

한국 한자어의 형태소 분포 조사

주 지 연 *

본고에서는 김창섭(2001, 2013), 노명희(2005)에서 주장한 한자어의 “2자어 3자어 문법”을 바탕으로 한국 한자어의 형태소 분포를 조사하였다. 한국어 한자 형태소가 출현할 수 있는 위치로 자립적 1음절어(W), 선행 차사(α), 선행 원사(A), 후행 원사(B), 후행 차사(β), ‘-하다’ 앞의 요소(R) 등의 분포를 포함하여 총 6가지를 설정하였다. 이 6개 분포 위치의 가능한 조합은 논리적으로 64가지이지만 실제로는 20종의 분포 부류만이 관찰되었다. 가장 많은 한자 형태소가 속한 분포 부류는 AB유형으로 이는 전체 한자 형태소의 37% 정도를 차지하였다. 두 번째로 많은 한자 형태소가 속한 부류는 AB β 유형으로서 이는 10% 정도에 해당한다. 한편 한국어 한자 형태소가 출현할 수 있는 최대의 분포 위치를 보여주는 것은 RW α AB β 유형으로 전체 한자 형태소의 1% 정도에 해당된다.

핵심어: 한자, 한자 형태소, 한자어, 원사, 차사, 분포, 통계적 접근

1. 서론

본고의 목적은 한국어에서 사용되는 한자 형태소¹⁾의 분포 양상을 밝

* 명지대학교 강사

- 1) 현재 한국어에서 사용되는 모든 한자는 일정한 음, 기호, 의미를 갖추고 있다. 이를 고려할 때 모든 한자가 고유한 음성/기호적 형식과 의미를 지니므로 한자 1글자 1글자를 각각 하나의 형태소인 것으로 파악한다. 그러나 1글자에 대하여 복수의 음, 훈을 가진 경우에는 별도의 형태소로 파악하였다. 그러나 ‘석가모니’와 같은 불교 용어의 음역, ‘파사’, ‘불란서’와 같은 국가명의 음역 등에서 나타

히는 것이다. 한국어 연구 분야에서는 꾸준히 한자어에 대한 연구가 이루어져 왔으며 특히 한자어 연구에서 큰 관심을 받아온 주제는 고유어와 한자어를 통합적으로 고유어 문법의 틀에서 분석할 것인가, 분리해서 분석할 것인가 하는 문제였다. 이와 더불어 한자어에 참여하는 한자 형태소들의 형태, 통사적 자격을 규정하는 논의도 활발하게 이루어졌다. 특히 접미, 접두 한자어에 속하는 소수의 한자 형태소를 중심으로 한 심도 있는 연구가 다수 발표되었다. 한국어 한자어에 참여하는 한자 형태소의 자격에 대한 대표적 논의로는 노명희(2005)를 꼽을 수 있다.

그러나 상대적으로 한국어 한자 형태소의 전반적인 분포에 대한 연구는 활발히 이루어지지 않았다. 특히 다수의 형태소를 대상으로 한 본격적인 연구는 더욱 드물다.

그러나 한국 한자어에 대한 심도 있는 연구와 국어/한국어 교육을 위해서는 한국 한자어에 사용되는 한자 형태소들 전반의 분포를 파악할 필요가 있다. 예를 들어 한국어 화자가 한자 형태소 ‘復(부)’를 어떻게 운영하는가를 알아보기 위해 ‘復(부)’의 분포를 파악해야 한다는 것이다. 한국어에서 ‘復(부)’는 2자어의 앞 글자로만 분포한다. 2자어의 뒤쪽 위치나 3자어의 맨 앞, 뒤에 사용된 예가 발견되지 않으며, 자립적으로 1자어로 사용되지 않고 ‘-하다’ 앞에 나타나지 않는다. 한편 ‘跡(적)’은 2자어의 뒤 위치에만 분포한다. 2자어의 앞 글자 위치, 3자어의 맨 앞이나 뒤에 사용된 예가 발견되지 않으며, 자립적으로 1자어로 사용되지 않고 ‘-하다’ 앞에 나타나지 않는다. 이들은 상당히 제약된 분포를 보이는 유형이다. 반면, ‘惡(악)’은 ‘악하다’와 같이 ‘-하다’ 앞에 분포하여 한 단어로 쓰이고, 1음절어의 자립적 용법도 있으며, ‘악인’, ‘죄악’과 같이 2자어의 선, 후행 위치 분포가 자유로울 뿐 아니라 ‘악감정’, ‘사회악’과 같이 3자어의 양 끝 분포도 자유롭다. 이처럼 한자 형태소는 저마다 다른 분포를 보이며, 이러한 분포의 양상은 한국어에서 한자 형태소

나는 경우, ‘배추’(<백채=白菜)와 같이 음이 바뀐 경우에 해당하는 경우에는 한자 형태소로 인정하지 않았다.

가 운용되는 방식과 밀접한 관계가 있다.

본고의 논의는 위와 같은 인식에서 출발한 시론적 성격을 띤다. 물론 남윤진(1999), 강현화(2001), 강현화·김창구(2001), 이영희(2007) 등과 같이 한국어 한자 형태소의 분포에 대한 연구가 없지는 않으나 이들 논의는 “2자어 3자어 문법”에 기반하지 않았다는 점에서 본고의 논의는 이들 논의와는 다른 접근을 취한다고 할 수 있다. 2자어와 3자어가 그 내부 구조 및 운용과 관련하여 큰 차이를 보이는 단위라는 것은 김창섭(2001), 노명희(2005) 등에서 충분히 논의된 바 있다. 2자어는 대체로 고유어의 문법이 적용되기 어려운 데 비해, 3자어의 경우는 고유어 문법에 의해서도 그 성분 간의 관계가 분석이 된다. 이는 한국어 화자가 한자 형태소를 운용할 때 2자어와 3자어를 별도의 단위로 처리하고 있음을 시사한다. 따라서 이러한 현상을 반영한 분석이 한국어 화자가 한국어 한자 형태소를 운용하는 원리를 밝히는 데 더 유용한 방법인 것으로 판단된다. 본고에서는 한국어 상용한자로 지정된 한자 1,800자²⁾를 분석함으로써 한국어 한자어에 사용되는 한자 형태소의 전반적인 분포 양상을 밝히고자 하였다. 물론 한국어에서 사용되는 한자가 1,800자뿐인 것은 아니다. 그러나 공인된 교육용 상용한자를 통해 일상생활에서 고빈도로 사용되는 한자어에 참여하는 한자 형태소의 전반적 특성을 파악하는 데 큰 무리가 없을 것으로 판단하였다.

2. 한자 형태소 분포 조사의 이론적 기반

본고에서 추구하는 한자 형태소의 분포 조사의 방향은 다음과 같다.

