COMMUNICATION SCIENCES & DISORDERS

ISSN 2288-1328 (Print)
ISSN 2288-0917 (Online)

Commun Sci & Dis 2014;19(3):274-284

Original Article

http://dx.doi.org/10.12963/csd.14168

Predicates in the Vocabulary of Typically Developing Preschoolers

Jea-Hyuk Oh[a], Jae-Eun Cha[b], Mi-Sun Yoon[c], Jung-Mee Kim[c], Moon-Soo Chang[d]

[a]*The Institute of Foreign Language Studies, Korea University, Seoul, Korea*
[b]*Department of Korean Language & Literature, Kyonggi University, Seoul, Korea*
[c]*Department of Communication Disorders, Korea Nazarene University, Cheonan, Korea*
[d]*Department of Computer Science, Seokyeong University, Seoul, Korea*

Correspondence: Mi-Sun Yoon, PhD
Department of Communication Disorders, Korea Nazarene University, 48 Wolbong-ro, Seobuk-gu, Cheonan 330-718, Korea
Tel: +82-41-570-1412
Fax: +82-41-570-7846
E-mail: msyoon@kornu.ac.kr

Received: July 1, 2014
Revised: August 5, 2014
Accepted: August 29, 2014

This work was supported by the National Research Foundation (NRF) of Korea (No. NRF-2011- B00202).

Objectives: The purpose of this study is to evaluate the types and tokens of predicates in the spoken language of Korean preschoolers. Predicates include verbs, adjectives, auxiliary verbs, and auxiliary adjectives. **Methods:** 300-400 utterances from 80 two- to five-year-old children, 20 children from each age group, were selected from the Nazarene spoken language corpus. Types and tokens of predicates were measured from a total of 27,485 utterances (65,783 eojeol). Differences in each word class of predicates, based on age groups, were evaluated and then arranged by frequency. **Results:** First, the ratio of types and tokens of predicates to total vocabulary was 17.4% and 20.1%, respectively. In predicates, the type-token verb ratio was 73.7%-60.2%, and that of adjectives was 23.4%-25.8%, respectively. Second, there were only significant differences in the types and tokens of predicates between the two-year-old group and the other groups. Third, only a limited number of types of verbs and adjectives were used from high frequencies in the word list. **Conclusion:** The ratio of predicates to total vocabulary was similar in each age group. The development of predicates in two- to three-years-old is significant. Preschoolers used several predicates repetitively. These results could be used in the evaluation and intervention of children with language difficulties.

Keywords: Speech corpus, Predicate, Verb, Adjective, Type-token, Vocabulary

용언은 체언과 더불어 문장의 필수적인 성분을 구성하는 어휘 부류이다. 체언이 주로 문장의 뼈대를 구성하는 주어, 목적어, 보어의 역할을 하는 어휘라면, 용언은 서술어로 쓰여 이러한 성분들을 지배하는 핵의 역할을 하는 어휘이다. 또한 체언이 지시물의 명칭이나 개념을 표현하는 어휘라면, 용언은 이러한 명칭이나 개념에 대한 화자의 의사를 표현하는 어휘이다. 국어에서 용언을 구성하는 품사는 동사, 형용사, 보조 동사, 보조 형용사이다. 동사와 형용사는 동작성과 상태성에 따른 구분이며, 묶어서 본용언이라고 한다. 본용언과 대비하여 보조 용언(보조 동사, 보조 형용사)은 문장(발화) 내에서 주로 양태의 의미를 더해주는 기능을 한다. 또한 국어의 품사 체계 내에서 용언은 어미를 취하여 활용하는 부류인 변화사이다(Kim, 2012; Ko & Ku, 2008).

일반 아동은 초기 어휘 습득 과정에서 대표적 체언인 명사보다 대표적 용언인 동사를 늦게 습득하며 이는 언어보편적인 현상으로 알려져 있다(Bassano, 2000; Caselli et al., 1995; Golinkoff, & Hirsh-Pasek, 2008; Kauschke, Lee, & Pae, 2007; Lee, 2004; McDonough, Song, Hirsh-Pasek, Golinkoff, & Lannon, 2011). 이러한 현상은 언어에 장애가 있는 아동들에게서 더욱 뚜렷이 나타난다. 언어 장애 아동들은 동사의 사용에서 명사보다 상대적으로 많은 어려움을 보인다. Lee와 Kim (2003)의 연구에서 언어 연령이 4;6-6;6 사이의 단순언어장애 아동은 같은 언어 연령의 일반 아동과 비교했을 때 명사보다 동사의 산출에서 특징적으로 더 많은 오류를 보였다. 자폐 아동의 경우에도 명사보다 동사의 이름대기 검사에서 수행력이 더 낮았다(Kwon & Pae, 2005). 이와 같은 결과는 청각장애 아동에서

Copyright © 2014 Korean Academy of Speech-Language Pathology and Audiology
This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0) which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

도 관찰된다. Park과 Pae (2003)에 따르면 70 dB 이상의 청력손실이 있는 청각장애 아동의 동사 범주 어휘의 습득률은 명사 범주 어휘의 습득률에 비해 낮았다. 이러한 결과는 어휘 습득 과정에서 용언의 발달이 체언보다 어려움을 시사한다.

언어 발달에서 용언이 갖는 중요성은 단순한 어휘 습득의 차원을 넘어 아동 언어의 구문론 및 형태론적인 발달과도 연결이 된다는 점이다. 아동 언어 발달 평가의 척도인 다양한 문법 형태소의 습득과 사용 또한 용언의 습득을 전제로 한다. 학령전기 아동 언어의 구문론적인 발달을 평가하는 지표로 사용되는 평균발화길이(mean length of utterance, MLU)는 용언의 습득과 높은 상관관계를 보이고 있다(Lee, Park, Park, & Kim, 2003). Lee (2003) 등은 4세 미만의 두, 세 낱말 발화기 아동 35명을 대상으로 자발화를 분석하여 용언과 평균발화길이의 관계를 보았다. 총 용언 수는 MLU-w와 .736, MLU-m과 .818의 높은 상관을 보였고, 용언 유형 수 또한 MLU-w와 .659, MLU-m과는 .711의 높은 상관을 보였다. 특히 이 연구에서는 생활 연령이 같고 MLU가 차이가 나는 집단에서 용언의 습득에 유의한 차이가 있음을 보여주며, 용언 어휘 습득이 구문 발달 단계와 더욱 관련이 있을 것으로 추정했다.

따라서 학령 전기 아동에게서 나타나는 용언의 어휘 빈도 및 그 사용 양상을 살피는 것은 아동의 의사 표현 발달 양상과 언어 발달 양상을 이해하고 언어 장애 아동의 평가와 치료를 위해 중요한 과제라 할 수 있다.

