

## 핵심어 분석을 통한 성별, 연령별 발화 특성 연구\*

- 세종 구어 말뭉치를 활용하여 -

전 지 은  
(연세대학교)

### Abstract

Jeon Ji-eun, 2014. A Study on Key Words Analysis according to Social Variables of Gender and Age using Sejong Korean Spoken corpus. *The Sociolinguistic Journal of Korea* 22(1). This study shows how a key words and key clusters analysis offers an empirical data of the significant characteristics in social variables of gender and age based on Sejong spoken corpus. "Key words" is a term for statistically significant lexical items which are most frequent in a given corpus compared with other corpus as a reference (Scott 1999). We also extend the keyness concept to word combinations, clusters which are found repeatedly together in each others' company, in sequence. We use the KeyWords facility in WordSmith tools (Scott 2012) which has been used in several studies as means for describing the characteristics of social variables. We extract (1) the male and female key words/key clusters in public and private speech, (2) the 20s and over 30s key words/key clusters in public and private speech. Therefore, we found that these empirical investigations verify the lexical saliency of utterance by gender and age. The gender key words show more prominent features than age key words. Moreover, The key words analysis is more efficient than key clusters analysis in sociolinguistic studies. This study is illustrative of the potentiality of the corpus-based research on social differentiation in the use of language.

**Keywords:** social variables, gender, age, key words, key clusters, Sejong Spoken Corpus, WordSmith Tools

---

\* 이 논문은 2013년 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2013S1A5B5A07044564).

본 논문의 완성을 위하여 유용한 조언을 해 주신 세 분의 심사자들께 감사드립니다.

## I. 서론

본 연구는 세종 구어 말뭉치에서 성별, 연령별 사회적 변인에 따라 핵심어, 핵심 어휘다발을 추출하여 남성과 여성, 20대와 30대 이후의 발화 차이를 비교 분석하는 데 목적이 있다. 핵심어(Key words)란 분석 대상이 되는 대상 코퍼스(target corpus)와 기준이 될 만한 참조 코퍼스(reference corpus)의 어휘 빈도 목록(word list)을 통계 기법을 통해 비교하여 대상 코퍼스에서 특징적으로 빈번하게 나타나는 어휘를 말한다. 예를 들어, 남성이 여성에 비해 특징적으로 더 많이 사용한 어휘를 분석하기 위해 남성 발화 텍스트와 여성 발화 텍스트를 통계적으로 비교하는 것이다. 또한 본 연구에서는 개별 어휘 차원을 넘어서 어구 차원의 핵심 어휘다발(key clusters)까지 분석 층위를 확대해서 성별, 연령별 발화 특징을 살펴보고자 한다.

지금까지 사회언어학에서 사회적 변인에 따른 언어 사용 차이에 대한 연구는 주로 성별을 중심으로 이루어져 왔는데, 구어 자료의 수집과 분석의 어려움 때문에 초기에는 방송, 드라마와 같은 준구어 자료를 바탕으로 이루어졌다(민현식, 1997; 임철성, 1997; 장영희, 2000). 점차 구어 코퍼스 구축이 확대되어 계량언어학적인 양적 분석 방법으로 통계 분석이 각광 받고 있는데, 성별에 따른 정도부사 및 부사(김혜영, 2009; 전지은, 2010), 보조용언(권순구 2007), 담화표지(임규홍, 2004; 김혜영, 2011), 맞장구(장영희·김순자 2008) 등 특정 품사 및 화용적인 기능에 초점을 맞춘 성별 발화 차이에 대한 연구가 활발하게 진행되고 있다. 또한 최근에 강범모(2013)에서는 세종 구어 말뭉치를 활용하여 성별에 따라 특징적으로 나타나는 어휘(형태소) 및 품사를 바탕으로 남성과 여성을 판별하는 연구를 한 바 있다. 한편, 연령별 코퍼스 활용 연구는 주로 대학생 대화를 중심으로 음운·어휘·형태·통사·담화의 차원에서 언어 사용 양상 차이를 개괄적으로 살펴본 연구(김혜숙, 2002; 서상규·구현정, 2005)를 제외하고는 전무한 실정이다. 또한 성별, 연령별 어휘 사용 특징을 함께 살펴 본 연구는 드라마 대사를 대상으로 한 연구(임철성 1997)와 초등학생부터 30대까지 대화 자료를 바탕으로 맞장구 수행에 대한 연구(장영희·김순자 2008)가 유일하다. 이처럼 기존의 연구들은 특정 언어 현상에 국한하여 한정된 자료를 바탕으로 성별, 연령별 발화 차이를 분석하였는데, 본 연구에서는 공적/사적 상황에 따라 성별 뿐 아니라 연령별 핵심어 분석을 통해 전반적인 품사별 분포 및 특징을 비교 분석하고 핵심 어휘다발 분석을 통해 어구 차원까지 종합적으로 살펴보고자 한다.

사회적 변인에 따른 발화의 차이에 대한 연구는 지속해서 연구할 만한 흥미로운 주제이나 어떤 방법으로 그 차이를 조사하고 조사 결과에 대한 분석을 어떻게 하는가가 중요하다(전혜영 2006: 67). 본 연구의 연구 방법인 핵심어 분석을 적용한 가장 대표적인

사회언어학 연구는 Rayson et al.(1997)이다. 이 연구에서는 영국 영어 코퍼스 BNC(British National Corpus) 중 구어 일부인 약 450만 단어 텍스트를 활용하여 발화자의 성별, 연령, 지역, 교육 수준에 따른 핵심어를 추출하여 어휘 사용 차이에 대해 분석하였다. 그 결과 남성은 비속어나 금기어(fucking, shit), 숫자(hundred, one, three, four), 간투사(er, yeah, aye, okay, ah, hmm), 관사(the), 전치사 그리고 일반명사를 더 많이 사용하고, 여성은 여성을 지칭하는 대명사(she, her, hers)나 일인칭 대명사(I, me, my, mine), 간투사(yes, mm), 부사(really) 그리고 고유명사를 더 많이 사용한다는 사실을 발견하였다. 또한 연령 별로는 35세를 기준으로 두 그룹으로 나누어서 35세 이하 발화자는 축약형(wanna, gonna, yep)과 비속어(fucking, fuck, crap), 형용사, 부사 등을 많이 사용한다고 주장하였다. 반면에, 35세 이상 발화자는 주저어(hesitation)인 'well, mm, er' 등과 'says, said'와 같은 단어를 더 많이 사용한다고 주장한 바 있다.

이러한 연구를 토대로 본 연구에서는 한국어의 성별, 연령별 핵심어, 핵심 어휘다발을 추출하여 아래의 연구 질문에 대한 대답을 하고자 한다.

- (1) 가. 공적/사적인 상황에서 성별, 연령별 핵심어는 어떤 품사에서 특징적으로 나타나는가? 구체적으로 품사별로 어떤 어휘들이 특징적으로 나타나는가?
- 나. 공적/사적인 상황에서 성별, 연령별 특징적으로 나타나는 어구가 있는가?
- 다. 성별, 연령별 핵심어, 핵심 어휘다발은 어떤 변인에 의해서 더 특징적으로 나타나는가? 변인간의 상관관계가 있는가?
- 라. 성별, 연령별 핵심어와 핵심 어휘다발은 서로 상관관계가 있는가?

본 논문의 구성은 다음과 같다. 2장에서는 본 연구 대상인 세종 구어 말뭉치의 사회적 변인별 구성 및 크기를 살펴보고, 연구 방법인 핵심어, 핵심 어휘다발 분석에 대한 구체적인 방법을 제시하였다. 3장에서는 성별, 연령별 핵심어를 품사별로 비교 검토하였고 4장에서는 성별, 연령별 핵심 어휘다발을 추출하여 그 특징을 살펴보았다. 5장은 본 논문의 결론이다.

## II. 연구 대상 및 방법

### 1. 연구 대상: 세종 구어 말뭉치

본 연구에서 활용한 자료는 21세기 세종계획 구어 전사 말뭉치 최종 배포판으로, 형태소 분석 말뭉치 약 80만 어절을 대상으로 하였다.<sup>1)</sup> 세종 구어 말뭉치는 현재 국내에서 구축된 구어 말뭉치 중 가장 규모가 크고, 대화 참여자의 성별, 연령 등의 사회적 변인 정보가 포함되어 있어서 사회언어학적 연구에 적합하다.

서상규·김형정(2005: 19-26)에 따르면, 구어 말뭉치의 구성은 크게 텍스트 유형과 사회적 변인으로 나누어 볼 수 있다. 텍스트의 유형은 담화 참여자들 사이의 '상호작용성'을 기준으로 대화와 독백으로, 담화 참여자간의 관계의 '공공성'에 따라 공적인 텍스트와 사적인 텍스트로 나뉜다. 텍스트 유형별 분포는 어절수를 기준으로 사적 대화가 가장 많은 비율(37.83%)을 차지하고, 공적 독백(27.14%), 공적 대화(22.44%), 사적 독백(12.59%) 순으로 나타난다. 그러나 독백의 경우 무심코 내뱉는 혼잣말을 제외하고는 절대로 혼자서 행하는 발화의 경우는 찾아보기 어렵고, 강연, 강의, 발표, 설교 등의 공적 독백과 수업, 토론, 방송 상담 등의 공적 대화를 철저히 분류하기가 어려운 점이 있으므로 본 연구에서는 보다 확연한 차이를 드러내는 공공성 측면인 공적/사적 상황에 한정하여 살펴보았다. 또한 강범모(2013: 5)에 따르면 “남녀 발화의 문체 차이가 공사 상황의 발화의 문체 차이보다 크지 않다는 것, 그리고 아울러 남녀 차이와 공사 차이가 완전히 독립적인 차원이 아니라는 점”을 지적한 바 있다. 따라서 단순히 성별, 연령별 발화 전체를 대상으로 분석하는 것보다 공적인 상황과 사적인 상황을 구분하여 살펴보는 것이 더 적합하다.

한편 사회적 변인을 기준으로 성별과 연령별 발화를 대상으로 활용하였는데, 성별의 경우, 어절수 기준으로 남성 발화(49.83%), 여성 발화(49.63%)가 비슷한 비율로 균형적인 반면에, 연령과 직업은 20대 학생이 대부분을 차지하는 것으로 나타났다.<sup>2)</sup> 세종 구어 말뭉치에서 연령은 10대, 20대, 30대, 40대, 50대, 60대 이후로 나뉘는데, 어절 수 기준으로

- 
- 1) 본 연구에서 사용한 자료는 2009년에 배포된 것으로 당시 최재웅 교수님께서 연구에 활용할 수 있게끔 자료를 추출하는 데 도움을 주셨다.
  - 2) 서상규 외(2013: 101)에서 “20대가 높은 비율을 보이는 것은 자료 수집의 용이성이라는 측면도 있지만, 20대 화자들이 다른 연령대에 비해 신조어, 유행어의 사용 등 역동성과 생산성에 민감하고 사회에서 언어적 영향력을 행사할 수 있는 가능성이 높다는 점 등이 고려했 것”이라고 언급한 바 있다.

20대가 거의 60%를 차지하고 40대(17.46%), 30대(12.96%), 50대(2.49%), 10대(2.18%), 60대 이후(0.83%) 순의 분포를 보였다.