- (1) ㄱ. 한자어의 내부 구조를 고려하여 한자 형태소의 분포를 규정한
다. (1자어, 2자어, 3자어에 참여하는 한자 형태소를 구별하여

2) 해당 목록의 한계와 의의에 대한 논의는 본고 3.1 참조.

처리)

- ㄴ. 한국 한자어의 전반적 특성을 반영할 수 있는 자료를 대상으로 한다. (대상 한자 형태소: 교육용 한자 1,800자, 대상 한자어: 《표준국어대사전》에 수록된 일반 용어)

본고에서 행해지는 자료의 분석은 김창섭(2001, 2013), 노명희(2005)의 “2자어 3자어 문법”에 기반 한다. 김창섭(2013: 168)에서는 “2자어 3자어 문법”에 대하여 다음과 같이 기술하고 있다.

- (2) 현대국어의 한자어 형성에는 이른바 2자어가 기본이 되고 그 앞이나 뒤에 다시 1자가 결합하여 3자어가 이루어진다고 말해져 왔다. 김창섭(2001)에서는 이 생각을 한층 구체화하여 한자어 2자어 3자어론을 제시한 바 있다. 그것은 다음과 같이 요약된다. (Ⅰ) 1자어는 형태론적으로 예외적이며 2자어가 기본적인 적격형식이다. 3자어는 특정한 1자가 1자어나 2자어를 작용영역으로 삼아 그 앞이나 뒤에 결합함으로써 형성된다. (Ⅱ) 고유어 문법은 2자어의 두 자는 의미적 존재로만 분석하고 3자어류의 어기와 1자는 의미 및 문법적 존재로 분석한다.

김창섭(2013)에서는 “2자어 3자어 문법”에 기초하여 한자어를 분석하는 데 다음과 같은 틀을 사용하였다.

- (3) 김창섭(2013)의 한자 형태소 분포 분류

W: 1자 한자어

α: AB나 W 앞의 차사

A: 원사(2자어에 참여하는 요소 중 앞의 것)

B: 원사(2자어에 참여하는 요소 중 뒤의 것)

β: AB나 W 뒤의 차사

김창섭(2013)에서는 위의 W는 단어, A, B는 원사, α , β 는 차사라는 용어를 사용한 바 있다. 본고에서는 이를 받아들여, 이에 더하여 ‘-하다’ 앞에 오는 한자형태소를 R로 따로 분류하여 분포 위치를 6개로 나누고 원사와 차사에 대하여 각각의 위치에 따라 선행, 후행을 구별한 용어를 사용하기로 한다.

(4) 본고에서 사용하는 한자 형태소 분포 분석의 틀

R: ‘-하다’에 선행하는 한자 형태소³⁾

W: 자립적으로 분포하는 한자 형태소⁴⁾

α : 선행 차사

A: 선행 원사

B: 후행 원사

β : 후행 차사

이에 의하면 ‘독서실’은 $AB\beta$ 로, ‘구하다’의 ‘구’는 R로, ‘가건물’은 αAB 로 ‘산’은 W로 분석된다.

위의 한자 형태소 분포 분석의 틀을 기반으로 할 때, 논리적으로 한자 형태소가 보일 수 있는 분포 부류의 가능성은 다음과 같이 63가지⁵⁾ +1(분포 없음)로 총 64가지에 달한다. 아래 기호들은 분포 가능한 위치를 나열한 것이다. 예를 들어 AB는 이에 해당하는 특정 형태소가 A 분포에도 B 분포에도 나타날 수 있음을 의미한다.

3) 편의상 R기호를 사용하고 있으나, 반드시 여기에 해당되는 대상이 ‘단어’가 아닌 ‘어근’일 필요는 없다. ‘-하다’ 앞에 분포하여 용언으로 사용될 수 있는가를 살펴보는 것을 목적으로 따로 나눈 것뿐이므로 본 조사에서는 R과 W가 배타적인 범주는 아니다. 예를 들어 ‘駄(토)’는 R이자 W일 수 있다.

4) 명사뿐 아니라, 관형사, 수사, 의존명사, 부사를 모두 포괄한다.

5) $\frac{6}{1} + \frac{6*5}{1*2} + \frac{6*5*4}{1*2*3} + \frac{6*5*4*3}{1*2*3*4} + \frac{6*5*4*3*2}{1*2*3*4*5} + \frac{6*5*4*3*2*1}{1*2*3*4*5*6} = 63$

(5) 한국어 한자 형태소의 분포 부류 가능성

번호	기호	번호	기호	번호	기호
1	R	22	RW α	43	RWAB
2	W	23	R α A	44	RWB β
3	α	24	RAB	45	R α AB
4	A	25	RB β	46	R α B β
5	B	26	W α A	47	RAB β
6	β	27	WAB	48	W α AB
7	RW	28	WB β	49	W α B β
8	W α	29	α AB	50	WAB β
9	α A	30	α B β	51	α AB β
10	AB	31	AB β	52	RW α B
11	B β	32	RWA	53	RW α B β
12	R α	33	R α B	54	R α A β
13	WA	34	RA β	55	W α A β
14	α B	35	W α B	56	RW α A β
15	A β	36	WA β	57	RW α AB
16	RA	37	α A β	58	RW α B β
17	WB	38	RWB	59	RWAB β
18	α β	39	R α β	60	R α AB β
19	RB	40	W α β	61	W α AB β
20	W β	41	RW β	62	RW α A β
21	R β	42	RW α A	63	RW α AB β

김창섭(2013)에서는 국어 한자 형태소의 부류를 다음과 같이 7가지만을 보인 바 있다. 그러나 김창섭(2013)에서도 이는 일부 예만을 보이는 것이라고 언급하였다.

(6)

부류	예	A,B자 용법	α 자 용법	β 자 용법	W자 용법
A. A/B류	威	威嚴, 國威	x	x	x
B. A/B/ β 류	國	國民 建國	x	強大國	x

c. A/B/β/W류	法	法律 司法	x	(選舉法)	法哲學 選舉法
d. A/B/W류	命	命令	x	x	命
e. A/a류	再	再建	再創造	x	x
f. B류	子	箱子	x	x	x
g. β류	的	x	x	知的 全知的	x

본고의 3장에서는 한국어 한자 형태소 분포 조사 결과 나타난 한자 형태소의 분포 부류를 제시하고, 그것이 의미하는 바를 살펴볼 것이다.