아동의 언어 발달 지표로 활용할 수 있는 어휘 빈도 정보를 제공하기 위해서는 몇 가지 고려해야 할 점들이 있다. 첫째, 아동의 어휘 빈도 정보는 균형성과 대표성이 있는 균형 말뭉치를 기반으로 도출해야 한다. 균형 말뭉치를 구축하기 위해서는 특정 과제나 특정 집단을 기반으로 발화를 수집하고 말뭉치를 구축하는 것은 지양해야 한다. 또한 말뭉치의 대표성을 담보하기 위해서는 적정 수 이상 아동의 자발화를 수집해야 한다. 둘째, 빈도로 제시할 어휘 분류 기준을 명확하게 설정하고 이에 따라 구체적인 어휘 목록을 제시할 필요가 있다. 국어학적 관점에서 보편 타당한 어휘 분류 기준을 따라야 하며, 가능한 한 구체적으로 어휘 목록을 제시해야 한다. 셋째, 품사별 빈도 정보 이상의 다양한 어휘 정보를 제시할 필요가 있다. 단순히 동사나 형용사의 유형 수와 사용 빈도를 객관적으로 제시하는 것을 넘어, 고빈도 어휘군이 지니고 있는 특성에 대한 다양한 어휘 정보가 제공되어야 한다. 영어권에서 사용하고 있는 MacWhinney (2000)의 CHILDES 프로젝트, Miller (2003)의 Systematic Analysis of Language Transcripts (SALT)는 정보의 균형성과 대표성을 갖춘 대표적인 말뭉치 자료이다. 국어에서는 '21세기 세종 말뭉치', '물결 21'과 같은 대규모 말뭉치가 있으나, 성인 언어, 특히 구어보다는 문어 중심으로 어휘 정보를 제공하고 있다. 학령 전기 아동의 어휘 정보에 대한 연구도 진행되었으나, 말뭉치로부터 어휘를 추출하는 방식으로 보기는 어렵다. Pae, Chang, Kwak, Sung과 Sim (2004)은 1,138명의 유아를 대상으로 어휘 발달을 연구하였다. 전국적으로 많은 아동을 대상으로 실시된 방대한 어휘 연구로 대표성을 충분히 확보한 연구이나, 자료 수집은 제시된 어휘 목록을 보고 부모가 표시하는 방식이었다. 이렇게 아동이 사용하는 어휘를 부모가 표시하는 형식으로 실시하는 평가는 목록에 포함되지 않은 어휘는 분석에서 제외되며 어떤 어휘를 자주 사용하는가에 대한 사용 빈도 분석은 하기 어렵다. 자발화 자료에서 용언을 분석한 연구에서도 제한점이 나타난다. 대상자가 소수거나(Lee et al., 2003), 보편 타당한 어휘 분류 기준으로 보기 어려운 점 등이다(Oh, 2014; Park & Rhee, 2011). Park과 Rhee (2011)는 형용사와 관형사를 동일한 품사로 분류하였고, Oh (2014)는 동사만을 대상으로 하여 형용사를 제외하였다는 점과, 본동사와 보조 동사를 구분하지 않은 점, '있다, 없다'와 같은 특정 어휘를 제외한 점은 이들 연구가 지니는 한계로 지적될 수 있다.

이런 점을 고려하여 본 연구에서는 대표성과 균형성을 확보한 아동 구어 자발화 자료를 대상으로 용언의 어휘 유형 수와 사용 빈도를 분석하고 고빈도 어휘 목록을 제시하는 것을 목적으로 하였다.

연구 방법

자료 수집

아동 언어에 나타나는 용언의 어휘 유형 수 및 사용 빈도의 특징을 밝히기 위해 본 연구에서는 나사렛 말뭉치 중 아동 구어 말뭉치를 이용하였다. 나사렛 아동 구어 말뭉치는 만 2세부터 만 5세까지, 각 연령에서 남녀 10명씩 20명, 총 80명 아동이 부모나 형제, 자매 및 주변인과 일상에서 나눈 자유로운 대화를 전사하여 구축한 말뭉치이다. 서울, 경기, 충청 지역에 거주하며 청각과 언어 발달에 문제가 없는 일반 아동으로 선정하였고, 아동의 부모에게 녹음기를 장착할 수 있도록 제작된 앞치마와 녹음기(Click Voice S300)를 지급한 뒤 일주일 동안 가정에서 일상적인 생활 및 놀이 시간에 자연스럽게 발화한 것을 10시간 이상 녹음하고 이를 수거하여 전사 도구를 활용하여 전사한 자료이다. 나사렛 말뭉치의 전사 및 발화 구분 원칙은 Yoon, Kim, Kim, Chang과 Cha (2013)에 제시되어 있다.

자료의 선별과 가공

본 연구에서는 나사렛 말뭉치의 발화 자료 중 2차 전사가 끝난 발화에서 아동 1인당 300여의 연속 발화를 무작위로 추출한, 80명 아

CSD COMMUNICATION SCIENCES & DISORDERS

동의 총 27,486발화를 대상으로 국어학 박사 3인이 3차 검토를 실시한 65,783어절을 조사 대상으로 하였다. 1인당 발화 수 평균은 343.1, 표준편차는 23.247이다.

형태소 사용 빈도 및 통계 자료 추출은 KSTARS (Kwak & Chang, 2014)에 내장된 형태소 분석기를 이용하여 1차로 자동 태깅을 하고, 이 결과물을 웹의 표준국어대사전(http://stdweb2.korean.go.kr/main.jsp)을 기준으로 2차로 수동 정련하여 최종 분석 자료를 얻었다. 수동 정련에서는 특히 우리말의 특성상 동사와 형용사 두 품사로 모두 통용되므로 자동 정련으로는 분류를 할 수 없는 '있다'의 산출 발화를 모두 찾아 동사, 또는 형용사로 분류하여 품사 태깅 정확도를 높였다. 진행의 양태 어미 'ㄴ다'와 결합한 경우나, 명령형 어미 '-어라', 청유형 어미 '-자' 등과 결합한 '있는다', '있어라', '있자' 등을 동사로 분류하고, '동작성의 의미가 없는 존재의 의미를 지닌 상태'로 사용된 '있다'에 대해서 형용사로 분류하였다. 아울러 형용사로서의 '있다'와 반의어 관계에 있는 '없다'도 본 연구에서는 모두 포함하여 분석하였다.

분석 방법

2-5세 아동 발화의 전체 품사에서 용언의 품사별 유형 수와 사용 빈도의 비중과 연령별로 용언을 구성하는 품사의 비율을 분석하였다. 또한 각 품사별로 유형 수와 사용 빈도의 연령 간 차이를 검증하기 위하여 SPSS version 12.0으로 일원분산분석을 실시하였다. 집단 간 차이의 사후검증은 Scheffé를 사용하였다. 마지막으로 각 품사별로 사용 빈도가 높은 어휘를 제시하였다.

연구 결과

용언의 유형 수와 사용 빈도

전체 품사 빈도 중 용언의 유형 수와 사용 빈도

이 연구에서 분석 대상으로 삼은 2-5세 아동의 품사 전체의 유형 수는 4,679개 유형이며, 사용 빈도는 105,284회이다. 그 중에서 용언의 유형 수는 813개로 전체 유형 수 중 17.4%, 사용 빈도는 21,130회로 전체 사용 빈도 중 20.1%의 비중을 차지한다(Table 1). 연령별 유형 수와 사용 빈도의 비중은 2세는 20.1%와 18.5%, 3세는 22.2%와 19.8%, 4세는 21.2%와 20.7% 5세는 21.3%와 20.5%로 유사하게 나타났다.