이와 관련하여 본 연구의 분석 방법인 핵심어 추출에서 가장 고려해야 할 사항은 비교하는 코퍼스간의 구성과 크기이다. 핵심어는 연구 대상 코퍼스(target corpus)와 비교가 되는 참조 코퍼스(reference corpus)간의 빈도 목록을 비교하여 통계적으로 두드러지는 어휘로, 참조 코퍼스의 구성과 크기에 따라 추출되는 목록이 달라질 수 있다. 일반적으로 크기 면에서는 참조 코퍼스가 대상 코퍼스와 크기가 비슷하거나 커야 한다. 구성면에서는 표준 비교와 맞비교가 있는데, “표준 비교는 여러 개의 대상 코퍼스를 분석할 때 일반적인 언어 현상을 모두 포괄할 만한 대규모 코퍼스를 참조 코퍼스로 비교하는 것이고, 맞비교는 서로 대조가 되는 구성으로 크기가 비슷한 두 개의 코퍼스를 서로 비교하는 것이다”(전지은 2011: 42). 남녀 성별 핵심어를 추출할 때는 맞비교를 하는 것이 일반적인데, 남성 핵심어를 추출할 때 남성 코퍼스를 대상 코퍼스로, 여성 코퍼스를 참조 코퍼스로 하는 것이다. 연령별 핵심어를 추출할 때는 10대, 20대, 30대, 40대, 50대, 60대 이후 각 연령별 대상 코퍼스와 대규모 참조 코퍼스와 표준 비교할 수도 있지만, Rayson et al.(1997)에서는 35세를 기준으로 이전과 이후로 나누어 맞비교하여 핵심어를 추출하였다. 이 기준을 참조하여 본 연구에서는 30세를 기준으로 30세 이전인 10대, 20대와 30대 이후인 30대, 40대, 50대, 60대 이후를 비교할 수 있는데, 10대의 발화는 2개의 텍스트에 국한되어 나타나므로 오히려 20대 발화에 초점을 맞추는 게 더 정밀한 연구가 될 것이라는 판단을 하여 20대와 30대 이후 발화를 비교하기로 하였다.<sup>3)</sup>

또한 각 변인별 맞비교 시 남성과 여성 발화 코퍼스는 크기가 비슷하나, 연령의 경우에는 공적인 상황에서는 30대 이후 발화의 크기가 더 크고, 사적인 상황에서는 20대 발화의 크기가 현저히 크다. 즉, 공적인 상황에서 20대의 발화는 약 8만, 30대 이후의 발화는 약 21만인데, 20대의 발화는 주로 ‘발표와 토론’으로 이루어져 있다. 따라서 30대 이후의 발화도 ‘토론, 발화, 워크샵, 강연’으로 한정하여 발화 크기와 구성을 동등하게 맞추었다. 또한 사적인 상황에서 20대 발화는 약 40만 정도이고 30대 발화는 약 5만 정도이므로 20대 학생 발화 중에 전화대화, 주점대화, 주제대화를 제외한 일상대화와 독백에 한정하였다. 따라서 공적인 상황 특히 발표와 토론 시에 20대와 30대 이후의 발화와 사적인 상황에서 20대와 30대 이후의 발화를 비교할 수 있다. 다음은 본 연구에서 활용한 남성과

3) 익명의 심사자께서 30대를 기준으로 구분한 것에 대해 요즘은 30대 이전과 이후는 일상생활의 언어 사용이 큰 차이가 없을 것이라는 지적을 하셨는데, 세종 구어 말뭉치는 20대 대화가 절반 이상을 차지하고 있는데, 핵심어 추출 시 비교하는 코퍼스의 크기가 비슷하거나 커야 하므로, 30대를 기준으로 분석한 것이다.

여성, 20대와 30대 이후 발화의 어절수를 제시한 표이다.

〈표 1〉 공적/사적 성별, 연령별 텍스트 어절수

		공적	사적
성별	남성	148,987	218,258
	여성	147,800	229,661
연령별	20대	58,398	55,710
	30대 이후	64,903	47,405

일반적으로 코퍼스를 활용한 연구에서 코퍼스의 크기는 어절수를 기준으로 나타내는데, 공적인 상황에서 성별에 따라 남성 발화와 여성 발화는 약 14만 어절씩, 사적인 상황에서는 약 22만 어절씩으로 구성되어 있다. 또한 연령별로 공적인 상황에서 20대와 30대 이후는 약 6만 어절씩, 사적인 상황에서는 약 5만 어절씩이다.

## 2. 연구 방법: 핵심어, 핵심 어휘다발 분석

대규모의 코퍼스를 보다 효율적으로 다루기 위해서는 언어 자료를 적절히 처리해주는 도구를 활용하는 것이 바람직하다. 일반적으로 핵심어 추출을 위해서는 WordSmith 6.0이 주로 사용된다. 핵심어를 추출하기 위해서는 세 가지 단계를 거쳐야 한다. 첫째, 연구 대상이 되는 대상 텍스트(target text)와 기준이 되는 참조 텍스트(reference text) 파일을 연다. 둘째, 각 텍스트별 빈도 목록(word list)을 생성한다. 셋째, 핵심어 추출 조건을 선택한다. 핵심어 추출 시 가장 중요한 단계로, 두 텍스트의 어휘 빈도 목록을 비교하는데 Log-likelihood 혹은 Chi-square의 통계식을 선정해야 한다. 그리고 통계적으로 유의미한 핵심어를 추출하기 위한 유의도 선정을 한다. 이때 어떠한 통계식, 유의도를 선정하느냐에 따라 핵심어의 목록이 약간씩 달라질 수 있는데, 본 연구에서는 워드스미스에서 초기 설정으로 지정되어 있는 Log-likelihood 통계식과 유의도  $p=0.000001$ 로 핵심어를 추출하였다.<sup>4)</sup>

또한 개별 어휘 차원인 핵심어 목록을 제시하는 것을 넘어서 구 차원에서 핵심 어휘다발도 살펴 볼 것이다. 핵심 어휘다발(key clusters)이란 코퍼스에서 빈번하게 발생하는

4) 이와 관련하여 워드스미스 매뉴얼에는 Log-likelihood가 더 유의미한 결과를 제공한다고 제시한 바 있고, 유의도는  $p=0.05$ 부터  $p=0.000001$ 까지 제한할 수 있는데 유의도가 낮을수록 추출되는 핵심어 개수가 줄어드는 엄격한 기준이 되므로 초기 설정을 그대로 따랐다. 자세한 논의는 전지은(2011)을 참조할 수 있다.

빈도수가 높은 연결 어휘 표현으로, McEnery et al.(2006)에 따르면 주로 3단어, 4단어 연속된 표현 중 4단어 연속체가 사용역(register)에 따른 언어 변이형을 가장 잘 드러낸다고 언급한 바 있다. 특히 대화체와 학술산문에서의 4단어 연속체를 비교한 Biber et al.(2003) 연구는 많은 연구들의 시작점이 된다. 핵심 어휘다발은 핵심어와 동일한 방식으로 추출되는데 분석 단위가 확장된 것으로 핵심어는 단어 단위라면, 핵심 어휘다발은 단어 연속체의 어구 단위이다. 즉, 자주 공기해서 나타나는 단어 연쇄를 계량적으로 추출한 것으로, 위드스미스 매뉴얼(Scott 2012)에 따르면 어휘다발의 크기는 최소 2단어에서 최대 8단어까지 설정할 수 있는데, 본 연구에서는 기존 연구 결과를 토대로 4단어 연속체를 기준으로 추출하였다.<sup>5)</sup>

한편, 한국어 핵심어, 핵심 어휘다발 추출 시에 분석 단위를 고려해야 하는데, 한국어는 영어와 달리 교착어의 특성을 지니고 있으므로 띄어쓰기 단위로 하면 조사나 어미 등의 특성을 반영할 수 없다. 따라서 띄어쓰기의 어절 단위가 아닌 형태소 단위로 추출하였다.<sup>6)</sup> 또한 세종 구어 말뭉치에는 형태소에 따라 다음과 같이 33개의 품사 및 문법 정보가 부착되어 있다.

(2) 세종 구어 말뭉치 형태소 품사 및 문법 정보

- ㄱ. 체언(5): 명사NN(일반명사NNG, 고유명사NNP, 의존명사NNB), 대명사NP, 수사NR
- ㄴ. 용언(5): 동사VV, 형용사VA, 보조용언VX, 지정사VC(궁정지정사VCP, 부정지정사VCN)
- ㄷ. 수식언(3): 관형사MM, 일반부사MAG, 접속부사MAJ
- ㄹ. 독립언(1): 감탄사IC
- ㅁ. 관계언(9): 격조사JK (주격JKS, 보격JKC, 관형격JKG, 목적격JKO, 부사격JKB, 호격JKV, 인용격JKQ) 조사사JX, 접속조사JC
- ㅂ. 의존형태(10): 어미(선어말EP, 종결EF, 연결EC, 명사형전성어미ETN, 관형형전성어미ETM) 접두사(체언접두사 XPN), 접미사(명사파생XSN, 동사파생XSV, 형용사파생XSA) 어근XR

이러한 정보를 바탕으로 추출된 성별, 연령별 핵심어는 체언, 용언, 수식언, 독립언,

- 5) 예비실험을 한 결과로도 3단어일 경우 추출된 어휘다발이 서로 중첩되는 부분이 많았고, 5단어인 경우에는 추출되는 어휘다발의 개수가 너무 한정적이므로 4단어가 가장 적절하였다.
- 6) 이와 관련하여 단어 중심으로 번역한 ‘핵심어’와 ‘핵심 어휘다발’이라는 용어를 형태소 차원으로 고려한다면 ‘핵심 형태소’, ‘핵심 형태소다발’이라 번역할 수도 있지만, 기존의 ‘핵심어’라는 개념이 특정 텍스트에서 주로 다루고 있는 ‘주제어’로서 그 텍스트의 특징을 대표할만한 단어로 할 때, 보편적으로 알려진 용어를 사용하기로 한다.

관계인, 의존형태의 대분류로 구분하여 비교 분석하였다.

### III. 사회적 변인별 핵심어 분석

#### 1. 성별 핵심어

성별 핵심어 추출 결과 공적인 상황에서 여성 64개, 남성 57개, 사적인 상황에서 여성 82개, 남성 82개가 추출되었는데, 성별에 따른 어휘 사용 차이뿐 아니라, 남성 혹은 여성이 공적/사적 상황에 따라 선호하는 어휘에도 차이가 있음을 알 수 있다. 우선, 공적/사적 상황별 성별 핵심어의 품사별 비율을 비교해 보면 다음과 같다.<sup>7)</sup>

##### (3) 공적/사적 성별 핵심어 품사별 비율

- ㄱ. 일반명사: 공적여성(56.25) > 사적남성(25.61) > 공적남성(22.81) > 사적여성(21.95)
- ㄴ. 대명사: 공적남성(5.26) > 사적남성(4.88) > 공적여성(3.13) > 사적여성(2.44)
- ㄷ. 동사: 사적남성(12.20) > 사적여성(10.98) > 공적남성(7.02) > 공적여성(4.69)
- ㄹ. 형용사: 사적여성(6.10) > 사적남성(2.44) > 공적남성(1.75)
- ㅁ. 보조용언: 사적여성(3.66), 사적남성(3.66) > 공적남성(1.75), 공적여성(1.56)
- ㅂ. 일반부사: 사적여성(13.41), 사적남성(9.76) > 공적남성(3.51), 공적여성(3.13)
- ㅅ. 접속부사: 사적여성(3.66) > 공적남성(1.75), 공적여성(1.56), 사적남성(1.22)
- ㅇ. 감탄사: 사적여성(12.20) > 공적남성(8.77) > 사적남성(4.88) > 공적여성(4.69)
- ㅈ. 종결어미: 공적남성(10.53) > 사적여성(9.76) > 공적여성(9.38) > 사적남성(3.66)

(3)의 분포를 보면, 공적/사적 성별에 관계없이 일반명사가 가장 많이 나타나고, 특히 공적 여성에서 56.25%로 핵심어의 절반 이상을 차지하였다. 일반적으로 기존 연구에 따르면 일반명사는 여성보다 남성이 더 많이 사용하는 것으로 알려져 있다(Rayson et al., 1997; 강범모, 2013). 그러나 공적/사적 상황을 구분하여 살펴보면, 발표, 강연과 같은 공적인 상황에서는 여성이 발화의 주제와 관련된 내용어인 일반명사를 남성보다 많이 사용하고, 사적인 상황에서는 남성보다 적게 사용한다는 점을 알 수 있다. 다음으로 용언인 동사, 형용사, 보조용언과 수식어인 일반부사, 접속부사는 남녀 모두 공적인 상황보다

7) (3ㄹ)의 형용사는 공적 여성에서 추출된 핵심어가 없었고, (3ㅁ)의 보조용언의 경우 공적남성(1.75), 공적여성(1.56)으로 남성의 비율이 약간 더 큰 듯하지만 해당하는 핵심어 개수가 1개씩이므로 부등호가 아닌 쉼표로 동등한 비율로 처리하여 제시하였다.

