	871	0.26	89.62	131	ㅂ-ㅂ	397	0.12	98.22
30	ㄴ-ㅂ	4,034	1.20	57.95	81	ㄷ-ㅂ	849	0.25	89.87	132	ㅂ-ㅂ	384	0.11	98.33
31	ㄴ-ㅂ	3,961	1.17	59.12	82	ㅂ-ㅂ	796	0.24	90.11	133	ㄱ-ㅂ	383	0.11	98.45
32	ㄷ-ㅇ	3,804	1.13	60.25	83	ㄱ-ㅂ	760	0.23	90.34	134	ㄱ-ㅂ	367	0.11	98.56
33	ㅂ-ㅂ	3,774	1.12	61.37	84	ㄷ-ㅂ	747	0.22	90.56	135	ㄱ-ㅂ	364	0.11	98.66
34	ㄷ-ㅂ	3,760	1.12	62.49	85	ㄱ-ㅂ	744	0.22	90.78	136	ㄷ-ㅂ	362	0.11	98.77
35	ㄷ-ㅂ	3,581	1.06	63.55	86	ㅂ-ㅂ	736	0.22	91.00	137	ㄷ-ㅂ	360	0.11	98.88
36	ㄱ-ㅂ	3,558	1.06	64.60	87	ㄷ-ㅂ	724	0.21	91.21	138	ㅇ-ㅂ	350	0.10	98.98
37	ㄴ-ㅂ	3,552	1.05	65.66	88	ㄱ-ㅂ	673	0.20	91.41	139	ㄱ-ㅂ	343	0.10	99.08
38	ㅇ-ㅂ	3,175	0.94	66.60	89	ㄱ-ㅂ	671	0.20	91.61	140	ㄱ-ㅂ	318	0.09	99.18
39	ㄱ-ㅂ	3,082	0.91	67.51	90	ㄱ-ㅂ	659	0.20	91.81	141	ㅂ-ㅂ	311	0.09	99.27
40	ㄴ-ㅂ	3,054	0.91	68.42	91	ㅂ-ㅂ	654	0.19	92.00	142	ㄴ-ㅂ	305	0.09	99.36
41	ㅂ-ㅂ	3,025	0.90	69.32	92	ㅂ-ㅂ	649	0.19	92.19	143	ㅂ-ㅂ	290	0.09	99.45
42	ㅅ-ㅂ	2,987	0.89	70.20	93	ㄷ-ㅂ	640	0.19	92.38	144	ㅂ-ㅂ	277	0.08	99.53
43	ㅂ-ㅂ	2,934	0.87	71.07		ㅂ-ㅂ	640	0.19	92.57	145	ㄴ-ㅂ	253	0.08	99.60
44	ㅅ-ㅂ	2,911	0.86	71.94	95	ㄱ-ㅂ	626	0.19	92.76	146	ㄷ-ㅂ	250	0.07	99.68
45	ㄷ-ㅂ	2,863	0.85	72.79	96	ㅂ-ㅂ	613	0.18	92.94	147	ㄴ-ㅂ	243	0.07	99.75
46	ㅂ-ㅂ	2,820	0.84	73.62	97	ㅅ-ㅂ	611	0.18	93.12	148	ㄴ-ㅂ	218	0.06	99.82
	ㅂ-ㅂ	2,820	0.84	74.46	98	ㅂ-ㅂ	603	0.18	93.30	149	ㅂ-ㅂ	176	0.05	99.87
48	ㅂ-ㅂ	2,641	0.78	75.24	99	ㅂ-ㅂ	600	0.18	93.48	150	ㄷ-ㅂ	121	0.04	99.90
49	ㅂ-ㅂ	2,563	0.76	76.00	100	ㅇ-ㅂ	595	0.18	93.65		ㅂ-ㅂ	121	0.04	99.94
50	ㅂ-ㅂ	2,471	0.73	76.74	101	ㅇ-ㅂ	589	0.17	93.83	152	ㄷ-ㅂ	114	0.03	99.97
51	ㅂ-ㅂ	2,275	0.67	77.41	102	ㄷ-ㅂ	586	0.17	94.00	153	ㅂ-ㅂ	89	0.03	100.0

부록 2. 초성 최소대립쌍 품사 일치율(품사 일치율 내림차순 정렬)