3. 자료 조사의 대상과 방법

3.1. 한자 형태소 목록 선정

앞서 언급했듯이 본고의 목적은 한국어에서 사용되는 한자형태소의 분포를 파악하기 위한 것으로, 교육부에서 발표한 상용한자 1,800자⁶⁾를 대상으로 하였다. 상용한자 1,800자는 지금까지 초, 중, 고등학교의 국어 및 한자 교육의 기준으로 사용되어 왔으나 지나치게 다양한 목적을 포괄하고 있다. 교육부에서 2000년 12월에 발표한 문건(“2001학년도부터 적용되는 한문 교육용 기초 한자 1,800자 이렇게 조정하였습니다”)에 의하면 조정의 기본 원칙이 다음과 같이 제시되어 있다.

- (7) 한문교육과 국어생활을 동시에 고려하고, 동북아 한자 문화권 형성에 능동적으로 대처하여 교육할 수 있는 한자를 선정하였습니다.
○동북아 한자문화권(한국·북한·일본·중국·대만)에서 널리 쓰이는 한자.

6) 하나의 한자에 대해 복수의 음과 훈이 존재하는 경우(‘樂’, ‘更’ 등)는 별도의 한자 형태소인 것으로 처리하였다.

○한문고전에 자주 쓰인 한자.

○국어생활에 자주 쓰이는 한자.

위와 같은 사실에 비추어볼 때 교육부 지정 상용한자 1,800자는 한국어 화자의 한자 형태소 운용 양상을 살펴보기에 가장 적절한 바탕을 제공한다고 보기 어렵다. 특히 ‘한문고전에 자주 쓰인 한자’ 중에는 ‘국어 생활에 자주 쓰이는 한자’와 겹치지 않는 부분이 존재한다.⁷⁾

그럼에도 불구하고 한국어 한자 형태소의 목록으로 선정된 것이 따로 존재하지 않은 상황이기 때문에 한국어 상용한자로 선정된 목록을 우선적으로 연구 대상으로 선정하였다. 추후, 본 연구와 같은 한국어 한자 형태소에 대한 기초 연구들을 통해 국어교육, 외국어로서의 한국어 교육, 한국어 한자어 연구 등, 각각의 목적에 맞는 한국어 한자어 형태소 목록의 선정에 대한 논의가 활발하게 이루어지기를 기대한다.

3.2. 대상 한자어 목록 선정

해당 한자 형태소가 한자어에서 어떤 분포를 보이는가 하는 것을 살펴보기 위해서는 한국 한자어를 관찰해야 한다. 그렇다면 관찰 대상이 될 한자어 목록을 어떻게 정할 것인가 하는 것이 문제가 된다. 본고는 기본적으로 《표준국어대사전》에 수록된 한자어를 대상으로 하되,⁸⁾ 다

7) 이러한 한계가 본고의 조사 결과에 반영된 내용은 본고 4장에서 언급하고 있다.

8) 조사 대상을 『표준국어대사전』으로 한 것에 대해 익명의 심사위원으로부터 각종 어휘빈도조사 자료에서 일정 빈도 이상을 보이는 고빈도 한자어들을 분석 대상으로 삼았다면 전문어 등을 제외하는 작업상의 수고를 덜면서도 더 연구 목적에 부합하는 목록이 나오지 않았을까 하는 아쉬움이 있다는 지적을 받았다. 본고에서는 한국어 한자 형태소의 가장 포괄적인 출현 양상을 파악하기 위하여 조사 대상을 『표준국어대사전』으로 정하였으나 이것이 항상 가장 적절한 자료가 될 수 있는 것은 아니다. 즉 목적에 따라서 심사위원의 지적과 같이 적절하게 여과된 자료를 사용하는 것이 더 효용이 높을 가능성이 높다고 생각된다. 예를 들어 학습자를 위한 한자 형태소 분포 빈도자료를 추출하고자 할 때는 학습자용 사전에 수록된 고빈도 한자어를 대상으로 하는 것과 같은 방법이

음과 같은 점을 고려하여 목록을 선정하였다.

(8) ㄱ. 전문용어를 포함시킬 것인가?

나. 방언과 북한어를 포함시킬 것인가?

다. 고유명사, 성씨를 포함시킬 것인가?

르. 천간에 해당하는 한자를 1음절 단어로 인정할 것인가?

마. 현대국어에서 사용되지 않는 수사를 1음절 단어로 인정할 것인가?

바. 사자성어나 4자 이상의 한문인용구, 한문투의 용법을 포함시킬 것인가?

위의 각각에 대해 본 조사에서는 다음과 같은 기준을 적용하였다.

(9) ㄱ. 전문용어를 원칙적으로 제외하되, 일상용어에 가까운 일부 용어는 조사대상에 포함시켰다.

(예) 활화산(活火山)(지리), 공공선(公共善)(철학), 외국환(外國換)(경제), 위도(緯度)(지리), 해(解)(수학), 가야금(伽倻琴)(음악), 무연탄(無煙炭)(광업), 가면무(假面舞)(예술) 정삼각형(正三角形)(수학), 연비(燃費)(기계) 외 다수.

위에 제시한 자료들은 전문 용어이다. 그러나 ‘활화산’, ‘공공선’ 등과 같은 한자어는 일상용어에 가까운 성격을 띠고 있으므로 논의 대상에 포함시킬 것이다.

또 다음과 같이 《표준국어대사전》에 등재된 범위 내의 다수의 전문 용어 내부에서 규칙적으로 같은 의미를 더하며 차사 용법으로 사용되는 경우에는 차사로 인정하였다.

사용될 수 있을 것이다. 이러한 방식의 조사가 여러 차원에 걸쳐서 이루어질 때, 한국어 한자 형태소의 운용 양상이 다각도로 관찰될 수 있을 것이다.

(예) 의학 전문용어 중, ‘항박테리아’, ‘항결핵’, ‘항말라리아’ 등에 사용된 ‘抗(항)’에 대하여 차사 분포를 인정.

예술 전문용어 중, ‘가면무’, ‘농악무’, ‘사자무’ 등에 사용된 ‘舞(무)’에 대하여 차사 분포를 인정.

나. 방언은 포함시키되, 북한어는 포함시키지 않았다.

다. 인명 등 고유명사와 성씨에 해당하는 것은 원칙적으로 조사대상에 포함시키지 않았지만 고유명사 내부에서 차사 용법으로 사용되는 ‘--山(산), --江(강), --湖(호)’와 같은 경우에는 차사로 인정하였다.

(예) 관악산, 낙동강, 나주호

르. 천간에 해당하는 한자는 《표준국어대사전》에서 단어(1음절어)로 등재⁹⁾되어 있으나 본 조사에서는 단어로 취급하지 않았으나 다음의 ‘甲(갑)’과 같이 일반 용어로서의 용법을 가지고 있는 경우에는 인정하였다.

마. 《표준국어대사전》에 한자 단어로 등재된 단어 중에는 다음과 같이 현대국어에서 사용하지 않는 수사가 20여 개 포함되어 있고 이것이 전문용어로 분류되어 있지 않은데, 본 조사에서는 이를 1음절 단어 목록에 포함시키지 않았다.