Table 1. Type and token ratios of predicates

Age (yr)	Type (predicates/total)		Token (predicates/total)	
	Number	Ratio (%)	Number	Ratio (%)
2-5	813/4,679	17.4	21,130/105,284	20.1
2	311/1,544	20.1	3,659/19,760	18.5
3	468/2,105	22.2	5,276/26,565	19.8
4	486/2,285	21.2	5,870/28,239	20.7
5	562/2,639	21.3	6,325/30,720	20.5

용언의 품사별 유형 수와 사용 빈도

용언의 품사별 유형 수와 사용 빈도는 Table 2에서 제시하였다. 2-5세 아동 말뭉치에서 동사는 599개 유형이 12,722회의 사용 빈도를 보여 유형 수에서는 전체 용언 중에 73.7%, 사용 빈도에서는 60.2%를 보였다. 형용사는 190개 유형이 5,459회의 사용 빈도를 보여 유형 수에서는 23.4%, 사용 빈도에서는 25.5%를 보였다. 보조 동사 및 보조 형용사의 유형 수와 사용 빈도는 매우 낮은데, 보조 동사는 18개 유형(2.2%)이 2,675회(12.7%) 사용되었으며, 보조 형용사는 6개 유형(0.7%)이 274회(1.3%) 사용되었다. Table 2에서 보듯이 용언의 품사별 순위는 유형 수와 사용 빈도 모두에서 '동사>형용사>보조 동사>보조 형용사' 순으로 나타났다. 동사의 유형 수와 사용 빈도의 비율이 형용사의 비율보다 두 배가 넘고, 보조 용언은 유형 수와 사용 빈도의 비율이 매우 낮았다. 이러한 비율은 2, 3, 4, 5세 연령 집단에서 동일하게 나타났다.

Table 2. Types and tokens of word classes in total predicates

Age (yr)	Word class	Type		Token	
		Number	Ratio (%)	Number	Ratio (%)
2-5	VV	599	73.7	12,722	60.2
	AJ	190	23.4	5,459	25.8
	VX	18	2.2	2,675	12.7
	AX	6	0.7	274	1.3
	SUM	813	100	21,130	100
2	VV	221	71.1	2,322	63.5
	AJ	75	24.1	919	25.1
	VX	12	3.9	398	10.9
	AX	3	1.0	20	0.5
	SUM	311	100	3,659	100
3	VV	325	69.4	3,263	61.8
	AJ	124	26.5	1,333	25.3
	VX	15	3.2	603	11.4
	AX	4	0.9	77	1.5
	SUM	468	100	5,276	100
4	VV	348	71.6	3,440	58.6
	AJ	119	24.5	1,558	26.5
	VX	15	3.1	782	13.3
	AX	4	0.8	90	1.5
	SUM	486	100	5,870	100
5	VV	403	71.7	3,696	58.4
	AJ	138	24.6	1,650	26.1
	VX	16	2.8	892	14.1
	AX	5	0.9	87	1.4
	SUM	562	100	63,265	100

VV = verb; AJ = adjective; VX = auxiliary verb; AX = auxiliary adjective.

http://dx.doi.org/10.12963/csd.14168

Table 3. Descriptive statistics and ANOVA results of types and tokens of verbs

Age (yr)	Type					Token				
	M	SD	Min	Max	*F*	M	SD	Min	Max	*F*
2	32.55	10.54	11	54	25.14**	116.15	35.69	31	168	18.39**
3	48.95	7.86	32	63		163.15	25.42	114	193	
4	52.85	10.61	36	74		172.05	32.00	121	219	
5	60.75	12.78	38	81		184.85	30.94	132	240	

**$p < .01$.

Table 4. Descriptive statistics and ANOVA results of types and tokens of adjectives

Age (yr)	Type					Token				
	M	SD	Min	Max	*F*	M	SD	Min	Max	*F*
2	14.55	5.19	7	30	20.33**	45.95	20.54	12	84	16.37**
3	21.15	3.82	15	31		66.65	14.68	40	97	
4	22.80	2.84	18	28		77.90	16.32	54	111	
5	24.85	5.35	15	34		82.50	19.81	51	114	

**$p < .01$.

Table 5. Descriptive statistics and ANOVA results of types and tokens of auxiliary verbs

Age (yr)	Type					Token				
	M	SD	Min	Max	*F*	M	SD	Min	Max	*F*
2	4.45	1.91	0	8	22.68**	19.90	12.77	0	48	15.51**
3	7.30	1.34	6	11		30.15	11.51	14	59	
4	7.30	1.46	3	10		39.10	12.46	17	63	
5	8.70	1.92	5	11		44.60	12.35	24	64	

**$p < .01$.

품사별 유형 수와 사용 빈도의 연령 간 차이

동사

각 연령 집단의 동사 유형 수와 사용 빈도의 평균과 표준편차 및 일원분산분석 결과는 Table 3과 같다. 유형 수에서 2세는 32.55 (± 10.54)개 유형, 3세는 48.95 (± 7.86)개 유형, 4세는 52.85 (± 10.61)개 유형, 5세는 60.75 (± 12.78)개 유형의 동사를 사용하였으며, 연령 간 차이는 통계적으로 유의하였다($F_{(3,76)} = 25.14$, $p < .01$). 연령에 따른 동사의 사용 빈도는 2세는 116.15 (± 35.69)회, 3세는 163.15 (± 25.42)회, 4세는 172.05 (± 32.00)회, 5세는 184.85 (± 30.94)회 동사를 사용하였다. 동사 사용 빈도의 연령 집단 간 평균은 통계적으로 유의한 차이가 있었다($F_{(3,76)} = 18.39$, $p < .01$). Scheffé 사후 검정 결과 유형 수의 연령 집단 간 차이는 2세가 모든 집단과 유의한 차이를 보였으며, 3세는 5세와 유의한 차이를 보였다. 사용 빈도는 2세만 모든 집단과 유의한 차이를 보였다.

형용사

각 연령 집단의 형용사의 유형 수와 사용 빈도의 평균과 표준편차 및 일원분산분석 결과는 Table 4와 같다. 유형 수에서 2세는 14.55 (± 5.19)개 유형, 3세는 21.15 (± 3.82)개 유형, 4세는 22.80 (± 2.84)개 유형, 5세는 24.85 (± 5.35)개 유형의 형용사를 사용하였으며, 연령 간 차이는 통계적으로 유의하였다($F_{(3,76)} = 20.33$, $p < .01$). 각 연령 집단의 형용사의 사용 빈도에서 2세는 45.95 (± 20.54)회, 3세는 66.65 (± 14.68)회, 4세는 77.90 (± 16.32)회, 5세는 82.50 (± 19.81)회의 형용사를 사용하였으며, 연령 간 차이는 통계적으로 유의하였다($F_{(3,76)} = 16.37$, $p < .01$). Scheffé 사후검증 결과 유형 수와 사용 빈도 모두 2세 집단은 모든 집단과 유의한 차이를 보였으나 다른 연령 집단은 집단 간의 차이가 유의하지 않았다.