사적인 상황에서 핵심어가 더 많이 나타났는데, 이 중에서 동사만 여성보다 남성에서 더 많은 비율로 나타났고, 나머지는 다 사적 여성에서 가장 특징적으로 나타났다. 그러나 공적인 상황에서는 성별에 따른 큰 차이가 나타나지 않았다. 그러므로 기존 연구에서 여성이 남성에 비해 형용사와 부사, 감탄사를 더 빈번히 사용한다는 주장(민현식, 1997; 장영희, 2000; 전해영, 2006)은 사적인 상황에서 더 확연하게 드러남을 알 수 있다. 한편, 대명사, 종결어미의 경우에는 공적 남성에서 가장 많은 비율로 나타났는데, 특히 대명사는 기존 연구(Rayson et al., 1997; 민현식, 1997)에서 여성의 특징으로 보는 것과 달리 공적, 사적 상황 모두에서 남성이 더 많이 사용하는 것으로 나타났다.

이처럼 공적/사적 사용역별 성별 핵심어 목록을 보다 체계적으로 분석하기 위해 체언, 용언, 수식언, 감탄사, 관계언, 의존형태 품사별로 구분하여 해당하는 핵심어를 상세히 살펴보고자 하겠다.<sup>8)</sup> 다음은 공적/사적 성별 핵심어를 품사별로 제시한 표이다.<sup>9)</sup>

〈표 2〉 공적/사적 성별 핵심어 품사별 분포

	품사	여성 핵심어	남성 핵심어
공 적	일반명사	집단, 그림, 대화, 에너지, 근대, 변인, 친구, 서비스, 엄마, 광고...	기도, 말씀, 발음, 개발, 팀, 땅, 활동, 사업, 지원, 문제, 소리...
	대명사	이거, 그거	요것, 저희, 요거
	의존명사	없음	헤르쯔, 차, 개, 분, 키로헤르쯔, 년
	동사/ 형용사	보, 듣, 알 / 없음	들어가, 넘어지, 잡, 치 / 그렇
	보조용언	보	가지
	일반부사	굉장히, 정말	아마, 지금
	접속부사	그까	하여튼
	감탄사	음, 이제, 예	그, 네, 아, 예, 뭐

8) 공적/사적 사용역별 성별 핵심어 중 다음의 예는 분석에서 제외하였다. 강범모(2013)에 따르면 접미사의 경우 남녀 성별 판별에 거의 기여하지 못한다고 주장한 바 있다.

ㄱ. 공적여성: 명사파생접미사(이), 어근(중요), 고유명사(진아, 조선)

ㄴ. 공적남성: 명사파생접미사(자, 제), 어근(하), 관형사(요), 수사(일, 백)

ㄷ. 사적여성: 명사파생접미사(님), 선어말어미(었, ㄱ ㅁ)

ㄹ. 사적남성: 명사파생접미사(적, 이, 화, 간), 동사파생접미사(하), 관형형어미(르, 란), 고유명사(미국, 명수), 수사(구, 이, 투, 일, 원)

9) 본 논문에서는 지면 관계상 각각의 개별 빈도와 핵심도의 수치를 제공하지 못한 아쉬움이 있다. 본 논문에서 추출한 핵심어는 모두 최소 빈도 3회 이상 나타난 것으로 텍스트 전반에 고루 분포되어 있는지에 대한 산포도도 고려하여 특정 부분에서 치우쳐서 나타나는 어휘는 배제하였다.

	종결/ 연결어미	아요, 어요, 거든요, 죠, 면 / 라구, 구, 면서, 면	습니다, ㅂ니다, 습니까, 느냐, 냐, 지 / 고
	조사	으루 / 두	로, 으로/께서/마는/이라든가, 라든 가
사 적	일반명사	언니, 오빠, 엄마, 아빠, 집, 아저씨, 여행, 남자, 선생, 째즈...	형, 군대, 술, 싸끼, 공군, 나라, 동아 리, 정치, 행사, 전쟁, 경우, 동기...
	대명사	저, 제	저거, 니, 뭐, 거
	의존명사	없음	수, 놈, 개월, 년
	동사/ 형용사	사, 맛, 먹, 듣, 다니, 넣, 내, 이리, 나타나 / 예쁘, 맛있, 좋, 싫, 어리	쏘, 치, 없애, 나오, 막, 나가, 넘어가, 만들, 굶, 피하 / 있, 야하
	보조용언	갖, 싶, 하	가지, 버리, 나
	일반부사	되게, 너무, 정말, 너무너무, 그냥, 진짜, 잘, 같이, 막, 이렇게, 이제	우선, 존나, 좀, 지금, 솔직히, 물론, 떡, 갑자기
	접속부사	그래서, 근까, 그러구	그럼
	감탄사	음, 어, 어머, 네, 아, 예, 허, 하, 아 우, 헤	똥, 저, 씨발, 씨
	종결/ 연결어미	어요, 아요, 대요, 애요, 예요, ㅏ요, 구, ㅑ요 / 구, 면서, 나, 다구	지, 냐, 더라구 / 면은, 나면은, 지만, 다, 고
	조사	에게 /을 / 두, 요, 는	으로, ㄴ

성별 핵심어 목록에서 가장 큰 부분을 차지하는 것은 일반명사인데, 공적인 상황에서는 성별에 따른 차이보다 발화 주제와 내용의 차이가 더 크게 나타났다. 그럼에도 불구하고 일반명사 ‘엄마’는 공적/사적 여성 핵심어 목록에 모두 나타남으로 발화 상황에 관계없이 여성이 선호한다고 볼 수 있다. 사적인 상황에서는 일반명사 핵심어가 성별에 따라 의미적인 차이가 나타나는데, 여성은 ‘언니, 오빠, 엄마, 아빠, 집’ 등과 같이 가족 관련 명사가 주로 나타나고 남성은 ‘군대, 공군, 나라, 정치, 전쟁’ 등 국가 관련 명사와 ‘형, 술, 싸끼, 동아리, 동기’ 등 친구와의 친목 관련 명사가 특징적으로 나타난다.

다음으로 대명사를 살펴보면, 기존 연구에서는 여성 언어의 특징으로 보는데, Rayson et al.(1997)은 여성이 일인칭 대명사를 더 많이 사용한다고 주장하였고, 민현식(1997)에서는 여성이 ‘요것, 고기, 조것’ 과 같은 지시대명사를 특징적으로 사용한다고 보았다. 그런데 본 연구 결과 공적인 상황에서 남성과 여성 모두 지시대명사가 특징적으로 나타났고, 사적인 상황에서는 여성이 일인칭 대명사를, 남성은 이인칭 대명사, 지시 대명사의 쓰임이 두드러졌다. 먼저, 공적인 상황에서 여성은 ‘이거, 그거’, 남성은 ‘요것, 요거’를 주로 사용하는데, 일반적으로 ‘요것, 고것, 요기, 조기, 요거, 고거’ 등의 지시사는 작고 귀여운

어감의 어휘로 여성이 선호하는 어휘(전혜영 2006: 54)라고 주장하지만, 공적인 상황에서 남성이 사용한 ‘요것, 요거’는 설명 도중에 ‘이것’을 강조하며 낮잡아 이를 때 사용하는 것으로 보인다.

- (4) 가. 맞아. 그럼 이거 이거 약분하면 뭐야? (공적 여성)
- 나. 요것과 요것이 이렇게 와서 이 전자가 이쪽으로 이렇게 와서 페어 쌍을 이루면서.. (공적 남성)
- 다. 요게 움직여 가지고 요게 물을 갖다가 차곡차곡 움직여요. (공적 남성)

이처럼 강연, 발표 등 상황에서 화자가 자신이 말하고자 하는 대상을 지칭할 때 사용하는 지시대명사가 성별에 따라 차이를 보임을 알 수 있다. 한편, 사적인 상황에서 여성은 자신을 상대방에 비해 낮추어 지칭하는 ‘저, 제’의 일인칭 대명사의 쓰임이 두드러지는 특징을 보인다. 반면에 남성은 친구나 자신보다 어린 상대방을 지칭할 때 사용하는 이인칭 대명사이며 방언인 ‘니’가 핵심어로 추출되었다. 또한 남성은 모르는 사실이나 사물을 가리키는 지시 대명사 ‘무어’의 줄임말 ‘뭐’와 ‘거기’를 구어적으로 이르는 줄임말인 ‘거’의 쓰임이 두드러지는데 다음과 같은 예에서 나타난다.

- (5) 가. 아까 우리두 그 맥도날드 가서 뭐 먹을라구 그러다가 (사적 남성)
- 나. 그거 뭐, 등록금 비싼 거 그거 뭐 (사적 남성)
- 다. 아, 거 되게 궁금하네. (사적 남성)
- 라. 거기는 요새 거 뭐지? 백조같은 그런 보트 있잖아.(사적 남성)

(5가)에서 ‘뭐’는 어떤 사물을 지칭하는 지시 대명사로 쓰였는데, (5나)에서 ‘뭐’는 시간을 벌기 위한 담화표지로 사용되고 있다. 또한 (5다)의 ‘거’는 이야기 대상이 되는 것을 구어적으로 지칭하는데, (5라)에서처럼 이야기 대상이 생각이 잘 안 날 때, 발화를 이어가기 위해서 주로 사용되기도 한다. 이와 같이 사적인 상황에서 남성이 지시 대명사를 다양하게 사용한다는 것을 알 수 있다.

한편, 체언 중 의존명사의 쓰임은 공적/사적인 상황 모두 남성에서만 나타났다는 점이 흥미롭다. 의존 명사는 수와 관련된 세는 단위가 주로 나타나는데, 수사와 함께 의존명사는 남성이 선호하는 품사이다.