(예) 예를 들어 ‘漠(막)’은 “渺(묘)의 10분의 1이 되는 수. 또는 그런 수의. 즉, 10^{-12} 을 이른다”고 하면서도 또 “예전에,

9) 예를 들어 ‘庚(경)’은 다음과 같이 등재되어 있다.

경06 발음 듣기(庚)명사」《민속》

「1」천간(天干)의 일곱째.

「2」=경방03(庚方).

「3」=경시02(庚時).

묘의 억분의 1이 되는 수를 이르던 말. 즉, 10^{-40} 을 이른다”고도 되어 있어 가리키는 수가 명확하지 않을 뿐 아니라 현대국어에서 전혀 사용하지 않는다. 또 ‘埃(애)’ 역시 수사, 관형사로 등재되어 ‘塵(진)’의 10분의 1이 되는 수. 또는 그런 수의. 즉, 10^{-10} 을 이른다’고 하고 또 두 번째 뜻으로 ‘예전에, 진의 억분의 1이 되는 수를 이르던 말. 즉, 10^{-24} 을 이른다.’고 하고 있다. 이러한 일군의 수사는 1 음절 단어로 받아들이지 않았다.

ㄴ. 사자성어, 4자 이상의 한문 인용구에 나타나는 한자는 일반적인 한국어 한자 형태소의 분포와 직접적인 관련이 없는 것으로 보고, 고려하지 않았다.

(예) 연목구어(緣木求魚), 동가식서가숙(東家食西家宿)

또한 표준국어대사전의 뜻풀이에서 한문 투의 용법이나 의고적인 용법, 예전의 용법 등으로 현대국어와 거리가 있음을 명시한 경우에는 이를 대상에서 제외하였다.

(예) 웅(雄)

한문투로 우두머리를 이르는 말.

(예) 유(留)

예전에, 화살이 과녁의 아래를 맞힌 것을 이르던 말

(예) 고(孤)

예전에, 왕이나 제후가 자기를 낮추어 이르던 일인칭 대명사

(예) 사-하다(死하다)

‘죽다’를 한문 투로 이르는 말.

(예) 찬-하다(讚하다)

(예스러운 표현으로) 칭찬하거나 찬양하다.

(예) 무-하다(無-하다)

‘없다’를 예스럽게 이르는 말.

3.3. 분포 인정 기준의 선정

대상 한자어의 목록을 선정하고 나면, 이제 한자 형태소의 분포를 인정할 기준을 선정해야 한다. 이 때 우리는 다음의 선택에 직면하게 된다.

- (10) ㄱ. 해당 한자 형태소가 단 하나의 한자어에만 나타나더라도 해당 위치에 분포하는 것으로 파악한다.
- ㄴ. 해당 한자 형태소가 n개 이상의 한자어에 나타날 때 해당 위치에 분포 가능한 것으로 파악한다.

예를 들어 ‘佳(가)’는 ‘가자제(佳子弟)’ 단 하나의 용례에서만 선행차사(α) 용법을 보인다. 이 경우 (10ㄱ)의 입장을 취하면 ‘佳(가)’에 α용법(선행 차사 용법)이 있다고 볼 것이고, (10ㄴ)의 입장을 취할 경우 ‘佳(가)’에 대하여 α용법(선행 차사 용법)을 제외하여 파악하게 된다.

본고에서는 (10ㄴ)의 입장을 취하여 본고에서 정한 6가지 한자어 분포 위치(R, W, α, A, B, β)에 대하여 해당 한자 형태소가 2회 초과(3회 이상) 나타나는 경우에만 해당 분포 용법을 인정하기로 하였다. 본고에서 한자어 조사 대상의 범위로 결정한 《표준국어대사전》은 완전한 현대국어 사전으로 보기 어려우며 비교적 이전 시기의 어휘를 포함하고 있다. 특히 한자어의 경우 더욱 그러한 경향이 강하여, 1-2개의 어휘에만 나타나는 분포까지 모두 인정할 경우, 현대국어 화자의 한자어에 대한 직관과 먼 결과가 도출될 수 있다. 기준을 특히 ‘2회’로 설정한 것은 어떠한 현상의 누적이 2회 이상인 경우 그것이 단발적/비정착적인 것으로 보기 어렵기 때문이다. 맥락은 다르나, 언어 자료에서 관찰되는 현상이 ‘2회’를 초과하는 것을 우연한/비정착적인 것으로 볼 수 없다는 인

식의 바탕에서 이루어진 논의로 Baayen and Renouf(1996: 75-76)을 들 수 있다. Baayen and Renouf(1996: 75-76)는 타임(Time)지에 나타난 단어들의 유형 빈도와 사전 등재 여부를 비교한 결과 1, 2회 출현 어까지는 사전에 등재되지 않은 신어가 상당한 비율로 관찰되지만, 3회 출현어부터는 대부분이 사전에 등재되지 않은 단어의 비율이 무시할 수 있을 정도로 적음을 보였다. 어떤 언어 단위/현상의 누적이 2회 까지는 단발적/비정착적인 것으로 볼 여지가 있으나 3회 이상인 경우 일정한 경향이 있는 것으로 보고 분석할 필요가 있다는 것이 본고의 입장이다.

다만, 해당 위치에 나타나는 단어가 1개나 2개이더라도 해당 단어가 말뭉치 통계¹⁰⁾에서 100회 이상의 출현 빈도를 보이는 아주 많이 쓰이는 단어일 경우, 이를 고려하여 해당 위치에 대한 분포를 인정하였다.

- (예1) ‘遍(편)’의 B 분포 단어는 ‘보편’ 뿐이지만 ‘보편’의 말뭉치 통계 빈도가 317이므로 B위치를 인정.
- (예2) ‘亦(역)’의 A 분포 단어는 ‘역시’와 ‘역연’으로 2개 뿐이지만 ‘역시’의 빈도가 4855이므로 A위치 인정
- (예3) ‘泊(박)’의 A위치에 해당하는 단어는 2개인데 이에 해당하는 ‘박지(泊地)’, ‘박회(泊懷)’는 빈도 100이상이 아니기 때문에 A위치에 두 단어가 있지만 A위치 분포를 인정하지 않음.
- (예4) ‘幹(간)’의 α위치에 해당하는 단어는 간기술¹¹⁾뿐인데, 이는 빈도 100이상이 아니기 때문에 α위치 분포를 인정하지 않음.

10) 국립국어원에서 제공하는 ‘2010형태분석현대문어말뭉치통계’

11) 간-기술(幹技術)

「명사」

생년월일을 가지고 간지(干支)에 의하여 사람의 운세를 점치는 기술.