보조 용언

연령 집단별 보조 동사의 유형 수와 사용 빈도의 평균과 표준편차 및 일원분산분석 결과는 Table 5와 같다. 유형 수에서 2세는 4.45

CSD COMMUNICATION SCIENCES & DISORDERS

Table 6. Descriptive statistics and ANOVA results of types and tokens of auxiliary adjectives

Age (yr)	Type					Token				
	M	SD	Min	Max	*F*	M	SD	Min	Max	*F*
2	0.55	0.69	0	2	10.59**	1.00	1.52	0	6	6.49**
3	1.45	0.69	0	3		3.85	2.91	0	10	
4	1.70	0.86	0	3		4.50	3.32	0	11	
5	2.00	1.12	0	4		4.35	3.38	0	13	

**$p < .01$.

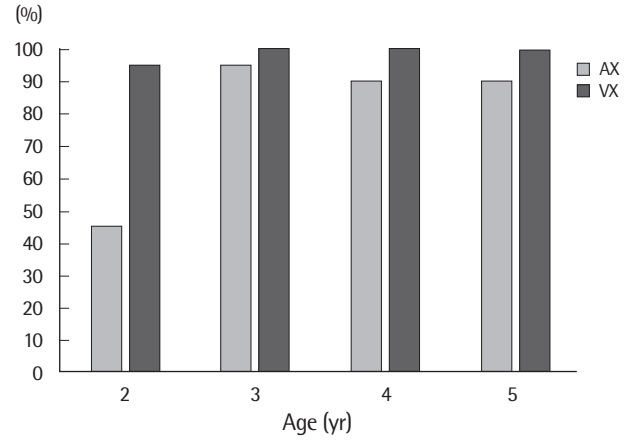

Figure 1. Ratios of children who use auxiliary predicates. AX = auxiliary adjective; VX = auxiliary verb.

(± 1.91)개, 3세는 7.30 (± 1.34)개, 4세는 7.30 (± 1.46)개, 5세는 8.70 (± 1.92)개의 보조 동사 유형을 사용하였으며, 연령 간 차이는 통계적으로 유의하였다($F_{(3,76)} = 22.68$, $p < .01$). 각 연령 집단의 보조 동사 사용 빈도 평균과 표준편차 및 일원분산분석 결과는 2세는 19.90 (± 12.77)회, 3세는 30.15 (± 11.51)회, 4세는 39.10 (± 12.46)회, 5세는 44.60 (± 12.35)회의 보조 동사를 사용하였으며, 연령 간 차이는 통계적으로 유의하였다($F_{(3,76)} = 15.51$, $p < .01$). Scheffé 사후검증 결과는 보조 동사의 유형 수는 2세와 다른 모든 집단 간의 차이가 유의하였고, 사용 빈도는 2세와 4세, 2세와 5세, 3세와 5세 집단 간에서 유의한 차이를 보였다.

각 연령 집단의 보조 형용사의 유형 수와 사용 빈도의 평균과 표준편차 및 일원분산분석 결과는 Table 6과 같다. 유형수에서 2세는 0.55 (± 0.69)개, 3세는 1.45 (± 0.69)개, 4세는 1.70 (± 0.86)개, 5세는 2.00 (± 1.12)개의 보조 형용사를 사용하였으며, 연령 간 차이는 통계적으로 유의하였다($F_{(3,76)} = 10.59$, $p < .01$). 보조 형용사의 사용 빈도에서 2세는 1.00 (± 1.52)회, 3세는 3.85 (± 2.91)회, 4세는 4.50 (± 3.32)회, 5세는 4.35 (± 3.38)회의 보조 형용사를 사용하였으며, 연령 간 차이는 통계적으로 유의하였다($F_{(3,76)} = 6.49$, $p < .01$). Scheffé 사후검증 결과는 2세가 모든 집단과 유의한 차이를 보였으며 다른 연령 집단은 집단 간 차이가 유의하지 않았다.

동사와 형용사는 모든 아동에게서 산출되었으나, 보조 동사와 보조 형용사는 모든 연령 집단에서 산출되지는 않았다. 보조 동사는 2세 집단 1명이 산출하지 않았으며, 보조 형용사는 2세 11명, 3세 1명, 4세 2명, 5세 2명이 산출하지 않았다. 이에 따라 각 연령별로 보조 동사와 보조 형용사를 산출한 아동의 비율을 제시하면 Figure 1과 같다.

빈도에 따른 용언 목록

빈도에 따른 용언 목록은 품사별로 고빈도 사용 단어를 중심으로 제시한다. 동사와 형용사는 고빈도 사용 순위 30위까지, 유형 수가 많지 않은 보조 동사는 사용 순위 10위까지, 보조 형용사는 전체 유형(6개)을 모두 제시하였다(Appendixes 1-4).

동사

사용 빈도 1순위는 {하-}로 1,888회가 사용되어 전체 중에서 14.8%의 비율을 나타냈다. 2순위는 {되-}로 1,070회, 8.4% 비율을 나타냈으며, 3순위는 {먹-}으로 765회, 6.0%의 비율을 나타냈다. 이 세 동사의 누적 사용 빈도가 29.3%이다. 누적 빈도 상위 50%까지 해당하는 동사는 {하-, 되-, 먹-, 보-, 가-, 주-, 알-, 만들-, 나-, 넣-}으로 모두 11개이다. 동사의 사용 빈도 순위에 따른 어휘 목록은 Appendix 1에 수록하였다.

형용사

가장 높은 사용 빈도를 보인 형용사는 {있-}으로 903회 사용되어 16.5%를 나타냈다. 2순위는 {이렇-}으로 666회, 12.2%의 비율을 나타냈으며, 3순위는 {그렇-}으로 396회, 7.3%의 비율을 나타냈다. 이 세 형용사의 사용 비율이 전체 형용사 중에서 36%를 나타냈다. 전체 190개 유형 중에서 {있-, 이렇-, 그렇-, 없-, 아니-, 어떻-} 6개의 형용사가 누적 빈도 52.2%를 나타냈다. 형용사의 사용 빈도 순위에 따른 어휘 목록은 Appendix 2에 제시하였다.

http://dx.doi.org/10.12963/csd.14168

보조 용언

보조 동사의 사용 빈도 1위는 {보-}로 638회 사용되어 전체 보조 동사 중에 23.9%의 비율을 나타냈으며, 2위는 {주-}로 589회 사용되어 22.0%의 비율을 나타냈다. 18개의 유형 중에서 이 두 보조 동사 유형이 45.9%로 사용되었다. 보조 동사의 사용 빈도 순위에 따른 어휘 목록은 Appendix 3에 제시하였다.

보조 형용사는 조사 대상 말뭉치에서 6개 유형밖에 나타나지 않았다. 그 중에서 {싶-}이 169회 사용되어 61.7%의 비율로 1순위였으며, {않-}이 54회로 19.7%의 비율을 보여 2순위를 나타냈다. 이 두 보조 형용사가 전체 보조 형용사 사용 빈도 중에서 81.4%의 비율을 보였다. 보조 형용사의 사용 빈도 순위에 따른 어휘 목록은 Appendix 4에 제시하였다.