순위	초성 쌍	최소대립쌍 개수	품사 일치쌍 개수	품사 일치율	순위	초성 쌍	최소대립쌍 개수	품사 일치쌍 개수	품사 일치율	순위	초성 쌍	최소대립쌍 개수	품사 일치쌍 개수	품사 일치율
1	ㄱ-ㅋ	343	322	93.88	52	ㄱ-ㅈ	9,978	8,824	88.43	103	ㄱ-ㅈ	659	555	84.22
2	ㄱ-ㆁ	626	587	93.77	53	ㄷ-ㅌ	3,581	3,162	88.30	104	ㄹ-ㄹ	1,557	1,310	84.14
3	ㄱ-ㄷ	889	831	93.48	54	ㄱ-ㄱ	1,524	1,345	88.25	105	ㄹ-ㄷ	525	441	84.00
4	ㆁ-ㄷ	654	611	93.43	55	ㄴ-ㄴ	4,401	3,882	88.21	106	ㄹ-ㄴ	1,861	1,559	83.77
5	ㆁ-ㄱ	736	684	92.93	56	ㅅ-ㅅ	796	702	88.19	107	ㄱ-ㅅ	547	458	83.73
6	ㄱ-ㄴ	871	807	92.65	57	ㅅ-ㅈ	4,810	4,241	88.17	108	ㄷ-ㄹ	411	344	83.70
7	ㄱ-ㄷ	558	514	92.11	58	ㄷ-ㄷ	919	810	88.14	109	ㄷ-ㄷ	935	780	83.42
8	ㄷ-ㄷ	2,275	2,093	92.00	59	ㄷ-ㅈ	3,774	3,323	88.05	110	ㄱ-ㅈ	562	468	83.27
9	ㄷ-ㅅ	747	687	91.97	60	ㄱ-ㅈ	8,826	7,771	88.05	111	ㄹ-ㄷ	508	423	83.27
10	ㄷ-ㄷ	1,281	1,172	91.49	61	ㄷ-ㄷ	993	874	88.02	112	ㄷ-ㅅ	583	485	83.19
11	ㄷ-ㄷ	1,714	1,568	91.48	62	ㄴ-ㅈ	4,034	3,540	87.75	113	ㄷ-ㄷ	951	791	83.18
12	ㄷ-ㅅ	450	411	91.33	63	ㄴ-ㄷ	6,685	5,862	87.69	114	ㄷ-ㅈ	397	330	83.12
13	ㅅ-ㄷ	1,307	1,193	91.28	64	ㄱ-ㄴ	8,140	7,135	87.65	115	ㄴ-ㄷ	253	210	83.00
14	ㄴ-ㄷ	3,054	2,780	91.03	65	ㄱ-ㄷ	744	652	87.63	116	ㅅ-ㄷ	649	538	82.90
15	ㄱ-ㅅ	367	333	90.74	66	ㅅ-ㅈ	2,820	2,471	87.62	117	ㄱ-ㄷ	760	630	82.89
16	ㄴ-ㅅ	2,173	1,971	90.70	67	ㅅ-ㆁ	1,381	1,210	87.62	118	ㅋ-ㅅ	640	528	82.50
17	ㄷ-ㅅ	2,641	2,395	90.69	68	ㄷ-ㅈ	5,530	4,822	87.20	119	ㄹ-ㄷ	1,504	1,240	82.45
18	ㄷ-ㅅ	2,987	2,708	90.66	69	ㄱ-ㄷ	576	502	87.15	120	ㄹ-ㄷ	586	480	81.91
19	ㄷ-ㄷ	524	475	90.65	70	ㅅ-ㅈ	5,374	4,683	87.14	121	ㄷ-ㄷ	724	592	81.77
20	ㅅ-ㄷ	2,934	2,659	90.63	71	ㄷ-ㄷ	5,328	4,636	87.01	122	ㄷ-ㅅ	505	410	81.19
21	ㄴ-ㄷ	2,027	1,837	90.63	72	ㄷ-ㄷ	5,415	4,710	86.98	123	ㅋ-ㅅ	311	252	81.03
22	ㄷ-ㄷ	8,294	7,515	90.61	73	ㄷ-ㄷ	528	459	86.93	124	ㄴ-ㄷ	305	247	80.98
23	ㄷ-ㅅ	4,518	4,086	90.44	74	ㄱ-ㄷ	5,777	5,014	86.79	125	ㄷ-ㅈ	403	326	80.89
24	ㄱ-ㄷ	7,084	6,401	90.36	75	ㄴ-ㄷ	5,792	5,027	86.79	126	ㄴ-ㄷ	539	435	80.71
25	ㅅ-ㅅ	2,563	2,314	90.28	76	ㄴ-ㅅ	2,149	1,860	86.55	127	ㄴ-ㄷ	243	196	80.66