3.4. 분포 조사의 실제

본 절에서는 위에서 설정한 기준들을 토대로 실제 한자 형태소의 분포 부류를 어떻게 결정하였는가를 몇 실제 사례를 통해 살펴보기로 한다.

‘姿(자)’에 대한 조사 결과는 다음과 같다.

(11)

W분포	분포 없음
R분포	분포 없음
a분포	분포 없음
A분포	자세, 자태 등 다수의 단어에서 분포가 인정됨
B분포	방자, 교자 등 다수의 단어에서 분포가 인정됨
β분포	부용자 1개 존재

이 결과를 바탕으로 할 때 분포 예를 찾을 수 없는 1자어 분포(W)와 ‘-하다’ 앞의 분포(R), 선행 차사 분포(a)는 제외하였다. 반면 선행원사 분포(A), 후행 원사 분포(B)는 인정하였다. 후행 차사(β) 분포의 경우에는 1개의 예가 존재하지만 본고 3.3에서 기술한 바와 같은 기준에 의해 인정하지 않았다. 만약 ‘부용자’가 고빈도어일 경우 재고의 여지가 있으나 그렇지 않으므로 한국어 한자 형태소로서의 ‘姿(자)’의 분포 부류는 AB부류인 것으로 결정하였다.

‘川(천)’에 대한 조사 결과는 다음과 같다.

(12)

W분포	분포 없음
R분포	분포 없음
a분포	분포 없음

A분포	천렵, 천면 등 다수의 단어에서 분포가 인정됨
B분포	하천, 개천 등 다수의 단어에서 분포가 인정됨
β분포	1) 간혈천, 곡류천 등 다수의 전문어(주로 지리 분야)에서 반복적으로 출현하며 이 때 ‘어떠한 특성을 가진 하천’이라는 의미를 규칙적으로 부가함. 2) 중랑천, 양재천 등 고유명사에서 반복적으로 출현하여 ‘하천’의 의미를 규칙적으로 부가함. 3) 《표준국어 대사전》에 접미사로 등재되어 있음.

이 결과를 바탕으로 할 때 분포 예를 찾을 수 없는 1자어 분포(W)와 ‘-하다’ 앞의 분포(R), 선행 차사 분포(α)는 제외하였다. 반면 선행원사 분포(A), 후행 원사 분포(B)를 인정하였다. 후행 차사(β) 분포의 경우, 그 인정 근거를 세 가지로 나누어 볼 수 있다. 첫째로 다수의 전문어에서 ‘하천’의 의미를 규칙적으로 부가하고 있으며, 둘째로 고유명사에 반복적으로 참여하여 ‘하천’의 의미를 규칙적으로 부가하고 있고 셋째로, 《표준국어대사전》에 접미사로 등재되어 있다. 위 세 가지 중 하나만 해당되어도 후행 차사(β) 인정이 가능한데, 모두 갖추고 있으므로 이를 인정하였다. 위를 근거로 하여 ‘川(천)’은 ABβ분포 부류에 속하는 것으로 판단하였다.

‘湯(탕)’에 대한 조사 결과는 다음과 같다.

(13)

W분포	분포 인정: 명사
R분포	분포 인정: 탕하다
α분포	탕거리 등 다수의 단어에서 분포가 인정됨
A분포	탕국 등 다수의 단어에서 분포가 인정됨
B분포	욕탕 등 다수의 단어에서 분포가 인정됨
β분포	설령탕, 매운탕 등 다수의 단어에서 분포가 인정됨

‘湯(탕)’의 경우 W, α, A, B, β의 분포를 무리 없이 인정할 수 있다. 다만 논의의 여지가 있는 부분은 ‘-하다’ 앞의 분포(R)이다. ‘탕하다’는

‘목욕하다’는 의미로 《표준국어대사전》에 등재되어 있으며 ‘온천에서 탕하다’와 같이 쓰일 수 있다고 되어 있다. 그러나 젊은 현대 한국어 화자는 ‘탕하다’를 많이 사용하지 않는다. 《표준국어대사전》에는 수많은 ‘1음절 한자형태소+하다’ 형의 단어가 등재되어 있다. 그러나 이 중 다수는 현대 한국어에서 거의 사용되지 않는 단어이다. 본고 3장에서는 이러한 사정으로 발생하는 왜곡을 보정하기 위하여 뜻풀이에서 ‘한문 투로 이르는 말’, ‘예스럽게 이르는 말’과 같이 규정한 경우 분포 용례에서 제외하기로 한 바 있다. 그 결과 ‘사하다(死-하다)’, ‘찬하다(讚-하다)’ 등 다수의 ‘1음절 한자 형태소+하다’형 용례가 제외되었다. 그럼에도 불구하고 ‘탕하다’와 같이 제외할 근거가 명시적으로 발견되지 않는 경우에는 연구자 개인의 의견에 따라 자의적으로 제외하는 처리를 지양하고, 원칙대로 분포 용례에 포함시켰다. 위를 근거로 하여 ‘湯(탕)’은 RWαABβ분포 부류에 속하는 것으로 판단하였다.

4. 한자 형태소 분포 조사의 결과

3장에서 제시한 바와 같은 기준과 절차에 의하여 이루어진 한국어 한자 형태소 분포 조사의 결과는 다음과 같다.

- (14) 수학적으로 가능한 한자형태소의 분포부류와 실제로 출현하는 한자형태소의 분포부류의 비교 (음영부분이 실제 출현 부류¹²⁾)

번호	기호
1	R

번호	기호
22	RWα

번호	기호
43	RWAB

12) 분포 부류로 분류된 한자가 2개 이하인 부류는 제외하였다. 위의 음영 처리한 분포 부류는 19가지이며, 이에 ‘분포 없음’ 부류를 더하여 20가지 부류가 실제 자료를 통해 확인되었다.

2	W
3	α
4	A
5	B
6	β
7	RW
8	W α
9	α A
10	AB
11	B β
12	R α
13	WA
14	α B
15	A β
16	RA
17	WB
18	$\alpha\beta$
19	RB
20	W β
21	R β

23	R α A
24	RAB
25	RB β
26	W α A
27	WAB
28	WB β
29	α AB
30	α B β
31	AB β
32	RWA
33	R α B
34	RA β
35	W α B
36	WA β
37	α A β
38	RWB
39	R $\alpha\beta$
40	W $\alpha\beta$
41	RW β
42	RW α A

44	RWB β
45	R α AB
46	R α B β
47	RAB β
48	W α AB
49	W α B β
50	WAB β
51	α AB β
52	RW α B
53	RWA β
54	R α A β
55	W α A β
56	RW α β
57	RW α AB
58	RW α B β
59	RWAB β
60	R α AB β
61	W α AB β
62	RW α A β
63	RW α AB

한국 한자 형태소의 가장 기본적인 용법은 2자어에 참여하는 원사 용법이다. 이는 김창섭(2013)에서 언급된 바 있다. 본고에서는 원사와 다른 분포의 관계를 자료를 통하여 확인한 바, 위의 분포에서 분석되는 다음과 같은 사실에 의해 지지된다.