용언의 경우에는 공적인 상황보다 사적인 상황에서 더 많이 나타나고, 동사는 여성에 비해 남성이, 형용사는 사적 여성에서 보조용언은 여성과 남성이 비슷한 분포로 나타났다. 동사의 경우, 공적인 상황에서 여성은 ‘보다, 듣다, 알다’의 인지지가 동사가 특징적으

로 나타난 반면에, 남성은 ‘들어가다, 넘어지다, 잡다, 지다’와 같이 동작과 관련된 동사가 두드러진다. 또한 사적인 상황에서 여성은 ‘사다, 먹다, 다니다, 내다’와 같이 물건을 구매하는 쇼핑 관련 동사와 ‘듣다, 이리다’와 같이 다른 사람들의 말을 전하고 이야기하는 것과 관련된 동사가 주로 나타났는데, 특히 ‘듣다’는 공적/사적 상황에 관계없이 여성이 선호하는 경향을 보인다. 남성은 ‘쏘다, 치다, 없애다, 막다, 피하다’처럼 공격과 방어를 나타내는 단어가 특징적으로 나타났다.

형용사의 경우, 공적인 상황에서 여성은 핵심어로 추출된 것이 없으나 남성은 상태, 모양, 성질을 나타내는 대응표현 ‘그렇다’가 특징적으로 나타났는데, 예를 들어 ‘그렇죠, 뭐, 그럴 수도 있지 않겠습니까?’와 같이 선행 발화를 지칭하는 경우에 사용된다. 공적인 상황보다 사적인 상황에서 성별 차이가 잘 드러나는데, 여성은 ‘예쁘다, 맛있다, 좋다, 싫다, 어리다’와 같이 자신의 느낌과 감정을 표현하는 형용사, 남성은 ‘있다, 야하다’의 쓰임이 두드러졌다.

보조용언의 경우, 공적/사적 상황과 성별에 관계없이 이유를 나타내는 구어 표현인 ‘-어/아 가지고’의 쓰임이 특징적으로 나타나는데, 여성은 ‘-어/아 갖고’와 같이 발음상의 특징이 눈에 띈다. 보조 용언 ‘가지다’의 ‘이’ 모음 탈락과 연결어미 ‘고’가 ‘구’로 음운 변이를 보이는데, 김혜숙(2002: 22)에 따르면, 이러한 음운론적 변이 경향은 구어 발화자의 자신없음, 주체성 결여, 고집스러움 등을 보이는 특성이 있다고 한다. 또한 여성은 공적인 상황에서 시도를 나타내는 ‘-어/아 보다’의 공손 표현이 주로 쓰이고, 사적인 상황에서는 바람을 나타내는 ‘-고 싶다’의 표현이 두드러진다. 사적인 상황에서 남성은 어떤 행위가 완전히 끝남을 나타내는 표현 ‘-어/아 버리다’, 행위를 계속 이어나감을 나타내는 표현인 ‘-어/아 나가다’ 표현이 주로 쓰인다.<sup>10)</sup>

수식언의 경우도 공적인 상황보다 사적인 상황에서 남성보다 여성이 주로 사용하는데, 일반부사의 경우 여성은 공적/사적인 상황에 관계없이 강조부사의 쓰임이 두드러진다는 점이 특징적이다. 그러나 같은 강조부사임에도 공적인 상황에서는 ‘굉장히’와 같이 보다 문어체적인 어휘가 사용된 반면에, 사적인 상황에서는 ‘되게, 너무, 너무너무, 진짜, 잘, 막’ 과 같이 구어체의 특징이 잘 드러나고 감성을 나타내는 부사들이 빈번히 나타난다. ‘정말’의 경우는 공적/사적인 상황에 관계없이 여성이 선호하는 것으로 나타났는데 장영

10) ‘-어/아 버리다’는 ‘숙제를 끝내 버리다’의 예처럼 ‘부담의 제거’의 긍정적인 양태 의미와 ‘아이 스크림이 녹아 버렸다’의 예처럼 ‘아쉬움’을 나타내는 부정적인 양태 의미로 쓰일 수 있다. 권순구(2007)에 따르면 남성은 ‘부담의 제거’의 의미일 경우 여성보다 더 즐겨 쓰는데, 이는 갈등 상황의 종결이 함축되어 있고, 그에 따른 실제 행동이 있었음을 의미하므로 남성이 여성보다 돌발적이고 과단성이 있는 측면이 있다고 주장한 바 있다.

회(2000)은 남성보다 여성이 극단적인 감정 표현에 더 민감하기 때문이라고 주장한 바 있다. 남성은 공적/사적인 상황 모두에서 '지금'이라는 시간 부사의 쓰임이 두드러지고, 공적인 상황에서는 말하는 사람의 태도를 나타내는 양태 부사 '아마', 사적인 상황에서는 양태부사 '솔직히, 물론' 우선, 좀, 갑자기'의 쓰임이 특징적으로 나타났다.<sup>11)</sup> 특히 남성은 '존나, 딱'과 같이 비속어와 은어 표현의 부사가 여성과 달리 빈번히 사용하는 것으로 나타났다.

접속부사의 경우, 여성은 공적/사적 상황과 관계없이 원인과 결과를 나타내는 접속부사 '그러니까'의 축약형인 '그까, 근까'를 선호하며, 남성은 의견을 정리할 때 공적인 상황에서는 '하여튼', 사적인 상황에서는 '그럼'을 주로 사용하는 것으로 나타났다. 특히 사적인 상황에서 여성이 접속부사의 쓰임이 두드러지는데, 이야기를 계속 이어나갈 때 주로 쓰는 '그래서, 그러구'가 핵심어로 추출되었고, 사적인 상황에서도 접속부사가 표준형이 아닌 축약형과 모음변이를 보인다는 점이 두드러진다.

감탄사는 공적인 상황에서는 남성이, 사적인 상황에서는 여성이 특징적으로 나타났다. 먼저 여성은 공적인 상황에서 '음, 이제, 예', 사적인 상황에서 '음, 어, 어머, 네, 아, 예, 허, 하, 아우, 헤'로 나타났다. 특히 여성은 공적/사적 상황에 관계없이 군말에 해당하는 '음'을 선호하는데, 말할 때 주저하거나 머뭇거림 혹은 대화를 듣고 이해하고 공감하고 있음을 나타내는 맞장구 표현이다. 또한 여성은 긍정적 대담을 할 때 공적인 상황에서는 '예', 사적인 상황에서는 '네'가 주로 나타난다. 장영희(2000: 111)에 따르면, 응답어로서 여성은 '네', 남성은 '예'를 잘 쓰는 경향이 있다고 하는데, 공적/사적 상황에 따라 다르게 나타남을 알 수 있다. 그 밖에도 사적인 상황에서 다양한 감탄사가 추출되었는데, 놀람의 감정을 나타내는 '어머, 어, 아우', 숨 들이는 형태의 '허, 하, 헤' 등은 전형적으로 여성이 특징적으로 사용하는 감탄사이다. 반면에 남성은 공적/사적 상황에 관계없이 '뉘'를 선호하고, 공적인 상황에서는 '그', 사적인 상황에서는 '저'가 특징적으로 나타났다. 이와 관련하여 임규홍(2004)은 어원적으로는 지시 대명사에서 왔지만 실제 담화에서는 지시적 의미보다는 '시간별기' 기능을 하며 여성보다 남성이 더 많이 사용한다고 주장한 바 있다. 또한 사적인 상황에서 남성은 욕설인 '씨발, 씨'가 핵심어로 추출되었다.<sup>12)</sup>

11) 이는 한국어 부사의 성별 사용 차이를 연구한 전지은(2010)의 결과와도 상통하는데, 전지은(2010)은 공적/사적 사용역을 구분하지 않고 남녀 성별에 따른 부사 사용 차이를 살펴보았는데, 여성은 강조부사를, 남성은 시간부사와 양태부사의 쓰임이 특징적으로 나타남을 제시한 바 있다.

12) 그 밖에도 공적인 상황에서 '네, 아, 예'와 같은 감탄사가 핵심어로 추출되었는데, 이는 사적 여성에서도 추출된 것으로 어떤 의미로 사용되는지 구체적인 용례를 살펴야 할 것이다. 예를 들어, 감탄사 '아'의 경우, 감동, 응답, 부름, 놀람 따위의 느낌을 나타내는 것으로 맥락에 따라

마지막으로 문법적 기능을 하는 의존 형태인 어미와 관계언인 조사의 쓰임을 살펴보면, 종결어미의 경우, 여성은 공적/사적인 상황과 관계없이 비격식체인 해요체(아요, 어요, 애요, 예요, ㅏ요, ㅑ요)의 쓰임이 두드러진다. 반면에 남성은 공적인 상황에서는 격식체인 합쇼체(습니다, ㅂ니다, 습니까)가, 사적인 상황에서는 해체(지, (느)냐)가 특징적으로 나타나며, 특히 ‘습니까, (느)냐’처럼 의문문의 형태도 나타난다는 점은 남성 발화의 특징이다. 또한 공적인 상황에서는 여성과 남성 모두 상대방의 의견을 확인할 때 사용하는 ‘지요’가 특징적으로 나타나는데 여성은 해요체 축약형인 ‘쥬’, 남성은 해체인 ‘지’의 형태로 나타난다. 한편, 사적인 상황에서 여성은 다른 사람이 한 말을 전할 때 사용하는 ‘대요’, 남성은 자신의 경험을 회상하며 강조하면서 말할 때 주로 사용하는 ‘더라구’의 쓰임이 두드러진다.

연결어미의 경우, 여성은 다른 사람의 말이나 사실을 직접 인용할 때 쓰이는 ‘라구/다구’, 두 가지를 동시에 하는 것을 나타내는 ‘면서’, 여러 가지를 열거 하는 ‘구’가 두드러진다. 남성에서도 공적/사적 상황과 관계없이 열거하는 연결어미 ‘고’의 쓰임이 특징적인데, 여성은 모음변화가 일어난 점이 특징이다.

조사의 경우에 공적 남성이 가장 다양한 조사가 특징적으로 나타났는데, 부사격 조사 ‘로, 으로’ 주격조사 ‘께서’ 보조사 ‘마는’ 접속조사 ‘이라든가, 라든가’가 핵심어로 추출되었다. 특히 부사격조사 ‘(으)로’는 사적 여성만 제외하고 모두 핵심어로 추출되었는데, 공적인 상황에서 여성은 ‘으루’로 모음변이가 일어난다는 점이 특징적이다. 또한 여성은 더함의 뜻을 지닌 보조사 ‘도가’ 공적/사적 관계없이 ‘오’에서 ‘우’로 모음 변이가 일어나서 ‘두’가 특징적으로 나타났다.

이처럼 성별 핵심어를 품사별로 살펴 본 결과, 남성과 여성이 특징적으로 사용하는 어휘와 그 의미적인 특징을 비교할 수 있었다. 특히 대화의 격식이 비교적 자유로운 사적인 상황에서 성별 핵심어의 차이가 더 특징적으로 나타났는데, 격식적인 공적인 상황에서도 남성과 여성이 선호하는 품사와 어휘가 다르게 나타난다는 점도 알 수 있었다. 구체적으로 여성은 사적인 상황에서 형용사, 보조용언, 일반부사, 접속부사 감탄사의 쓰임이, 남성은 공적인 상황에서 동사, 대명사, 종결어미, 조사의 쓰임이 두드러졌다. 또한 동일 품사 내에서도 공적/사적, 성별 추출된 핵심어의 의미적 특징이 다르게 나타났다.

---

의미가 결정된다. 따라서 담화 차원에서 기능을 살펴 볼 필요가 있다.