논의 및 결론

본 연구는 학령 전기 아동의 자발화 말뭉치에서 나타나는 용언의 어휘 빈도 양상을 고찰하여 언어 치료를 위한 언어 평가 및 중재에 활용할 수 있는 기초 자료를 제공하는 것을 목적으로 하였다. 이를 위하여 나사렛 말뭉치 중에서 2-5세 아동의 일상 자발화를 대상으로 용언의 어휘 유형 수 및 사용 빈도, 용언의 품사별 유형 수와 사용 빈도의 연령 간 차이를 밝히고 사용 빈도가 높은 어휘를 제시하였다. 본 연구는 총 27,486발화를 대상으로 65,783어절 규모를 분석하였는데, 이 정도의 규모는 문어 말뭉치 기준으로 5만 어절이면 대략 300페이지 책 한 권 규모이므로 구어 말뭉치로서 대표성을 가진다 할 수 있다.

2-5세 시기 아동에게서 나타나는 용언 빈도 정보의 전반적인 특징

나사렛 말뭉치의 2-5세 아동의 자발화를 분석한 본 연구 결과, 용언의 비중은 전체 유형 수 중 17.4%, 사용 빈도는 20.1%로 나타났다. 연령별로 유형 수와 사용 빈도는 2세 20.1%-18.5%, 3세 22.25-19.8%, 4세 21.2%-20.7%, 5세 21.3%-20.5%로 나타났다. 이러한 비중은 최근 선행 연구와 비교하여 낮은 수준이다. Park과 Rhee (2011)는 경남, 부산 지역의 3-5세 정상 발달 아동, 연령별 6명씩 총 18명을 대상으로 일상적 이야기, 사진보고 대화하기 등 5가지 공통 활동 중심의 자발화를 30-40분씩 수집하여 연령에 따른 품사별 어휘 수, 연령별로 새롭게 출현한 어휘 수, 연령에 따른 어휘 빈도를 제시하였다. 품사별 어휘 수에서는 전체 품사 중 동사가 29.4% (3세), 30.9% (4세), 31.1% (5세), 형용사가 12.3% (3세), 12.0% (4세), 12.4% (5세)를 보인다는 것을 보고하였다. 본 연구에서 나타난 용언의 유형 수 비중이 이들의 연구보다 낮은 이유는 자발화의 수집 방법에 따른 차이에서 기인했을 것으로 보인다. 본 연구의 자발화 수집 상황은 일정한 활동으로 국한하지 않고 가정에서 아동과 친숙한 성인 1인이 자유롭게 상호작용하는 상황이었다. 따라서 선행 연구와 같이 특정 활동을 제시하고 자발화를 수집하는 경우보다 다양한 어휘를 사용하게 되므로, 상대적으로 열린 집합의 성격을 지닌 체언의 비중이 높아져 용언의 비중이 낮아진 것으로 추정된다.

용언을 이루는 품사별 정보를 보면 이 시기 아동에게서 관찰되는 용언의 품사별 빈도 정보는 유형 수와 사용 빈도 모두에서 동사의 비율이 형용사의 비율보다 두 배가량 높았으며, 보조 동사와 보조 형용사는 유형 수와 사용 빈도 모두에서 매우 적게 나타났다. 이는 성인의 언어 자료 분석에서도 동일하게 나타나는 특징이며(Kang & Kim, 1997, 2009), 동사나 형용사가 실질적 어휘 의미를 나타내는 단어인 반면에 보조 용언은 주로 양태 의미를 나타내는 범주라는 것을 상기할 때 이 시기의 아동에게서 보조 용언의 유형 수 및 사용 빈도가 낮게 나타나는 것은 자연스러운 결과이다.

한편 보조 동사와 보조 형용사의 사용 양상에서 다소 차이가 있었는데, 보조 형용사는 유형 수와 사용 빈도의 비율이 모두 2% 미만으로 극히 적었으나, 보조 동사는 유형 수가 2.2%였던 것에 반해 사용 빈도는 12.7%로 나타났다. 즉 보조 동사는 적은 유형이 반복적으로 많이 사용되는 양상을 보였다. 이는 {이/가, 을/를, 에서, 로}와 같은 한정된 문법 형태소가 열린 집합으로 구성된 어휘 형태소와 반복적으로 결합하여 사용되는 것과 비슷한 양상을 보인다. 즉 어휘 형태소에 비해 상대적으로 높은 기능을 부담하고 있는 문법 형태소의 사용 양상과 그 특징이 유사하다.

연령에 따른 용언의 습득 양상

연구 결과 용언의 유형 수와 사용 빈도는 대부분 2세와 그 이후 집단 간에 차이를 보였다. 유형 수에서는 동사, 형용사, 보조 동사, 보조 형용사 모두 2세가 다른 집단과 차이를 보였으며, 동사에서만 3세와 5세의 차이가 더 있었다. 사용 빈도에서는 동사, 형용사, 보조 형용사에서 2세가 다른 집단과 차이를 보였으며, 보조 동사에서만 2세와 3세의 차이가 없고 3세와 5세의 차이가 있었다. 이를 통해 용언의 어휘 유형 습득과 사용 빈도 증가는 모두 2세에서 3세 시기에 걸쳐 유의하게 증가하며, 3세부터 5세까지는 평균적으로 어휘 유형 습득이 증가하고 사용 횟수가 빈번해지기는 하나 유의하지는 않다고 할 수 있다.

2세에서 3세에 용언의 습득과 사용이 모두 유의하게 증가하는 것은 이 시기가 기본 문법기에 들어서며 구문적인 발달과 문법 형태소 발달이 진행되는 시기라는 견해를 뒷받침한다(Pae, 2006). 또

한 본 연구와 같은 말뭉치 자료를 분석한 Jung과 Yoon (2013)에서 아동의 MLU가 2-3세에서만 유의한 차이가 있었던 것과 연관하여 용언의 발달과 구문론적 발달의 관계를 추론해 볼 수 있을 것이다. 연령에 따른 용언의 사용 빈도 증가는 Park과 Rhee (2011)에서도 나타나나 이들의 연구는 3세부터 시작하여 본 연구에서 나타난 것과 같은 2-3세 사이의 용언의 급격한 증가를 확인할 수는 없었다. 반면 Lee, Choi, Chang과 Lee (2008)는 24개월에서 36개월의 아동에서 동사 표현이 급격히 증가하였다고 보고하여 본 연구 결과와 일치하나, 이들은 18개월부터 36개월까지의 아동을 대상으로 하였다.

보조 용언의 산출 시기는 용언과 차이가 있었다. 즉 동사와 형용사는 모든 아동이 산출하였으나, 보조 용언은 산출을 하지 않는 연령군이 있었다. 보조 동사와 보조 형용사의 연령별 산출 시기에도 다소 차이가 있었다. 보조 동사는 2세 아동 1명만이 산출을 하지 못했으나, 보조 형용사는 2세에는 50% 미만의 아동이 산출하고 3세에는 90% 이상의 아동이 산출하나 5세에 이르러도 모든 아동이 산출하지는 않는다는 것을 알 수 있다.