- (15) ㄱ. 선행 원사(A) 분포 없이 존재하는 선행 차사(α)의 분포는 없다.
 ㄴ. 후행 원사(B) 분포 없이 존재하는 후행 차사(β)의 분포는 없다.

ㄷ. 원사(A or B) 분포 없이 존재하는 1음절 한자어(W)의 분포는 없다.

ㄹ. 원사(A or B) 분포 없이 존재하는 ‘-하다’ 앞(R)의 분포는 없다.

다음의 표(16)은 실제로 출현하는 한자 형태소의 분포 부류의 빈도를 순위에 따라 제시한 것이다. 각 분포 부류의 빈도(백분율)와 함께 이에 해당하는 구체적인 한자 형태소의 예를 함께 수록하였다.

(16) 한자형태소 분포부류의 빈도 순위

부류	갯수	백분율	예 ¹³⁾
AB	685	37.2%	肯(궁): 궁정, 수궁
			拒(거): 거부, 항거
			出(출): 출입, 돌출
ABβ	188	10.3%	群(군): 군중, 일군, 식물군
			錄(록): 녹음, 기록, 회의록
			產(산): 산고, 출산, 중국산
WABβ	184	10.0%	糖(당): 당(명사), 당뇨, 혈당, 포도당
			刑(형): 형(명사), 형벌, 사형, 교수형
			例(예): 예(명사), 예외, 사례, 적용례
WAB	172	9.4%	個(개), -개(의존명사), 개별, 날개
			功(공): 공(명사), 공로, 무공
			便(변): 변(명사), 변비, 소변
RAB	90	4.9%	加(가): 가하다, 가감, 증가
			緊(긴): 긴하다, 긴급, 요긴
			鈍(둔): 둔하다, 둔감, 우둔
WαABβ	77	4.2%	金(금): 금(명사), 금가락지, 금탑, 황금, 지원금
			窓(창): 창(명사), 창머리, 창문, 객창, 검색창
			門(문): 문(명사), 문고리, 문틀, 대문, 개선문
A	68	3.7%	頻(빈): 빈도, 빈출
			郵(우): 우편, 우표
			佳(가): 가인, 가약
αAB	55	3.0%	異(이): 이물질, 이견, 차이
			多(다): 다가구, 다양, 파다
			好(호): 호시절, 호인, 기호

RWAB	49	2.7%	富(부): 부하다, 부(명사), 부자, 거부
			憤(분): 분하다, 분(명사), 분노, 격분
			興(흥): 흥하다, 흥(명사), 흥취, 감흥
WaAB	44	2.4%	近(근): 근(관형사), 근거리, 근경, 최근
			全(전): 전(관형사), 전국민, 전국, 온전
			初(초): 의존명사, 초가을, 초간, 단초
B	41	2.2%	憶(억): 기억, 추억
			況(황): 경황, 상황
			寧(녕): 안녕, 강녕
RWABβ	40	2.2%	占(점): 점하다, 점(명사), 점집, 독점, 그림자점
			曲(곡): 곡하다, 곡(명사), 곡조, 춤곡, 교향곡
			比(비): 비하다, 비(명사), 비교, 대비, 가로세로비,
없음 ¹⁴⁾	28	1.5%	兮 어조사 헤
			孰 누구 숙
			則 곧 즉 ¹⁵⁾
αABβ	19	1.0%	食(식): 식생활, 식초, 채식, 유동식
			原(원): 원자재, 원칙, 기원, 식수원
			醫(의): 의약품, 의료, 명의, 전공의
RWαABβ	18	1.0%	限(한): 한하다, 명사, 한생전, 한계, 제한, 최저한
			毒(독): 독하다, 명사, 독거미, 독기, 중독, 손톱독
			漆(칠): 칠하다, 명사, 칠그릇, 칠기, 덧칠, 기름칠
RaAB	16	0.9%	輕(경): 경하다, 경범죄, 정도, 감경,
			急(급): 급하다, 급정거, 급체, 갈급
			準(준): 준하다, 준회원, 준결승, 준거, 기준
RWαAB	15	0.8%	實(실): 실하다, 실(명사), 실세계, 실현, 현실
			要(요): 요하다, 요(명사), 요주의, 요건, 중요
			貴(귀): 귀하다, 귀(관형사), 귀공자, 귀감, 부귀
WA	10	0.6%	諸(제): 제(관형사), 제가
			每(매): 매(관형사), 매시
			永(영): 영(부사), 영구

αA	8	0.4%	閏(윤): 윤달, 윤선달
			再(재): 재혼, 재결합
			最(최): 최대, 최고가
RABβ	8	0.4%	畢(필): 필하다, 필경, 군필, 검사필
			論(론): 논하다, 논의, 결론, 가격결정론
			記(기): 기하다, 기록, 표기, 여행기
Bβ	7	0.4%	率(율): 비율, 성공률
			芽(아): 발아, 재생아 ¹⁶⁾
			值(치): 가치, 가중치
RA	3	0.2%	質(무): 무하다, 무역
			殉(순)순하다, 순교
			練(련): 연하다(상복을 빨아다듬다), 연복(상복중하나),
RaABβ	2	0.1%	難(난): 난하다, 난문제, 난산, 가난, 식량난
			親(친): 친하다, 친손자, 친족, 선친, 본생친(남아준부모)
			籍(적): 명사, 호적
WB ¹⁷⁾	2	0.1%	泊(박): 의존명사(1박 2박), 숙박
WBβ	2	0.1%	狀(상): 상(명사), 가상, 거울상
			者(자): 자(의존명사), 흑자, 살인자
αB ¹⁸⁾	1	0.1%	見(현): 현구고 알현
RWA	1	0.1%	兼(겸): 겸하다, 겸(의존명사), 겸비
RWaA ¹⁹⁾	1	0.1%	卽(즉): 즉하다(의거하다), 부사, 즉가지, 즉각
W ²⁰⁾	2	0.1%	乙(을): 둘째를 나타내는 명사
			丙(병): 셋째를 나타내는 명사
WaA ²¹⁾	1	0.1%	零(영): 영(명사), 영영눈금, 영점
	1837	100.0%	