## 2. 연령별 핵심어

연령별 핵심어는 공적인 상황에서 20대 57개, 30대 이후 45개, 사적인 상황에서 20대 36개, 30대 이후 34개가 추출되었는데, 공적/사적 상황별 연령별 핵심어의 품사별 비율을 비교해 보면 다음과 같다.<sup>13)</sup>

### (6) 연령별 핵심어 품사별 비율

- ㄱ. 일반명사: 사적20대(38.89) > 사적30대 이후(33.33), 공적20대(33.33) > 공적30대 이후(15.56)
- ㄴ. 대명사: 사적20대(16.67) > 공적30대 이후(6.67) > 사적30대 이후(3.03) > 공적20대(1.75)
- ㄷ. 동사: 사적30대 이후(12.12) > 공적30대 이후(8.89) > 공적20대(5.26) > 사적20대(2.78)
- ㄹ. 형용사: 사적 30대 이후(6.06) > 공적20대(1.75)
- ㅁ. 보조용언: 사적30대 이후(6.06) > 공적30대 이후(2.22) > 공적20대(1.75)
- ㅂ. 일반부사: 사적30대 이후(6.06) > 사적20대(5.56) > 공적30대 이후(4.44) > 공적20대(1.75)
- ㅅ. 접속부사: 공적30대 이후(4.44) > 공적20대(3.51)
- ㅇ. 감탄사: 공적30대 이후(13.33) > 공적20대(7.02) > 사적30대 이후(6.06) > 사적20대(2.78)
- ㅈ. 종결어미: 사적30대 이후(9.09) > 사적20대(8.33) > 공적20대(7.02) > 공적30대 이후(4.44)

연령별 핵심어의 품사별 분포를 보면, 체언인 일반명사와 대명사는 사적 20대에서 가장 많이 나타났다. 용언인 동사, 형용사, 보조용언과 종결어미는 공적인 상황보다 사적인 상황에서 추출된 핵심어가 더 많았는데, 특히 사적 30대 이후에서 가장 많이 나타났다는 점에서 체언과 대조적이다. 수식언인 일반부사는 공적/사적 상황에 관계없이 20대보다 30대 이후에서 더 선호하고, 접속부사는 공적인 상황에서만 핵심어로 추출되었다는 점이 특징적이다. 한편 감탄사와 종결어미는 공적 30대 이후에서 가장 특징적으로 나타났는데, 특히 감탄사는 사적인 상황에 비해 공적인 상황에서, 20대보다 30대 이후에서 핵심어가 더 많이 추출했다는 점이 주목할 만하다. 감탄사는 감정을 표현하는 기능을 하기도 하지만, 머뭇거림을 채우기 위한 담화 기능도 주로 나타나는데, 사적인 상황보다 공적인 상황에서 머뭇거림이 더 많이 발생하고, 20대보다 30대 이후에서 더 다양하게 나타난다.

이처럼 공적/사적 사용역별 연령별 핵심어 목록을 보다 체계적으로 분석하기 위해 체언, 용언, 수식언, 감탄사, 관계언, 의존형태 품사별 분포를 살펴보도록 하였다.

다음은 공적/사적 연령별 핵심어를 품사별로 제시한 표이다.<sup>14)</sup>

13) 각 품사별로 연령별 핵심어가 2개 이상 추출된 것만을 보인 예이다.

14) 다음의 예는 공적/사적 사용역별 연령별 핵심어로 추출되었지만 분석에서 제외하였다.

〈표 3〉 공적/사적 연령별 핵심어 품사별 분포

	품사	20대 핵심어	30대 이후 핵심어
공 적	일반명사	언어, 프로, 세계, 애기, 사회, 주 장, 소비, 시청, 종교, 개념...	한국어, 개발, 작업, 말씀, 분과, 소개, 수
	대명사	그거	요거, 요기, 여러분
	의존명사	없음	분, 년, 개
	동사/ 형용사	보, 나오, 나타나/ 같	되, 드리, 하, 관하/ 없음
	보조용언	갖	계시
	일반부사	되게	아마, 요렇게
	접속부사	근까, 그리구	그래서, 또는
	감탄사	음, 예, 예예, 뭐	자, 그, 아, 어, 으, 그죠
	종결/ 연결어미	거든요, 아요, 구요, 는데요	비니다, 습니다
조사	와, 이나, 망 / 으로써, 서 / / 의 / 요	이라든가, 라든가 / 께서 / 은	
사 적	일반명사	오빠, 학기, 학교, 이번, 군대, 친 구, 전공, 교수, 졸업, 대학...	신문, 집, 언니, 신혼, 살, 김밥, 사 과, 음식, 담, 아침, 외국...
	대명사	나, 내, 너, 뭐, 저거, 애	자기
	의존명사	없음	씨
	동사/ 형용사	듣 /없음	먹, 뛰, 짝, 읽/ 그렇, 맛있
	보조용언	없음	갖, 가지
	일반부사	솔직히, 왜	사실, 너무나무
	접속부사	없음	없음
	감탄사	응	음, 그
	종결/ 연결어미	어, 야, 는데	드라구, 어요, 여
조사	대로	을	

먼저, 일반명사를 보면, 성별 핵심어와 마찬가지로 공적인 상황에서는 발화 주제와 관련된 사회적이고 전문적인 영역의 일반명사들이 나타나서 연령에 따른 차이는 나타나지 않는다. 김혜숙(2002: 20)에 따르면, 대학생들이 주로 사용하는 어휘와 표현은 신어, 유행어, 은어, 비속어 등이 포함되어 있을 거라는 예상과는 달리 공적 상황에서는 발화 주제의 차이만을 보인다. 또한 사적인 상황에서는 20대는 친구 또래와 대학 생활과 관련

- ㄱ. 공적 20대: 명사파생접미사(화, 적, 성), 동사파생접미사(되, 시키), 선어말어미(었, 관형형 전성어미(라는, 다는), 명사형전성어미(ㄹ), 부정지정사(아니), 관형사(그런))  
 ㄴ. 공적 30대 이후: 명사파생접미사(님), 관형사(요, 고), 보조사(은)  
 ㄷ. 사적 20대: 명사파생접미사(님), 선어말어미(시, ㄱ ㅈ), 관형사(내), 보조사(ㄹ)  
 ㄹ. 사적 30대 이후: 관형사(그, 그런), 고유명사(한국, 일본), 목적격조사(을)

된 명사 ‘학기, 학교, 친구, 전공, 교수, 졸업, 대학, 수업, 아르바이트, 선생, 점수’ 등이 주를 이루는 반면, 30대 이후는 주거 생활과 관련된 명사 ‘신문, 짐, 신혼’과 음식과 관련된 명사 ‘김밥, 사과, 음식’이 특징적으로 나타났다.

대명사의 경우, 공적인 상황에 비해 사적인 상황에서 연령 관계없이 일인칭 대명사 ‘나, 내, 너, 자기’가 주로 나타난다는 점이 특징적이다. 일인칭 대명사는 사적 여성 핵심어에서도 추출된 바 있는데, 여성은 자신을 낮추어 지칭하는 일인칭 대명사 ‘저, 제’가 특징적으로 나타났다. 또한 20대는 어떤 대상을 지칭할 때 자신과 가까운 것보다는 멀리 있는 것을 지칭하는 ‘그거, 저거’가 더 두드러지는 반면에, 30대 이후는 공적인 상황에서 가까운 것을 지칭하는 ‘이것’의 변이형인 ‘요거, 요기’가 특징적으로 나타났다. 이는 공적 남성 핵심어에서도 추출되었는데, 임규홍(2004)에 따르면 말하는 사람에 중심으로 자신과 가까이 있는 대상을 지시하는 대명사의 사용은 더 내성적이고 자기중심적이라고 하는 일반적 성향과 서로 깊은 관련이 있다고 봤을 때, 30대 이후가 20대에 비해 내성적인 특징을 보인다고 볼 수 있다. 또한 공적 30대 이후는 듣는 이가 여러 사람일 때 그 사람들을 높여 이르는 이인칭 대명사 ‘여러분’이 핵심어로 추출되었다.

의존명사의 경우에는 숫자와 관련된 수사와 함께 30대 이후 발화에서 특징적으로 나타났는데, 남성 핵심어와 동일한 특징을 보인다. 특히 사적인 상황에서 30대 이후는 사람 이름을 부를 때 ‘씨’라는 호칭을 20대에 비해 더 많이 사용하는 것으로 나타났다.

용언의 경우에도 연령 변이 보다는 공적/사적 상황에 따라 의미상의 차이가 더 큰데, 그룹에도 불구하고 20대 보다는 30대 이후에 동사의 쓰임이 비교적 더 잘 드러난다. 먼저 동사의 경우, 공적인 상황에서 20대는 무언가 밖으로 드러나는 것을 의미하는 ‘보다, 나오다, 나타나다’ 동사가 특징적으로 나타났고, 30대 이후는 가장 쓰임이 일반적인 대표 주동사 ‘하다, 되다’와 공손표현인 ‘-어/아 주다’의 청자 높임인 ‘-어/아 드리다’의 쓰임이 주로 나타난다. 예를 들어, 다수의 청증을 위해 ‘알려 드리겠습니다, 얘기해 드리겠습니다, 공고를 해 드리겠습니다’ 등으로 나타난다. 특히 화자가 자신이 말하는 주제를 지칭할 때 사용하는 ‘관하다’라는 동사는 ‘정치외교에 관한, 청개구리에 관한, 우정에 관한’ 등의 쓰임으로 나타난다. 사적인 상황에서는 20대는 ‘듣다’, 30대 이후에는 ‘먹다, 뛰다, 짝다, 읽다’처럼 일상적인 동사가 핵심어로 추출되었다.

형용사는 공적 20대에서 ‘같다’의 쓰임이 두드러지는데 추측하여 자신의 의견을 단정하기 보다는 완화하여 표현하는 ‘-(으)ㄴ/는/(으)ㄹ 것 같다’ 문형에서 주로 나타났다. 이는 공적인 상황에서 20대가 자신의 의견에 대해 확신을 가지고 강하게 말하기 보다는 강도를 낮춤으로써 공손성을 확보하는 말하기 전략을 구사한다고 볼 수 있다. 사적 30대 이후에는 상태, 모양, 성질을 나타내는 ‘그렇다’의 쓰임이 특징적인데 긍정의 뜻으로

호응하며 대답할 때 쓰는 ‘그래’, ‘그렇구나’ 등과 이유와 담화의 진행을 나타내는 담화 표시 ‘그래 가지고’의 쓰임으로 주로 사용된다. ‘그렇다’는 공적 남성 핵심어로도 추출되었는데, ‘그렇죠, 그렇잖아요?’와 같이 상대방의 확인이나 동의를 구할 때 주로 사용되었다는 점에서 차이를 보인다.

보조동사의 쓰임도 흥미로운데, 공적인 상황에서 20대와 사적인 상황에서 30대 이후에서 ‘가지(갖)’의 쓰임이 눈에 띈다. 이유를 나타낼 때 사용하는 ‘-어/아 가지고(갖고)’의 문형으로 비격식적인 구어 발화의 특징인데, 20대는 공적인 상황에도 이런 표현이 이어진다는 점을 알 수 있다. 공적인 상황에서 30대 이후에는 동사와 마찬가지로 상대방을 높이는 보조동사 ‘계시’의 쓰임이 두드러지는데 ‘참가하고 계신 선생님, 앉아 계신, 일하고 계신, 맡고 계신, 오시고 계신’ 등의 예로 나타난다.