2-5세 아동에게서 관찰되는 고빈도 용언의 특징

고빈도 용언의 특징을 보면 소수의 어휘가 높은 사용 빈도를 보이고 있는 점이다. 즉 동사의 경우 599개 유형 중 상위 빈도 11개의 동사가 약 50%의 누적 빈도를 차지했다. 형용사는 190개 유형 중 상위 빈도 6개가 누적 빈도 약 50%를 차지했다. 보조 동사와 보조 형용사의 경우는 각각 2개와 1개의 어휘가 누적 빈도 약 50% 수준으로 나타났다. 이렇게 소수의 어휘가 높은 누적 빈도를 보이는 것은 성인 및 청소년 말뭉치의 연구에서도 나타난다. 문어 말뭉치를 대상으로 한 결과이지만 1,500만 어절 세종형태의미분석 말뭉치를 기반으로 분석한 Kang과 Kim (2009)에서는 {하-, 있-, 되-, 보-} 등 49개의 동사가 50%를 보였으며, {없-, 같-, 그렇-, 크-, 많-, 좋-, 어떻-, 이렇-} 8개의 형용사가 누적 빈도 상위 50%를 보였고, {있-, 하-, 않-} 3개의 보조 용언이 상위 57.2%를 보였다. 초·중·고등학생의 구어 어휘 조사인 Chang 등(2012)에서도 동사는 {하-, 되-, 가-} 등 초등학교 저학년 17개, 고학년 16개, 중학생 15개, 고등학교 16개의 어휘가 누적 빈도 상위 50%를 보였고, 형용사는 초등학교 저학년 {있-, 좋-, 없-, 그렇-} 4개, 고학년, 중학생 5개({같-} 추가), 고등학생 6개({많-} 추가)가 누적 빈도 상위 50%를 보였으며, 보조 용언은 초등학교 저학년, 고학년 {보-, 있-, 주-, 가지} 4개, 중학생, 고등학생 5개({않-} 추가)가 누적 빈도 상위 50%를 보였다.

동사의 사용 빈도 순위에서 1, 2순위의 {하-}, {되-}는 실질적이고 구체적인 의미를 지닌 동사라기보다는 추상적이고 포괄적인 의미를 지닌 동사이다. 3위부터 8위까지는 {먹-, 보-, 가-, 주-, 오-, 알-}인데 이는 인간의 생활에 가장 기본이 되는 동작 의미를 지닌 동사이며, 9위부터 11위까지는 {만들-, 나-, 넣-}은 아동의 생활과 밀접한 동작 의미를 지닌 동사이다. 그 뒤를 이어 {맞다/모르다}와 같이 지식에 대한 앎의 여부를 표현하는 동사, {나오-, 찾-, 놀-, 타-, 보이-, 사-, 잡-, 쓰-, 자-, 붙이-, 가지-, 빼-, 놓-}과 같이 아동의 생활, 혹은 놀이와 관련되는 동사가 주로 사용되는 양상을 나타냈다.

형용사의 사용 빈도 6위 안에 있는 {있-, 이렇-, 그렇-, 없-, 아니-, 어떻-}은 구체적이고 실질적 의미를 지닌 형용사가 아니라 추상적이고 포괄적인 의미를 지닌 형용사라는 점에서 동사와 비슷한 특징을 드러낸다. 사용 빈도 10위 내에 있는 실질적 의미를 지닌 형용사는 {싫-, 좋-, 아프-, 예쁘-}인데 가장 기본적인 상태 표현과 관련된 형용사라는 점이 흥미롭다. 사용 빈도 30위 내 고빈도 형용사는 반의어 관계를 나타내는 짝({있다/없다}, {덥다/춥다}, {크다/작다}, {싫다/좋다}, {이렇다/그렇다/아니다})이 함께 나타난다는 특징을 보이며, 기초적인 상태 표현의미를 지닌 단어들({무섭다, 재미있다, 괜찮다}, {아프다, 배고프다} {같다, 똑같다, 많다}, {뜨겁다, 세다, 맛있다, 예쁘다, 이상하다})이 나타나고 인사와 관련되는 표현({미안하다/고맙다})이 관찰된다는 것이 특징적이다.

보조 동사에서는 {보-, 주-}가 전체 보조 동사의 사용 빈도에서 45%를 나타냈다. {보-}는 '가 보다, 먹어 보다, 입어 보다'처럼 보조 동사로 쓰일 때 '무엇을 시도해 보다'라는 의미로 사용된다. 또한 {주-}는 '보여 주다, 먹여 주다, 고쳐 주다'처럼 '다른 사람을 위해 어떤 행동을 하는 것'이라는 의미로 사용된다. 아동에게서 가장 빈번한 의미로 사용되는 양태 표현이 '무엇을 시도하는 것', '누구를 위해 행동을 하는 것'이라는 점을 알 수 있다.

보조 형용사는 {싶-}이 1위였으며, 전체 중에 61.7% 비율로 사용되었다. {싶-}은 '먹고 싶다, 자고 싶다, 가고 싶다'에서 알 수 있듯이 욕구를 나타내는 표현이다. 아동은 '-고 싶다'와 같은 기본적인 구문 구조를 통해 자신의 욕구를 표현한다는 것을 알 수 있다.

이 연구는 대표성과 균형성을 확보한 말뭉치를 기반으로 2-5세 아동의 용언 빈도 정보를 제시하고 그 특징과 양상을 논의하였다. 2-5세 아동에게서는 본 용언과 보조 용언으로 구분할 때 보조 용언의 유형 수와 사용 빈도가 극히 낮게 나타났으며, 본 용언에서는 동사의 유형 수 및 사용 빈도가 형용사에 비해 두 배 가량 높게 나타난다는 것을 확인할 수 있었다. 연령에 따른 용언의 발달 양상에서는 동사와 형용사 모두 2-3세 시기의 증가가 뚜렷하게 관찰되었다. 이 시기 아동에게서 관찰되는 고빈도 동사는 구체적인 어휘 의미를 지닌 것보다 추상적이고 포괄적인 어휘 의미를 담고 있는 것이나, 인간 생활의 가장 기본 동작을 표현하는 동사, 아동의 생활이나 놀이를 표현하는 동사가 빈번하게 사용된다는 것을 알 수 있었다.

http://dx.doi.org/10.12963/csd.14168

형용사에서도 추상적이고 포괄적인 의미를 지닌 어휘와 가장 기본적인 상태 표현을 나타내는 어휘가 빈번하게 사용되었으며, 고빈도 사용 형용사에는 반의어 관계에 있는 단어들이 관찰된다는 것이 특징적이었다.

본 연구의 결과로 제공한 학령 전기 일반 아동의 연령별 용언 발달 특성은 언어 장애 아동의 용언 발달을 평가하는 기초 자료로 활용할 수 있다. 또한 용언의 품사별로 제시한 고빈도 어휘 목록은 언어 장애 아동의 치료 목표 어휘 선정 시 중요한 근거 자료가 될 것으로 기대된다. 지면상 제시하지 못한 전체 어휘에 대한 정보는 추후 출판할 예정이다.