- 13) 편의상 분포 부류별로 해당 한자 형태소의 예를 3개씩 보이고 각 분포 위치에 해당하는 한자어를 표시하였다. 그 예가 3개 미만인 부류의 경우에는 전수를 보였다.
- 14) ‘없음’ 유형은 위의 3.1에서 지적한 것처럼 본고에서 조사 대상으로 한 교육부 지정 상용한자 1,800자가 한국어 한자 형태소의 목록과 불일치하기 때문에 나타난 것들이다.
- 15) ‘則(칙)’은 한국 한자어에 사용되지만 ‘則(즉)’은 사용되지 않으며, ‘즉시’와 같은 단어에 나타나는 ‘즉’은 ‘卽(즉)’이다.
- 16) β용법은 식물용어에 반복적으로 등장하며, 해당 식물 용어는 전문용어이지만

가장 많은 한자 형태소가 속한 분포 부류는 단언 AB유형이다. AB유형은 원사로서의 용법만을 지니며 2자어의 앞, 뒤에 자유롭게 나타날 수 있는 유형이다. 이 유형이 한국 한자 형태소의 분포 부류 중 가장 다수이다. ‘肯(긍)’은 A위치에서 ‘긍정’, B위치에서 ‘불긍’과 같은 용례를 보여준다. 1음절어로서의 W용법이나 차사로서의 α , β 용법, 어근으로서의 R용법은 나타나지 않는다. ‘拒(거)’와 ‘出(출)’ 역시 AB분포 부류에

표준국어대사전에 같은 유형의 차사 용법이 사용된 단어가 다수 등재되어 있으므로 본고의 방침에 따라 인정한다.

- 17) ‘籍(적)’은 선행 원사 분포가 가능한 단어가 ‘적산’, ‘적입’ 2개가 있으나 본고에서 2개를 초과하는 별개의 단어에 출현할 때 해당 분포 위치를 인정하고 있어 선행 원사 분포를 인정받지 못한 특수한 예에 해당한다. ‘泊(박)’ 역시 선행 원사 분포가 2단어에서 가능하였으나 본고의 2개 초과 기준을 충족시키지 못하였다. 이 부류에 속하는 한자 형태소 2개 모두 이러한 사정이 있다는 점을 고려할 때, 이 WB분포 부류는 한국어 한자 형태소의 전형적인 분포 부류로 인정되기는 어렵다고 해야 할 것이다.
- 18) 선행원사(A, B)에 현량, 현알 2개 단어의 분포가 발견된다는 점에서 α AB로 볼 여지가 남아 있기는 하지만 역시 2개를 초과하는 별개의 단어에 출현할 때 해당 분포 위치를 인정받을 수 있다는 본고의 기준에 의해 α B로 분류되었다. 그런데 α B의 분포 조합은 본고에서 살펴본 다른 분포 부류의 일반적인 경향과 매우 다른 것이다. 현대 한국어 한자 형태소로서의 ‘見(현)’의 용법은 상당히 제한적이고, 의고적 표현에 한정된다는 점에서 볼 때 α B분포는 한국어 한자 형태소의 전형적인 분포 부류로 인정되기는 어렵다.
- 19) 선행 차사(α) 위치에 한자 형태소 ‘卽(즉)’이 분포하는 단어는 ‘즉기지, 즉출급, 즉기시’의 3단어로 2개를 초과하는 별개의 단어에 출현할 때 해당 분포 위치를 인정한다는 본고의 기준을 ‘간신히’ 통과하였다. 그러나 해당 단어들의 사용 빈도가 극히 낮다는 점을 고려할 때 이 부류 역시 한국어 한자 형태소의 전형적 분포 부류로 인정하기 어렵다.
- 20) 《표준국어대사전》에서는 천간 중에서 ‘甲(갑), 乙(을), 丙(병)’까지는 민속용어가 아니라 일반용어로서 순서를 나타내는 말로 처리하고 있고 ‘丁(정)’ 이후로는 민속 전문 용어로 보고 있다. 그 중 ‘甲(갑)’은 명사로서의 분포 외에 다양한 분포를 보여 WAB β 유형으로 분류되었으나 ‘乙(을), 丙(병)’은 민속전문용어로서 ‘을시’, ‘병일’ 등과 같은 분포만을 보일 뿐, 일반어에 참여하는 것이 발견되지 않아 W 유형으로 분류되었다.
- 21) 후행 원사(B) 위치에 ‘조령’, ‘표령’이 발견되나 본고의 기준에 의해 B위치가 인정되지 않았다. 1~10까지 숫자를 나타내는 한자 형태소는 WaAB나 WAB의 분포 부류에 해당하는데 ‘零(영)’만이 특수한 분포를 보인다.

속하는 한자 형태소이다. 이 중 ‘出(출)’은 한국어 단어에서 출현 빈도가 매우 높은 편이나, AB 분포 외에 기타 위치의 분포를 보이지 않는다. 이는 분포 위치의 범위와 해당 한자 형태소의 사용 빈도가 반드시 일치하지 않음을 보여준다.

두 번째로 많은 형태소가 속한 분포 부류는 AB β 유형이다. AB β 유형은 선행 원사(A) 후행 원사(B)로서의 분포와 더불어 2자어(AB)나 1자어(W) 뒤에 나타나는 후행 차사(β)의 분포를 보이는 유형이다. 예를 들어 ‘産(산)’은 ‘산고’, ‘산통’에서 A분포를, ‘출산, 난산’ 등에서 B분포를, ‘중국산’, ‘자연산’에서 β 분포를 보인다. AB β 유형은 AB유형에 이어 한국어 한자 형태소의 분포에서 두 번째로 많이 나타나는 유형이지만, 이에 비하여 원사(AB) 및 선행 차사(α)의 분포를 동시에 보이는 α AB유형의 경우 약 3%정도로 AB β 유형에 비하여 1/3에 불과하다는 점을 주목할 만하다.

한편 한국어 한자 형태소가 출현할 수 있는 최대의 분포 위치를 보여주는 것은 RW α AB β 부류이다. 이들 부류는 ‘-하다’ 앞에 출현할 수 있고 홀로 1자어로 사용될 수 있으며, 선행, 후행 원사로서 2자어의 앞 뒤에 모두 출현할 수 있고, 선행 후행 차사로서 2자어의 앞, 뒤에 붙어 3자어를 이룰 수 있다. 예를 들어 ‘限(한)’의 경우 ‘한하다’와 같이 ‘-하다’ 앞에 나와 동사 어근으로 사용될 수 있다. 또 명사로서 1자어로 사용되기도 한다. ‘한계’, ‘한정’과 같이 선행 원사로서의 분포를 가지며 ‘제한’, ‘기한’과 같이 후행 원사로서의 분포도 가진다. 또 ‘한평생’과 같이 선행 차사 위치에도 분포하고 ‘최저한’, ‘최고한’과 같이 후행 차사 위치에도 분포한다. 이러한 부류의 한자 형태소들은 본고에서 조사 대상으로 삼은 한자 형태소 중 약 1% 정도에 해당한다. 주목할 만한 것은 분포 가능한 위치가 다양한 한자 형태소가 가장 빈도가 높은 한자 형태소인 것은 아닐 것으로 보인다. 앞에서 살펴본 것처럼 한국어에서 한자 형태소 ‘出(출)’이 ‘漆(칠)’보다 빈도가 높은 형태소일 것이 분명하나, 분포 가능 위치는 ‘漆(칠)’이 ‘出(출)’에 비해 훨씬 넓다.