다음으로 수식어인 일반 부사를 보면, 공적인 상황에서 20대는 비격식 구어체의 특징인 강조부사 ‘되게’의 쓰임이 두드러지고, 30대 이후는 ‘아마, 요렇게’와 같이 자신의 태도를 나타내는 양태부사가 특징적으로 나타났다. 이와 반대로 사적인 상황에서는 20대가 ‘솔직히, 왜’와 같은 양태부사가 나타났고 30대 이후는 강조부사 ‘너무너무’의 쓰임이 두드러진다. 또한 사적인 상황에서 자신의 의견의 진위 여부를 드러낼 때 20대는 ‘솔직히’라는 감정적인 표현을 쓰는데, 30대 이후에는 보다 객관화 시켜 ‘사실’이라는 부사를 쓴다는 점도 흥미롭다. 접속부사의 쓰임에 있어서는 공적인 상황에서만 핵심어가 추출되었는데, 20대는 ‘근까, 그리구’와 같이 축약형과 음운 변화가 특징적인데, 30대 이후에는 ‘그래서, 또는’의 쓰임이 두드러진다.

감탄사의 경우에는 사적인 상황에 비해 공적인 상황에서 연령 관계없이 핵심어가 더 많이 추출되었는데, 주로 공적인 상황에서 말하기를 주저하고 시간을 끌 때 사용하는 예들이 나타났다. 특히 20대에는 긍정을 나타내는 ‘예, 예예’의 쓰임이 30대 이후에는 자신의 말하기를 이해하는지 확인할 때 사용하는 ‘그죠’의 쓰임이 눈에 띈다.

다음으로 기능어에 해당하는 의존형태인 어미와 관계어인 조사의 쓰임이다. 어미의 경우에는 종결어미에서 연령별 특징을 확인할 수 있다.<sup>15)</sup> 공적인 상황에서 20대는 해요체(거든요, 아요, 구요, 는데요)가, 30대 이후에는 합쇼체(ㅂ니다, 습니다)가 특징적으로 나타난다. 또한 사적인 상황에서 20대는 반말체인 해체(어, 야, 는데)가 나타나며 30대 이후는 해체(드라구, 여)뿐 아니라 해요체(어요)도 함께 나타난다. 특히 20대는 공적/사적 상황에 관계없이 ‘-는데(요)’의 쓰임이 두드러지는데, 이는 어떤 사실에 대해 알게 되거나 느끼게 되어 놀라움을 감탄하듯 말할 때 혹은 어떤 상황을 전달하면서 듣는 사람의 반응

15) 연결어미는 공적 20대 발화에서 대등의 의미를 나타내는 ‘구’가 핵심어로 추출되었다.

을 기대함을 나타낸다. 또한 사적인 상황에서 30대 이후가 특징적으로 사용한 ‘더(드)라구’ 표현은 과거에 직접 경험하여 새롭게 알게 된 사실을 상대방에게 전할 때 쓰는 것으로 사적 남성 핵심어로 추출된 바 있다.

마지막으로 조사의 분포를 살펴보면 사적인 상황에 비해 공적인 상황에서 핵심어가 더 많이 나타났는데 그 예는 다음과 같다. 공적인 상황에서 20대는 여러 대상을 예를 들거나 열거할 때 접속조사 ‘와, 랑, 이나’를 쓴 반면에 30대 이후는 ‘(이)라든가’를 선호하는 점이 특징적이다.<sup>16)</sup> 또한 주격조사의 경우에도 20대는 ‘이/가’, 30대 이후는 상대 높임 ‘께서’가 나타난 것을 보아 30대 이후가 더 격식적으로 말한다는 것을 유추할 수 있다. 또한 공적인 상황에서 20대가 보조사 ‘요’를 쓴다는 점이 흥미로운데, 예를 들어 ‘저희 삼촌에서는요, 보시면요, 이거는요’ 등과 같이 체언이나, 용언, 부사어 등에 ‘요’가 붙어서 청자에게 존대의 뜻을 나타냄을 알 수 있다.

이처럼 연령별 핵심어를 살펴본 결과, 성별 핵심어에 비해 연령 변인에 의한 차이보다 공적/사적 상황에 따른 차이가 더 많이 나타났다. 또한 성별 핵심어와는 반대로 사적인 상황에 비해 공적인 상황에서 비교적 20대와 30대 이후의 차이가 더 잘 드러났다. 공적인 상황에서 30대 이후는 지시 대명사, 의존명사, 동사, 일반부사, 접속부사, 감탄사, 종결어미, 조사의 쓰임이 특징적인데, 20대에 비해 더 격식적인 어휘들이 나타났다. 또한 연령별 핵심어 목록에는 성별 핵심어로 추출된 것과 유사한 어휘들이 나타났는데, 20대 핵심어는 여성 핵심어와, 30대 이후 핵심어는 남성 핵심어와 비슷한 특징이 발견되었다는 점이 흥미롭다. 예를 들어, 20대와 여성은 일인칭 대명사, 강조부사, 접속부사 ‘근까, 그리구’ 축약형과 음운변이, 종결어미 해요체의 특징이 비슷하게 나타났고, 30대 이후와 남성은 지시 대명사 ‘요거’, 의존명사와 수사 선호, 동사 선호, 형용사 ‘그렇’, 양태 부사, 종결어미 합쇼체, 조사 ‘라든가’의 쓰임이 유사하게 나타났다. 따라서 성별 핵심어와 연령별 핵심어는 서로 상관관계를 나타내는데, 20대와 여성은 비격식적(비형식적)인 표현이 30대 이후와 남성은 격식적(형식적)인 표현이 선호된다고 볼 수 있다.<sup>17)</sup>

16) ‘(이)라든가’는 앞서 공적 남성 핵심어로도 추출된 바 있다.

17) 이와 관련하여 익명의 심사자께서 20대와 30대 이후의 성별 비율에 의문을 제기하셨는데, 20대의 약 52%는 여성이고, 약 48%는 남성으로 나타났다. 또한 30대 이후의 약 47%는 여성이고 약 53%는 남성으로 20대는 여성이 약간 더 많고 30대 이후는 남성이 약간 더 많은 비율을 차지하는 것으로 나타났으나, 핵심어 결과에 큰 영향을 미칠 정도는 아니라고 판단하였다.

## IV. 핵심 어휘다발 분석

### 1. 성별 핵심 어휘다발

성별 핵심 어휘다발 추출 결과, 공적인 상황에서는 여성, 남성 모두 6개씩, 사적인 상황에서는 여성이 13개, 남성이 14개가 추출되었는데 다음은 정리하여 나타낸 표이다.<sup>18)</sup>

〈표 4〉 공적/사적 성별 핵심 어휘다발 목록

	여성 핵심 어휘다발	남성 핵심 어휘다발
공적	1) 을 알 르 수 2) 알 르 수 있 3) 르 수 두 있	1) 2) 하 ㄱ 가지 고 <sup>19)</sup>
	4) 한번 하 ㄱ 보	3) 하 는 것 이 4) 르 것 이 나
	5) 이 ㄴ 거 이	5) 으 로 되 ㄴ 있
	6) 예 를 들 면	6) 그 리 하 ㄴ 부분
사적	1) 음음음음, 2) 음음음어, 3) 어어어어, 4) 음음어어 5) 음어음음	1) 저 뭐 이 야, 2) 저저 뭐 이 야, 3) ㄴ 뭐 이 나 면 은
	6) 그렇 ㄱ 갖 구, 7) 하 ㄱ 갖 구, 8) 그렇 ㄱ 갖 고, 9) 그렇 ㄱ 가지 구	4) 같은 경우에, 5) 은 경우에는, 6) 나같은 경우
	10) 는 거 이 예요 11) ㄴ 거 이 예요	7) ㄴ 거 이지, 8) 거 이지 근데
	12) 생각 이 안 나	9) 내 가 보 르, 10) 내 가 보 ㄴ ㅅ 11) 가 보 ㄴ ㅅ 을
	13) 사람 들 이 다	12) 하 ㄱ 가지 구
		13) 일 이 있 었
		14) 개 네 들 이

〈표 4〉에 따르면 공적/사적 성별 핵심 어휘다발은 공통적으로 1) -어/아 가지고, 2) -(으)ㄴ/는 것이다 3) 담화 표지 4) 사람들이 / 개네들이 5) 예시 표현 6) 판단 표현이 특징적으로 나타난다.

18) 본 연구에서 핵심 어휘다발은 4개의 형태소 연속체로, 예를 들어 공적 여성 핵심 어휘다발 목록은 1위 을/JKO 알/VV 르/ETM 수/NNB, 2위 르/ETM 수/NNB 두/JX 있/VA, 6위 알/VV 르/ETM 수/NNB 있/VA 가 추출되었는데, 보다 효율적인 분석을 위해 한 눈에 보기 좋게 품사표지는 제외하고 관련 어구끼리 정리하여 보인 것이다.

19) 공적 남성 핵심 어휘다발 목록에 '하 ㄱ 가지 고'가 두 개 추출되었는데, 하나는 일반명사에 동사파생접미사 '하다'가 결합된 '하/XSV ㄱ/EC 가지/VX 고/EC'와 다른 하나는 부사의 수식을 받는 '집중/NNG 적/XSN 으 로/JKB 하/VV ㄱ/EC 가지/VX 고/EC'이다.

첫째, 이유와 시간의 흐름을 나타내는 구어 표현인 ‘-어/아 가지고’를 보면, 공적 여성인 경우만 제외하고 공적 남성, 사적 여성, 사적 남성에서 모두 나타난다. 그런데 남성은 ‘하다’ 동사와 결합하여 ‘해 가지고(구)’라는 표현을, 여성은 ‘그렇다’ 형용사와 결합하여 ‘그래 가지고(갖고, 갖구)’ 표현을 선호한다는 점이 특징적이다. 앞서 핵심어 목록에서도 보조용언 ‘가지다’가 공적 여성에서는 안 나타났고, 사적 여성에서 ‘갖’, 공적/사적 남성에서 ‘가지’로 추출된 것과도 관련이 깊다.

- (7) ㄱ. 프로젝트를 신청해 가지고 / 집중적으로 해 가지고 (공적 남성)
  - ㄴ. 상대가 천안이잖아. 그래 갖구, 이젠, 그래 갖구 친해졌다. 그래 가지고 미팅을 하자 그래, 그래갖구 만났어. (사적 여성)
  - ㄷ. 레슬링 해 가지고 체격이 끝내줬거든요. (사적 남성)

(7)의 예를 보면, 공적/사적 상황에 관계없이 남성은 ‘해 가지고(구)’를 선호하는데 시간상의 흐름, 이유의 의미를 나타낸다. 한편, 사적인 상황에서 여성이 선호하는 ‘그래 가지고(갖고, 갖구)’는 담화 표지의 기능을 지니고 있다. 임규홍(2004: 104)에서도 ‘그래 가지고’는 여성이 선호하는 담화 표지로, 어원적으로 선행 정보와 후행 정보를 이어주는 구실을 하는 ‘그래서’가 어휘적 의미를 상실하고, 담화 진행을 이끄는 구실을 한다고 주장한 바 있다.

둘째, 어떤 사실이나 사물에 대한 말하는 사람의 생각을 객관화하여 강조할 때 쓰이는 ‘-(으)ㄴ/는 것이다’를 살펴보겠다. 공적 여성은 ‘이 ㄴ 거 이’, 공적 남성 ‘하 는 것 이, ㄴ 것 이 냐’, 사적 여성 ‘는 거 이 예요, ㄴ 거 이 예요’, 사적 남성 ‘ㄴ 거 이지, 거이지 근데’의 표현이 특징적으로 나타나났는데, 그 예는 다음과 같다.