용언의 발달을 살펴보며 성별에 따른 차이를 분석하지 않은 것은 본 연구의 제한점이다. 어휘 발달에서 성별의 영향을 보고한 선행 연구들이 있는 만큼 향후 연구에서는 성별과 연령에 따른 어휘 발달을 분석할 필요가 있을 것이다. 또한 용언의 온전한 사용 양상을 살피기 위해서는 어미의 활용 빈도도 함께 살피는 것이 중요하나 본 연구에서는 용언의 어간에 대한 분석만을 하였다. 이후 용언의 어간과 어미의 결합 빈도 및 사용 양상에 대한 연구는 후고를 기약하도록 한다.

REFERENCES

Bassano, D. (2000). Early development of nouns and verbs in French: exploring the interface between lexicon and grammar. *Journal of Child Language, 27,* 521-559.

Caselli, M. C., Bates, E., Casadio, P., Fenson, J., Fenson, L., Sanderl, L., & Weir, J. (1995). A cross-linguistic study of early lexical development. *Cognitive Development, 10,* 159-199.

Chang, K. H., Lee, S. H., Lee, P. Y., Kim, M. H., Kim, J. S., & Jeon, E. J. (2012). *Research on students' spoken vocabulary.* Seoul: Jisikkwakyoyang.

Golinkoff, R. M., & Hirsh-Pasek, K. (2008). How toddlers begin to learn verbs. *Trends in Cognitive Sciences, 12,* 397-403.

Jung, Y. H., & Yoon, M. S. (2013). Mean length of utterance for typically developing children of 2 to 4 years. *Korean Journal of Early Childhood Special Education, 13,* 55-73.

Kang, B. M., & Kim, H. K. (1997). *Analysis of type and token frequency of Korean morpheme.* Seoul: Research Institute of Koran Studies.

Kang, B. M., & Kim, H. K. (2009). *Token frequency of Korean: analyze of 15 million words in Sejong corpus.* Seoul: Hankookmunhwasa.

Kauschke, C., Lee, H. W., & Pae, S. (2007). Similarities and variation in noun and verb acquisition: a crosslinguistic study of children learning German, Korean, and Turkish. *Language and Cognitive Processes, 22,* 1045-1072.

Kim, Y. K. (2012). *Word class: introduction to Korean grammar.* Seoul: Jisikkwakyoyang.

Ko, Y. G., & Ku, B. G. (2008). *Korean grammar.* Phaju: Jipmoondang.

Kwon, E. G., & Pae, S. Y. (2005). Noun & verb naming abilities of the children with autism spectrum disorder. *Korean Journal of Communication Disorders, 10,* 32-42.

Kwak, S. D., & Chang, M. S. (2014). CosmoScriBe 2.0: The development of Korean transcription tools. *Journal of the Korean Institute of Intelligent Systems, 24,* 323-329

Lee, H. J., Park, C. I., Park, E. S., & Kim, H. H. (2003). A study on the relationship between the children's development of predicates and syntactic ability. *Korean Journal of Communication Disorders, 8,* 20-40.

Lee, H. R., Choi, Y. R., Chang, Y. K., & Lee, S. B. (2008). The early verb acquisition in Korean children. *Journal of Speech & Hearing Disorders, 17,* 65-77.

Lee, H. Y. (2004). Sprachvergleichende studie zum benennen und verstehen von nomen und verben bei deutschen und koreanischen kindem. *Dogil Munhak, 92,* 365-384.

Lee, Y. K., & Kim, Y. T. (2003). Word (noun and verb) production in children with specific language impairment. *Korean Journal of Communication Disorders, 8,* 1-19.

MacWhinney, B. (2000). *The childes project: tools for analyzing talk* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

McDonough, C., Song, L., Hirsh-Pasek, K., Golinkoff, R. M., & Lannon, R. (2011). An image is worth a thousand words: why nouns tend to dominate verbs in early word learning. *Developmental Science, 14,* 181-189.

Miller, J. F. (2003). *Systematic analysis of language transcripts (SALT, version 8).* Madison, WI: University of Wisconsin-Madison, Waisman Center, Language Analysis Laboratory.

Oh, K. A. (2014). *Semantic categorization of preschooler's verbs in spontaneous conversation* (Master's thesis). Nazarene University, Cheonan, Korea.

Pae, S. Y. (2006). Grammatical development of Korean young children. *Association for Korean Linguistics, 31,* 31-45.

Pae, S. Y., Chang, Y. K., Kwak, K. J, Sung, H., R., & Sim, H. O. (2004). MCDI-K referenced expressive word development of Korean children and gender differences. *Korean Journal of Communication Disorders, 9,* 45-56.

Park, H. J., & Pae, S. Y. (2003). Early expressive vocabulary development of Korean children with hearing impairment. *Korean Journal of Communication Disorders, 8,* 66-81.

Park, M. H., & Rhee, J. A. (2011). A study for characteristics of vocabulary in spontaneous language sample aged from 3 to 5. *Korean Society of Education for Hearing-Language Impairments, 2,* 1-13.

Yoon, M. S., Kim, S. J., Kim, J. M, Chang, M., S., & Cha, J. E. (2013). Reliable sample size for mean length of utterance analysis in preschooler. *Communication Sciences & Disorders, 18,* 368-378.

http://dx.doi.org/10.12963/csd.14168

Appendix 1. List of verbs with high frequencies

Order	Verb	Frequency	Ratio (%)	Cumulated ratio (%)
1	하	1,888	14.8	14.8
2	되	1,070	8.4	23.3
3	먹	765	6.0	29.3
4	보	622	4.9	34.2
5	가	551	4.3	38.5
6	주	462	3.6	42.1
7	오	291	2.3	44.4
8	알	282	2.2	46.6
9	만들	224	1.8	48.4
10	나	165	1.3	49.7
11	넣	164	1.3	51.0
12	모르	155	1.2	52.2
13	나오	138	1.1	53.3
14	맞	136	1.1	54.3
15	찾	134	1.1	55.4
16	놀	122	1.0	56.4
17	타	118	0.9	57.3
18	보이	110	0.9	58.1
19	사	110	0.9	59.0
20	잡	104	0.8	59.8
21	쓰	100	0.8	60.6
22	자	92	0.7	61.3
23	붙이	90	0.7	62.0
24	가지	82	0.6	62.7
25	빼	77	0.6	63.3
26	놓	75	0.6	63.9
27	그리	70	0.6	64.4
28	그러	65	0.5	64.9
29	숨	65	0.5	65.5
30	입	65	0.5	66.0