26	ㄴ-ㅅ	3,961	3,574	90.23	77	ㄱ-ㅅ	1,225	1,059	86.45	128	ㄷ-ㄷ	482	386	80.08
27	ㄱ-ㄷ	4,228	3,811	90.14	78	ㄷ-ㄷ	849	732	86.22	129	ㄷ-ㄷ	423	337	79.67
28	ㄱ-ㅅ	9,064	8,162	90.05	79	ㄱ-ㄷ	928	800	86.21	130	ㅅ-ㅈ	290	231	79.66
29	ㅅ-ㄷ	7,466	6,718	89.98	80	ㄹ-ㅅ	1,036	893	86.20	131	ㄴ-ㄷ	589	468	79.46
30	ㄴ-ㄷ	3,552	3,195	89.95	81	ㄹ-ㅅ	640	551	86.09	132	ㄷ-ㄷ	930	736	79.14
31	ㅅ-ㄷ	5,289	4,756	89.92	82	ㅅ-ㅅ	277	238	85.92	133	ㄷ-ㄷ	498	392	78.71
32	ㅅ-ㄷ	8,632	7,754	89.83	83	ㄴ-ㄷ	1,652	1,418	85.84	134	ㄷ-ㄷ	360	283	78.61
33	ㄷ-ㅅ	422	379	89.81	84	ㄹ-ㅅ	1,851	1,588	85.79	135	ㄱ-ㄷ	673	529	78.60
34	ㄷ-ㄷ	2,911	2,610	89.66	85	ㄴ-ㅅ	475	407	85.68	136	ㄱ-ㄷ	403	315	78.16
35	ㄷ-ㄷ	1,476	1,322	89.57	86	ㄷ-ㅈ	3,760	3,218	85.59	137	ㄷ-ㄷ	611	473	77.41
36	ㄷ-ㅅ	527	471	89.37	87	ㄱ-ㄷ	3,082	2,632	85.40	138	ㄷ-ㄷ	561	434	77.36
37	ㅅ-ㅈ	2,471	2,208	89.36	88	ㄱ-ㄷ	499	426	85.37	139	ㄷ-ㄷ	362	280	77.35
38	ㄴ-ㄷ	4,588	4,095	89.25	89	ㄷ-ㄷ	613	523	85.32	140	ㄷ-ㄷ	176	135	76.70
39	ㄱ-ㄷ	10,023	8,944	89.23	90	ㅅ-ㅅ	603	514	85.24	141	ㄴ-ㅅ	350	268	76.57
40	ㅅ-ㄷ	2,820	2,516	89.22	91	ㄱ-ㄷ	364	310	85.16	142	ㅅ-ㅅ	89	68	76.40
41	ㄴ-ㅅ	6,917	6,170	89.20	92	ㄷ-ㄷ	121	103	85.12	143	ㄷ-ㅈ	419	319	76.13
42	ㄱ-ㄷ	6,052	5,398	89.19	93	ㄱ-ㄷ	671	571	85.10	144	ㅅ-ㅅ	497	378	76.06
43	ㅅ-ㅅ	3,025	2,698	89.19	94	ㄱ-ㄷ	318	270	84.91	145	ㅋ-ㅅ	2,043	1,551	75.92
44	ㄴ-ㄷ	4,209	3,750	89.09	95	ㄷ-ㄷ	3,804	3,224	84.75	146	ㅅ-ㅅ	434	329	75.81
45	ㄱ-ㅅ	3,558	3,168	89.04	96	ㄷ-ㄷ	1,990	1,685	84.67	147	ㄱ-ㅅ	383	289	75.46
46	ㄷ-ㄷ	12,931	11,509	89.00	97	ㄴ-ㅈ	4,632	3,914	84.50	148	ㄷ-ㄷ	505	380	75.25
47	ㅅ-ㄷ	7,068	6,280	88.85	98	ㅅ-ㄷ	1,171	988	84.37	149	ㄷ-ㅅ	250	187	74.80
48	ㄴ-ㄷ	4,283	3,804	88.82	99	ㄴ-ㄷ	595	502	84.37	150	ㄴ-ㅅ	218	163	74.77
49	ㄷ-ㅅ	2,863	2,540	88.72	100	ㄹ-ㄷ	1,291	1,089	84.35	151	ㄹ-ㅅ	114	84	73.68
50	ㄴ-ㄷ	3,175	2,810	88.50	101	ㅅ-ㅅ	600	506	84.33	152	ㅅ-ㅅ	384	277	72.14
51	ㅅ-ㄷ	459	406	88.45	102	ㄹ-ㅈ	1,543	1,300	84.25	153	ㄷ-ㄷ	121	83	68.60