- (8) ㄱ. 실제적인 응용인 거예요. / 문법의 차이인 거예요. (공적 여성)
  - ㄴ. 아스키라고 하는 것인데 / 우리가 어느 것을 선택할 것이냐가 (공적 남성)
  - ㄷ. 그 물을 마시는 거예요. / 너무 심한 거예요. (사적 여성)
  - ㄹ. 내 생각은 그렇다는 거지. 근데, 니 말도 일리가 있네. (사적 남성)

(8ㄱ)의 공적인 상황에서 여성은 어떤 사실을 설명할 때 ‘-(으)ㄴ 것이다’와 해요체가 결합하여 주로 사용하는데, 의존명사 ‘것’이 발음할 때 ‘ㅅ’ 탈락을 보인다. 반면에 (8ㄴ)의 공적인 상황에서 남성은 ‘-라고 하다’라는 다른 사람의 말을 전달하거나 가설이나 연구 결과 등을 제시할 때 주로 사용되고 음운변이는 나타나지 않는다. 사적인 상황에서는 (8ㄷ)의 여성은 일상생활의 기본 어휘인 동사, 형용사와 결합하여 말하는 화자가 자신의

감정이나 생각을 강조하며 나타낼 때 주로 쓰인다. 또한 여성은 ‘거예요’라는 표현이 특징적인데, (8ㄷ)의 남성은 ‘거이지’라는 표현으로 보다 자신의 의견에 대해 확신을 가지고 표현한다고 볼 수 있다. 이와 관련하여 사적 여성종결어미 핵심어로 ‘어요’, 사적 남성은 ‘지’가 추출된 것과도 상응하는 결과이다.

셋째, 사적인 상황에서 남성과 여성은 주저하거나 말더듬는 담화 표지가 특징적으로 나타나는데 여성은 ‘음음음음, 음음음어, 어어어어, 음음어어, 음어음음’으로, 남성은 ‘저 뭐야, 저거 뭐야 등’이다.

- (9) ㄱ. 이름이 생각이 안 나네, 어어 이렇게 어어어어어 맞아 맞아 맞아. (사적 여성)  
 ㄴ. 미치겠다 지... 저 뭐야 너 이름 뭐더라? (사적 남성)

이 표현은 말하기 도중에 뒤에 나오는 정보를 빨리 떠올리지 못할 때 시간을 버는 기능을 하는 것으로, (9ㄱ)의 여성은 ‘어어어어’로 나타나고, 핵심 어휘다발 중 ‘생각이 안 나’와 함께 쓰이면서 말하기 도중 잊어버린 것에 대해 걸으로 드러냄으로써 자신감이 결여 되어 있음을 보여 준다. 반면에 (9ㄴ)의 남성은 ‘저 뭐야’처럼 질문형으로 상대에게 먼저 물으면서 자신이 말할 내용을 생각하기 위한 시간을 버는 전략을 사용한다. 이는 담화 기능을 나타내는 감탄사 ‘어’는 사적 여성 핵심어이고, 대명사 저거, 뭐’는 사적 남성 핵심어로 추출된 것과 관련이 깊다.

넷째, 사적인 상황에서 여러 사람을 지칭하는 표현으로 여성 핵심어휘다발인 ‘사람들이 다’와 남성 핵심 어휘다발인 ‘개네들’을 비교할 수 있다.

- (10) ㄱ. 왜 사람들이 다 뭐라 그랬어, 사람들이 다 그래요. (사적 여성)  
 ㄴ. 그러니까 자기네 개 개네들이 왜 학생회장을 뽑냐면은... (사적 남성)

(10ㄱ)의 여성의 발화에서 여러 사람을 지칭한 ‘사람들이’라는 표현을 사용함으로써 다른 사람들의 시선과 생각을 중시 여긴다는 점이 드러난다. 한편 (10ㄴ)의 ‘개네들’은 화자와 청자가 아닌 3인칭 ‘그 아이’의 줄임말인 ‘개’에 무리의 뜻을 나타내는 ‘네’와 복수형 ‘들’이 붙은 표현으로 남성들이 자신과 동등하거나 어린 다른 사람에 대해 약간 부정적인 이야기를 할 때 주로 사용한다.

다섯 번째로, 공적/사적인 상황에서 성별에 따라 예를 들며 말할 때 차이를 보이는데 공적 여성은 ‘예를 들면’, 공적 남성은 ‘그러한 부분’, 사적 남성은 ‘같은 경우에’로 표현된다.

- (11) 가. 예를 들면 이 작품 여기 보면은... (공적 여성)
- 나. 그러한 부분에 대한 문제들이 있습니다. (공적 남성)
- 다. 음 이스라엘 같은 경우에는 여자가 이 년 갔다와야 되고 군대를. (사적 남성)

(11가)의 여성은 공적인 상황에서 설명하며 말할 때 예시를 들 경우 ‘예를 들면’이라고 명시적으로 드러낸다. 한편 남성의 경우 (11나)의 공적인 상황에서는 ‘그러한 부분’이라는 표현을 통해 앞서 설명한 예들을 지칭할 때 사용한다. (11다)의 사적인 상황에서는 구체적인 여러 사례 중 하나를 설명할 때 ‘같은 경우’를 사용한다.

마지막으로 발화자의 판단을 표현할 때 공적 여성은 ‘알 수 있다’, 사적 남성은 ‘내가 보 르, 내가 보 ㅏㅓ, 가 보 ㅏㅓ을’을 주로 사용하여 어떤 사실에 대한 자신의 의견이나 입장, 생각을 나타낸다.

- (12) 가. 서로 존칭을 사용하는 것을 알 수 있습니다. (공적 여성)
- 나. 내가 봤을 때는 주변에서 그런 걸 막 조장하는 거 같애. (사적 남성)

(12가)은 어떤 사실에 대해 화자가 판단한 확실성 또는 가능성의 정도를 나타내는 표현으로, 강연, 강의와 같은 공적인 상황에서 어떤 사실을 근거하여 발화자의 판단을 나타낸다. 반면에 (12나)은 사적인 상황에서 자신의 주관적인 의견 및 판단을 나타낼 때 사용한다. 이처럼 여성은 어떤 사실에 근거를 두고 판단하며 자신의 의견을 나타내고 남성은 자신의 주관적인 판단을 제시하는 것으로 보이는데 성별의 차이일 수도 있지만 공적/사적 상황적인 차이에서 기인하는 것일 가능성도 배제할 수 없다.

이처럼 성별에 따른 핵심 어휘다발을 비교 분석한 결과, ‘-어/아 가지고’, ‘-(으)ㄴ/는 것이다’처럼 구어에서 주로 쓰이는 표현이 주로 추출되어 핵심어에 비해 성별에 따라 선호하는 어구가 비교적 적게 나타났다. 그러나 동일 어구 표현이 성별에 따라 다르게 사용되는 것을 확인할 수 있고, 핵심 어휘다발을 이루고 있는 어휘 요소 중에는 핵심어도 포함되어 서로 관련이 있음을 알 수 있다.

## 2. 연령별 핵심 어휘다발

공적/사적 사용역별 연령별 핵심어휘다발의 목록이다. 공적인 상황에서는 20대와 30대 이후 모두 9개, 사적인 상황에서는 20대가 2개, 30대 이후 5개가 추출되었다.

〈표 5〉 공적/사적 연령별 핵심 어휘다발 목록

	20대 핵심 어휘다발	30대 이후 핵심 어휘다발
공 적	1) 알 르 수 있 2) 을 알 르 수 3) 르 수 있 었 4) 르 수 있 다는 5) 보 르 수 있	1) 이 라고 하 는 2) 라고 하 는 것 3) 다 라고 하 는 4) 라고 하 는 거 5) 하 는 것 은 6) 하 는 것 이
	6) 은 경우 에 는 7) 같 은 경우 는 8) 얘 기 르 하 고 9) 르 하 고 있	7) 하 ㄱ 가지 고 8) 하 ㄱ 주 시 9) 되 ㄴ다는 거 이
사 적	1) 되 르 거 같	1) 음 음 음 음 2) 어 어 어 어 3) 어 어 음 음
	2) 단 말 이 야	4) 는 거 이 야 5) 그렇 ㄱ 갖 구

〈표 5〉를 보면, 연령별 핵심 어휘다발 목록은 앞서 성별 핵심 어휘다발 목록에서 추출된 것이 일부 포함되어 있는데, 이를 비교하면서 살펴보도록 하겠다.

먼저 공적 20대 핵심 어휘다발 중 가장 많이 나타난 ‘알 수 있, 을 알 수’ 표현은 앞서 공적 여성 핵심 어휘다발 중 어떤 사실에 대한 화자의 판단을 나타내는 ‘알 수 있’과 쓰임이 동일하다. 그런데 공적 20대 핵심 어휘다발 목록에는 ‘알다’ 동사 외에도 ‘보다’와 다른 다양한 동사와 ‘-(으)르 수 있다’ 표현이 결합되어 나타난다는 점이 특징적이다.

- (13) 가. 여성이 보다 남성에 비해서 감성적임을 알 수 있었고.. (공적 20대)  
 나. 저희가 이제 발견할 수 있었구요. (공적 20대)  
 다. 그런 이미지를 상품과 연결시키고 있다는 것을 볼 수 있었습니다. (공적 20대)

(13)을 보면, 주로 발표 상황에서 발화자가 추측하고 발견하고 알게 된 사실을 판단하며 이야기할 때 ‘-(으)르 수 있다’ 표현이 사용되고 있다.

또한 공적 20대 핵심 어휘다발로 추출된 ‘은 경우에는, 같은 경우는’의 표현은 사적 남성 핵심 어휘다발에서도 추출된 것으로 여러 사례를 예를 들어 설명하는 기능을 하고 있다. 따라서 이 표현은 공적/사적 상황에 영향을 받지 않고 20대 남성이 선호하는 것으로 볼 수 있다.

공적 20대 핵심 어휘다발 중 ‘얘길하고, 하고 있’은 공적인 상황에서 발표를 할 때, 책이나 다른 사람의 글이나 말을 인용할 때 사용되고 있다.

- (14) ㄱ. 여기서는 (…중략…) 외국어 연구 사이의 알력들에 대해서 애길 하고 있습니다. (공적 20대)  
 ㄴ. 저자가 말하고자 하는 것은 (…중략…) 기독교 근본주의와 세계 교회주의에 대해서 이렇게 애길하고 있는데… (공적 20대)

(14ㄱ)의 ‘애길 하고’의 주어를 보면 ‘여기서는’으로 책 내용의 일부를 지칭하고 있고 (14ㄴ)에서는 ‘저자가 말하고자 하는 것은’으로 저자의 입장을 나타낼 때 사용되고 있다. 이와 같이 공적인 상황에서 인용을 할 때 사용되는 표현으로 공적 30대 이후에서는 ‘이라고 하는, 라고하는 것, 다라고하는, 라고하는 것’의 핵심 어휘다발이 특징적으로 나타났다.

- (15) ㄱ. 한국어 교육이라고 하는 특수 분야에 대한 이해를 배양하고… (공적 30대 이후)  
 ㄴ. 말뭉치라고 하는 것은 여기서 말하는 것은… (공적 30대 이후)  
 ㄷ. 말을 분석한다라고 하는 것은 결국 뭐냐면… (공적 30대 이후)

‘다고/라고’는 주로 어떤 말이나 문장을 그대로 직접 인용할 때 쓰는 표현으로 발표 상황에서 많이 사용하는데, (15)의 예를 보면, 대부분 전문 용어에 대해 풀어서 설명할 때 주로 사용된다.