Appendix 2. List of adjectives with high frequencies

Order	Verb	Frequency	Ratio (%)	Cumulated ratio (%)
1	있	903	16.5	16.5
2	이렇	666	12.2	28.7
3	그렇	396	7.3	36.0
4	없	359	6.6	42.6
5	아니	316	5.8	48.4
6	어떻	211	3.9	52.2
7	싫	188	3.4	55.7
8	좋	186	3.4	59.1
9	아프	144	2.6	61.7
10	예쁘	126	2.3	64.0
11	맛있	112	2.1	66.1
12	같	105	1.9	68.0
13	크	101	1.9	69.8
14	많	78	1.4	71.3
15	괜찮	62	1.1	72.4
16	춥	60	1.1	73.5
17	무섭	59	1.1	74.6
18	재미있	59	1.1	75.7
19	똑같	57	1.0	76.7
20	작	41	0.8	77.5
21	이상하	40	0.7	78.2
22	뜨겁	36	0.7	78.9
23	귀엽	33	0.6	79.5
24	덥	31	0.6	80.0
25	빨갛	31	0.6	80.6
26	배고프	30	0.5	81.2
27	세	30	0.5	81.7
28	고맙	29	0.5	82.2
29	미안하	29	0.5	82.8
30	안녕하	29	0.5	83.3

Appendix 3. List of auxiliary verbs with high frequencies

Order	Verb	Frequency	Ratio (%)	Cumulated ratio (%)
1	보	638	23.9	23.9
2	주	589	22.0	45.9
3	않	450	16.8	62.7
4	있	251	9.4	72.1
5	하	242	9.0	81.1
6	말	124	4.6	85.8
7	가지	122	4.6	90.3
8	지	104	3.9	94.2
9	놓	65	2.4	96.6
10	버리	32	1.2	97.8

Appendix 4. List of auxiliary adjectives with high frequencies

Order	Verb	Frequency	Ratio (%)	Cumulated ratio (%)
1	싶	169	61.7	61.7
2	않	54	19.7	81.4
3	보	41	15.0	96.4
4	죽[a]	5	1.8	98.2
5	생기[b]	4	1.5	99.6
6	뻔하	1	0.4	100.0

[a]'배고파 죽겠네'와 같은 용례가 있다.
[b]'죽게 생겼어'와 같은 용례가 있다.

CSD COMMUNICATION SCIENCES & DISORDERS

국문초록

2-5세 일반 아동의 어휘 발달: 용언

오재혁[1] · 차재은[2] · 윤미선[3] · 김정미[3] · 장문수[4]

[1]고려대학교 국제어학원, [2]경기대학교 국어국문학과, [3]나사렛대학교 언어치료학과, [4]서경대학교 컴퓨터과학과

배경 및 목적: 이 연구는 학령 전기 아동의 자발화에서 나타나는 용언의 어휘 빈도 양상을 고찰하여 언어 평가 및 중재의 기초 자료로 제공하는 것을 목적으로 한다. **방법:** 나사렛 말뭉치의 2, 3, 4, 5세 아동 각각 20명씩 총 80명에서 300-400발화씩을 선정하여, 총 27,485발화, 65,783어절에 대해 용언의 유형 수와 사용 빈도를 계량화하고 동사와 형용사의 고빈도 어휘를 제시하였다. **결과:** (1) 전체 품사별 유형 수와 사용 빈도에서 용언의 비중은 각각 17.4%와 20.1%였다. (2) 연령에 따른 용언의 발달은 유형 수와 사용 빈도 모두에서 2세와 3세 이상 집단에서 유의한 차이가 있었다. (3) 고빈도 어휘 목록에서 누적 빈도 상위 50%에 속한 동사는 11개, 형용사는 모두 6개였다. **논의 및 결론:** 2-5세 시기의 아동에게서 관찰되는 용언의 품사별 유형 수 및 사용 빈도의 구성 비율은 동일하였고 용언의 유형 수와 사용 빈도의 증가는 2-3세 시기에 두드러지게 나타나, 이 시기가 아동 언어 발달에서 용언의 발달이 뚜렷하게 이루어지는 시기임을 알 수 있었다. 또한 고빈도 어휘의 분석 결과 이 시기의 아동이 소수의 용언을 반복적으로 사용하는 것으로 나타났다. 본 연구의 결과는 향후 언어장애 아동의 평가와 치료의 기초 자료로 사용될 것으로 기대된다.

핵심어: 나사렛 말뭉치, 용언, 유형 수, 사용 빈도, 보조 용언, 동사, 형용사

이 연구는 한국연구재단의 지원을 받아 수행되었음(NRF-2011-B00202).

참고문헌

강범모, 김흥규(1997). *한국어 형태소 및 어휘 사용 빈도의 분석 2.* 서울: 고려대학교 민족문화 연구원.

강범모, 김흥규(2009). *한국어 사용 빈도: 1,500만 어절 세종형태의미분석말뭉치 기반.* 서울: 한국문화사.

고영근, 구본관(2008). *우리말 문법론.* 파주: 집문당.

곽선동, 장문수(2014). CosmoScriBe 2.0: 한국어 전사 도구의 개발. *한국지능시스템학회논문지, 24,* 323-329.

권유진, 배소영(2005). 자폐아동의 명사 및 동사 이름대기 능력. *언어청각장애연구, 10,* 32-42.

김원경(2012). *품사론. 쉽게 읽는 한국어학의 이해.* 서울: 지식과 교양.

박미혜, 이전아(2011). 자발화 표본에 나타난 3-5세 아동의 어휘 특성. *한국청각언어장애교육연구, 2,* 1-13.

박혜진, 배소영(2003). 청각장애 유아의 어휘 발달. *언어청각장애연구, 8,* 66-81.

배소영 (2006). 한국어 발달특성과 학령전기 문법형태소. *한국어학회, 31,* 31-45.

배소영, 장유경, 곽금주, 성현란, 심희옥(2004). MCDI-K를 통해 본 한국유아의 표현어휘발달과 성차. *언어청각장애연구, 9,* 45-56.

오경아(2014). *학령전기 아동이 자발화에서 사용하는 동사의 의미적 분류.* 나사렛대학교대학원 석사학위논문.

윤미선, 김수진, 김정미, 장문수, 차재은(2013). 평균발화길이 분석을 위한 발화 표본의 크기. *언어청각장애연구, 18,* 368-378.

이윤경, 김영태(2003). 단순언어장애 아동들의 낱말 산출 능력: 명사와 동사를 중심으로. *언어청각장애연구, 8,* 1-19.

이해욱(2004). 이름대기검사를 통한 학령 전 독일과 한국아동들의 명사, 동사 발달에 대한 연구. *독일문학, 92,* 365-384.

이현정, 박창일, 박은숙, 김향희(2003). 아동의 구문발달에 따른 용언의 발달. *언어청각장애연구, 8,* 20-40.

이희란 ,최유리 ,장유경, 이승복(2008). 부모보고를 통해 종단 관찰한 한국 아동의 초기 동사발달. *언어치료연구, 17,* 65-77.

장경희, 이삼형, 이필영, 김명희, 김태경, 김정선, 전은진(2012). *초·중·고등학생의 구어 어휘 조사.* 서울: 지식과 교양.

정윤희, 윤미선(2013). 2, 3, 4세 일반 아동의 평균발화길이. *유아특수교육연구, 13,* 55-73.

http://dx.doi.org/10.12963/csd.14168