한국어 음소 최소대립쌍의 계량언어학적 연구: 초성 자음을 중심으로

정 지 은

한국외국어대학교 국어국문학과

국문초록

이 연구의 목적은 한국어 음소의 최소대립쌍 출현 양상에 대해 계량언어학적으로 알아보는 것이다. 최소대립쌍은 한 언어에서 음소의 체계를 세우는 데 중요한 역할을 하고, 기능부담량의 측정에도 중요한 척도가 됨에도 불구하고 아직까지 한국어 음소의 최소대립쌍에 대한 전면적인 연구가 이루어지지 않았다. 이를 위해 『우리말샘』의 표제어 325,715개의 발음을 대상으로 초성 위치에서의 자음 최소대립쌍의 개수를 절대수치와 상대수치로 산출하고, 최소대립쌍을 이루는 두 단어의 품사 관계에 대해서 분석했다. 『우리말샘』을 연구의 대상으로 삼은 이유는 최소대립쌍 분석은 기본적으로 사전을 통해서 이루어져야 한다고 판단했고, 한국어 사전 중 규모가 가장 크기 때문이다. 연구 결과는 다음과 같다. 첫째, 최소대립쌍은 총 153가지, 337,135개였다. 개수가 많은 음소 쌍(/ㅅ-ㅈ/, /ㄱ-ㅈ/, /ㄱ-ㅊ/, /ㄱ-ㅍ/, /ㄱ-ㅎ/)은 평음의 비중이 높고, 개수가 적은 음소 쌍(/ㅃ-ㅍ/, /ㄹ-ㅃ/, /ㅆ-ㅍ/, /ㄸ-ㅍ/, /ㅉ-ㅍ/)은 경음의 비중이 높았다. 최소대립쌍 형성에 많은 역할을 담당하는 음소를 개별 음소 단위에서 살펴보면 /ㄱ, ㅅ, ㅈ, ㅍ, ㅊ/ 순으로 높게 나타났는데, 경구개음의 비율이 높게 나타난 것이 특징적이었다. 삼지적 상관속을 이루는 장애음의 최소대립쌍 관계에도 조음 위치와 조음 방법에 따라 차이가 나타났다. 최소대립쌍의 절대수치와 상대수치의 상관계수는 0.937로 높은 상관관계를 보였다. 둘째, 최소대립쌍을 이루는 두 단어의 품사는 ‘명사-명사’의 최소대립쌍이 70.25%로 가장 많았고, 그다음으로 ‘동사-동사’ 쌍이 14.77%로 나타나 이 두 유형이 전체 85% 이상을 차지했다. 초성 최소대립쌍의 품사 일치율은 87.91%로 나타나 최소대립쌍은 의미 형태적으로도 비슷한 범주로 묶일 수 있음을 확인할 수 있었다. 이 연구의 결과는 한국어 음소와 관련된 기초 자료로서 국어학, 언어 병리학, 언어 교육, 언어 습득, 음성 공학 등의 다양한 응용 분야에서 유용하게 활용될 수 있을 것이다.

핵심어: 최소대립쌍, 사전, 발음, 초성, 자음, 품사, 기능부담량

참고문헌

- 김태훈, 이재호, 장문수 (2014). 사전 기반 최소대립쌍 검색 도구. *한국지능시스템학회 논문지*, 24(2), 117-122.
- 김홍규, 강범모 (1997). *한글 사용빈도의 분석*. 서울: 고려대학교 민족문화연구소.
- 신지영 (2010). 한국어 사전 표제어 발음의 음소 및 음절 빈도. *언어청각장애연구*, 15, 94-106.
- 이용주, 김경태, 조철우, 이성구 (1987, 11월). 대용량 발음사전 표제어에 나타난 음소의 통계적 성질. *대한전자공학회 1987 학술대회 논문집*, 117-121.
- 이진호 (2010). 국어 최소대립쌍의 설정에 대하여. *어문학*, 107, 119-137.
- 이진호 (2017). *국어 음운론 용어 사전*. 서울: 역락.
- 정지은 (2019). 한국어 사전 표제어 발음형의 음소 빈도 연구. *언어와 언어학*, 83, 179-217.
- 진남택 (1996, 10월). 한국어 모음대립의 기능부담량과 음운변화. *제1회 서울 국제 음성학술대회 논문집* (pp. 169-175).