다음으로 공적 30대 이후 핵심 어휘다발 중 ‘해 가지고’는 공적 남성 핵심 어휘다발 목록 중 ‘해 가지고’와 동일한 것으로 보아, 이 표현은 30대 이후 남성이 공적인 상황에서 주로 사용한 것으로 판단된다.

그 밖에도 ‘-어/아 주다’ 표현인 ‘해 주시’는 다음과 같은 쓰임을 보인다.

- (16) ㄱ. 선생님께서 발표를 해 주시겠습니다. (공적 30대 이후)  
 ㄴ. 질의 응답에 응해 주시면 고맙겠습니다. (공적 30대 이후)  
 ㄷ. 제목으로 발표를 해 주시겠습니다. (공적 30대 이후)  
 ㄹ. 간략하게 말씀을 해 주시고 어 다음 순서를 진행하도록 하겠습니다. (공적 30대 이후)

(16)을 보면, 공적 상황 중 워크샵에서 전형적으로 쓰이는 표현들이 나타남을 알 수 있다. ‘-어/아 주다’라는 공손 표현에 주체 높임인 ‘시’와 함께 결합하여 공적인 상황에서 활용되고 있다.

사적인 상황을 살펴보면 가장 특징적인 것이 30대 이후 핵심 어휘다발은 대부분 ‘음음음음, 어어어어, 어어음음’의 주저하기 담화표지들이 주를 이룬다는 점이다. 이는 앞서 사적 여성 핵심 어휘다발과 동일한 표현으로 30대 이후 여성들이 사용한 것으로 보인다. 이와 관련하여 사적 30대 이후 핵심 어휘다발인 ‘는 거야, 그래 갖구’ 역시 사적 여성 핵심 어휘다발 중 일부이다. 따라서 사적인 상황의 30대 이후는 대부분 여성일 거라는 추측이 가능하다.

다음으로 사적 20대 핵심 어휘다발을 살펴보면 자신의 생각을 표현할 때 ‘될 거 같아’라는 표현이 특징적으로 나타났는데, 그 예는 다음과 같다.

- (17) 가. 정리가 되지 못하고 끝내야 될 거 같고 (사적 20대)  
 나. 어떻게 좀 잘 이렇게 넘어가면 될 거 같은데 (사적 20대)  
 다. 여름 방학 때 해도 될 거 같애. (사적 20대)

(17)에서 사용된 ‘(으)ㄹ 거 같다’ 표현은 일반적으로 추측하며 말하는 의미로 쓰이지만, 상대방에게 말하는 사람 자신의 생각이나 의견을 말할 때 강하게 단정하기 보다는 좀 더 부드럽고 겸손하게 말할 때 쓴다는 점에서 특징적이다. 이는 앞서 공적 20대 핵심어 중 형용사 ‘같다’가 추출된 것과 유사한 쓰임으로, 20대가 선호하는 표현으로 볼 수 있다. 마지막으로 사적 20대 핵심 어휘다발 중 ‘단 말이야’라는 표현이 있다.

- (18) 가. 따뜻한데서 먹으면 별로 맛도 없단 말이야. (사적 20대)  
 나. 보충이 있었던 말이야. (사적 20대)  
 다. 나같은 애들 뿐이 없단 말이야. (사적 20대)

주의 집중을 이끄는 담화표지 ‘말이야’는 주로 비격식 반말체와 함께 결합하여 구어적인 특징을 잘 드러내는데, 상대방의 주의를 끌거나 말을 강조하며 다짐하는 뜻을 나타낸다. 이 표현은 사적인 상황에서 20대가 주로 사용하는 것으로 나타났다.

이처럼 연령별 핵심 어휘다발을 살펴본 결과, 성별 핵심 어휘다발로 추출된 것과 유사한 표현들이 주로 나타났는데, 핵심어와 마찬가지로 성별 핵심 어휘다발에 비해 연령별 핵심 어휘다발의 특징이 덜 나타났다. 또한 성별 변인에 비해, 연령별 핵심어와의 상관관계도 크게 드러나지 않았다. 그럼에도 불구하고 공적 30대 이후와 사적 20대의 핵심 어휘다발은 핵심어 분석을 통해서 알 수 없는 어구 표현의 차이를 살펴 볼 수 있다.

## V. 맺음말

본 연구에서는 세종 구어 말뭉치를 활용하여 성별과 연령별 사회적 변인에 따른 발화 차이를 분석하기 위해 위드스미스의 핵심어 기능을 이용하여 핵심어, 핵심어휘다발을 통계적으로 추출하여 살펴보았다. 그 결과 요약하면 다음과 같다.

첫째, 공적/사적 상황에서 성별 핵심어가 연령별 핵심어에 비해 더 특징적으로 나타났는데, 특히 성별 핵심어는 사적인 상황에서 더 잘 드러나는 반면에, 연령별 핵심어는 공적인 상황에서 비교적 잘 나타났다. 즉, 남성과 여성의 발화 차이는 사적인 상황에서 더 잘 드러나고, 20대와 30대 이후의 발화차이는 공적인 상황에서 더 잘 나타났다.

둘째, 성별, 연령별 핵심 어휘다발을 살펴 본 결과, 남성과 여성, 20대와 30대 이후에서 ‘-어/아 가지고’, ‘-(으)ㄴ/는 것이다’, ‘-(으)ㄴ 수 있다’ 등의 어구가 주로 추출되었는데, 동일 어구라도 변인에 따라서 다르게 표현되고 사용됨을 알 수 있었다.

셋째, 성별 핵심어와 연령별 핵심어는 서로 상관관계가 나타나는데, 20대와 여성은 비격식적(비형식적) 표현을 30대 이후와 남성은 격식적(형식적) 표현을 선호하는 것으로 나타났다.

넷째, 성별, 연령별 핵심 어휘다발 분석은 핵심어 분석에 비해 변인에 따라 유의미한 차이가 크게 나타나지 않았지만 어구 차원의 성별, 연령별 발화 차이를 어느 정도 살펴 볼 수 있었다. 또한 핵심어로 추출된 어휘가 핵심 어휘다발의 요소로도 나타나는 것으로 보아, 핵심어와 핵심 어휘다발간의 상관관계가 있음을 알 수 있다.

따라서 본 연구에서는 공적/사적 성별, 연령별 발화의 특징을 핵심 어휘 차원뿐 아니라, 품사까지 고려한 통사 차원, 그리고 핵심 어휘 다발을 통해 어구의 화용적인 차원까지 종합적 분석을 수행함으로써 이러한 분석 방법이 사회언어학적 연구에 기여할 수 있는 가능성을 제시하였다는 점에 의의가 있다. 그러나 개별 핵심어와 핵심 어휘다발 자체로는 전체 문맥에서 사용된 의미나 특징을 명확히 알 수 없으므로 담화 차원에서 용례를 분석할 필요성이 있다는 점과 성별, 연령별로 다르게 선호하는 어휘와 표현에 대한 사회언어학적 해석에 대한 논의는 추후에 진행되어야 할 것이다.

## 참고문헌

- 강범모 (2013). “남성과 여성 발화의 어휘적 차이: 코퍼스 기반의 남녀 판별 연구.”  
한국어학 58, 1-30. 한국어학회.

- 권순구 (2007). “언어 표현과 인식에 있어서의 남녀 차이: 보조용언의 사용을 중심으로.” 인문학연구 34(3), 7-30. 충남대학교인문과학연구소.
- 김혜숙 (2002). “대학생 구어 활동의 모습과 언어학적 분석 방법 시안.” 동국어문학 14, 5-28. 동국대학교.
- 김혜영 (2009). 구어에서 나타나는 정도부사의 사용의미. 고려대학교 석사학위 논문.
- 김혜영 (2011). “남성과 여성의 사적 대화에서 발화 특성 연구: 담화 기능 분석을 중심으로.” 언어와 언어학 53, 89-108. 한국의국어대학교 언어연구소.
- 문금현 (2000). “대학생의 화법에 나타난 사회언어학적 특성.” 새국어교육 59, 223-250. 한국국어교육학회.
- 민현식 (1997). “국어 남녀 언어의 사회언어학적 특성 연구.” 사회언어학 5(2), 529-587. 한국사회언어학회.
- 서상규·구현정 (공편). (2005). 한국어 구어 연구 (2): 대학생 대화 말뭉치를 중심으로. 한국문화사.
- 서상규 외 (2013). 한국어 구어 말뭉치 연구. 한국문화사.
- 임규홍 (2004). “성별에 따른 국어 담화 표지 사용 모습: 중, 고등학생 이야기 대화를 바탕으로.” 어문학 83(3), 93-113. 한국어문학회.
- 임철성 (1997). “연령층별 성별 어휘 사용의 계량적 고찰: 방송국 ‘일출’의 대사를 대상으로.” 새국어교육 54, 253-276. 한국국어교육학회.
- 장경희·김순자 (2008). “연령과 성별 요인에 따른 맞장구 수행 실태 조사.” 텍스트언어학 25, 219-247. 한국텍스트언어학회.
- 장세은·이성민 (2012). “키클러스터 분석을 통한 학술목적영어 코퍼스의 어휘다발 특성 연구.” 언어학 20(4), 219-239. 대한언어학회.
- 장영희 (2000). “20대 남녀 사용 어휘의 대비적 고찰, 국어화법과 대화분석.” 화법연구 2, 93-114. 국어화법학회.
- 전지은 (2010). “성별에 따른 한국어 부사 사용 양상: 세종 구어 말뭉치를 활용하여.” 언어와 언어학 47, 191-218. 한국외국어대학교 언어연구소.
- 전지은 (2011). 핵심어 분석을 통한 영어 성별 어휘 사용 양상 연구. 고려대학교 박사학위 논문.
- 전혜영 (2006). “언어 사용자의 성별과 발화 특성.” 한국어학 31, 47-70. 한국어학회.
- Biber, D., S. Conrad and V. Cortes. (2003). Lexical bundles in speech and writing: An initial taxonomy. In A. Wilson, P. Rayson, and T. McEnery (eds.), *Corpus Linguistics by the Lune*, 71-93. Frankfurt, Main: Peter Lang.

- Baker, P. (2006). *Using Corpora in Discourse Analysis*. London: Continuum.
- Baker, P. (2010). *Sociolinguistics and Corpus Linguistics*. Edinburgh Sociolinguistics.
- McEnery, T., Xiao, R., and Tono, Y. (2006). *Corpus-Based Language Studies: An advanced resource book*. London and New York: Routledge.
- Rayson et al. (1997). Social Differentiation in the use of English vocabulary: some analyses of the conversational component of the British National Corpus. *International Journal of Corpus Linguistics* 2(1), 133-152.
- Sardinha, T. B., and Schimazumi. (2003). Schoolchildren writing: A corpus-based analysis. *Linguagem & Ensino* 6(1), 11-33.
- Schmid, H. J. (2003). Do women and men really live in different cultures? Evidence from the BNC. In A. Wilson, R. Rayson and T. McEnery (eds.), *Corpus Linguistics by the Lune*, 185-221. Frankfurt Peter Lang.
- Scott, M. (1999). WordSmith Tools Users Help File. Oxford: OUP.
- Scott, M. (2012). WordSmith tools help. Liverpool: Lexical Analysis Software.
- Scott, M., and Tribble, C. (2006). *Textual Patterns: Key words and corpus analysis in language education*. John Benjamins.

전지은

jejeon@yonsei.ac.kr

논문 접수일: 2월 15일

논문 심사일: 3월 1일~3월 31일

논문 확정일: 4월 10일