1음절어(W)로서의 분포를 가진 W-계 한자 형태소의 분포도 주목할 만하다. WAB β 분포는 1자어(W)로서의 분포와 선행 원사(A), 후행 원

사(B)의 분포, 2자어(AB)나 1자어(W) 뒤에 나타나는 후행 차사(β)의 분포를 보이는 유형이다. 예를 들어 ‘糖(당)’은 온전한 1자어 명사로서의 분포를 보일 뿐 아니라 ‘당뇨’, ‘당질’과 같이 선행 원사(A)의 분포, ‘혈당’, ‘과당’과 같이 후행 원사(B)의 분포, ‘포도당’과 같은 후행 차사(β)의 분포를 보인다. 이러한 WAB β 유형은 위에서 살펴본 AB β 유형에 비하여 추후 α 용법을 가지게 될 가능성이 상대적으로 열려있다고 볼 수 있다. 현재 사전에 등재되어 있지는 않지만 “당뇨병 환자가 ‘당조절’을 해야 한다”고 할 때의 ‘당’이 W의 용법인지 α 의 용법인지 명확하지 않을 수 있다. 이에 비하여 AB β 유형에 속하는 ‘錄(록)’의 경우에는 이러한 여지가 상대적으로 적다.²²⁾

역시 1음절어(W)로서의 분포를 가진 W-계 분포 부류 중 많은 한자 형태소가 속해있는 WAB부류는 원사로서 AB위치에 모두 분포 가능함과 동시에 1자어(W)로서의 용법도 가지고 있는 부류이다. 예를 들어 ‘個(개)’의 경우 의존명사로서의 용법을 가지므로 W분포가 인정된다. ‘개체’나 ‘개별’과 같은 단어에서 선행 원사로서의 A분포를, ‘날개’, ‘각개’와 같은 단어에서 후행 원사로서의 B분포를 보여준다. 현재 WAB분포를 보이는 한자 형태소는 쉽게 W α AB나 WAB β , W α AB β 분포 유형으로 변화할 수 있을 것으로 보이는데 특히 1자어(W)가 명사의 용법을 가지는 경우 그러한 경향이 더욱 높을 것으로 판단된다. 예를 들어 ‘功(공)’은 ‘공치사’와 같은 단어에서 α 용법을 보여주기도 하는데, 1개 단어에서만 α 용법에 나타나 본고의 집계에서는 제외되었다. 이는 앞서 살펴본 것처럼 1자어(W)로서의 용법, 특히 명사의 용법으로 인하여 분포의 확장 가능성이 매우 높은 것과 관계가 있다. 따라서 추후에도 우리가

22) 한자 형태소가 무엇인가에 따라 원천적으로 불가능하다고 보기는 어렵다. 한자 형태소의 분포는 개별 화자의 한자에 대한 지식이나 시대의 영향에 의해 상당히 유연하게 나타날 수 있기 때문이다. 그러나 WAB β 분포 유형에 속하는 한자 형태소에 비하여 그 가능성이 매우 적은 것은 사실이다. 이와 관련하여 오규환 학형과 이를 집미 형태소의 특성으로 파악할 수 있을 가능성에 대하여 긴 시간 토론하였는데 그것이 본 논의를 진행하는 데 큰 도움이 되었음을 밝혀둔다.

한자 형태소 분포 부류를 파악하거나, 분포의 변화 양상을 추적할 때는 1자어로서의 분포를 동시에 보이는 W계 분포 부류와 그렇지 않은 한자 형태소의 부류를 분리하여 보아야 할 것이다.

위와 관련하여 WαABβ유형에도 눈길을 줄 필요가 있다. 해당하는 한자 형태소는 1자어(W) 분포와 더불어 선, 후행 원사 분포(AB). 선, 후행 차사 분포(α, β)를 나타낸다. ‘-하다’ 앞에 분포하지 않는다는 점을 제외하면, 본고에서 설정한 분포 가능 위치에 모두 나타나는 것으로 이 부류에 해당하는 한자 형태소가 대체로 본래 명사 용법임을 고려할 때 실질적으로 가능한 최대 분포를 보이는 것으로 볼 수 있다. 이 부류에 속하는 한자 형태소에는 ‘금, 문, 창, 책, 철, 향, 산, 병, 벽’ 등 우리가 대표적 1자 한자어로 인식하고 있는 빈도 높은 1자어들이 포함되어 있다. 이는 고빈도 1자어의 용법이 원사용법으로부터 차사 용법으로의 확대를 촉진시켰을 가능성에 대해 고려할 필요가 있음을 보여준다.

이제 ‘-하다’ 앞에 분포할 수 있는 한자 형태소 부류(R-계 분포 부류)중 가장 빈도가 높은 RAB부류를 살펴보기로 한다. RAB유형은 해당 한자 형태소가 ‘-하다’ 앞에 분포하면서 선행 원사(A)와 후행 원사(B) 분포를 아울러 보이는 유형이다. 이전 시기 한문 투의 글에서 본래 용언의 용법을 가졌던 한자 형태소 중 일부가 현대 한국어에서도 R분포를 보인다. 그러나 이전 시기 한문 투의 글에서 본래 용언의 용법을 가졌다고 해서 현대국어에서도 모두 R분포를 나타내는 것은 아니다. 한자 형태소가 2자어를 이루지 않고 ‘-하다’ 앞에 분포하는 경향은 시간이 흐르면서 감소해 왔음은 여러 차례 지적되어 왔다. 이는 최근 외국어 용언이 유입될 때 ‘핫(hot)하다’와 같이 ‘-하다’ 앞에 분포하는 경향이 강하여 ‘외래어+하다’형 표현이 증가하는 추세와는 반대라고 할 수 있다. 이는 현재 한국어 한자 형태소의 경우 새로 유입되는 형태소가 거의 없는 닫힌 부류라는 점과도 관련이 있는 것으로 판단된다.

다음으로는 A분포, B분포만을 보여 원사로서 선, 후행 요소가 균형적이지 않은 유형을 살펴보기로 한다. 일반적으로 다수의 한자 형태소가 동시에 선행 원사, 후행 원사의 분포를 보이는 데 비해, 선행 원사로

서의 분포만을 보이거나 후행 원사로서의 분포만을 보이는 경우도 있다. 특히 선행 원사 A분포만을 보이는 A부류는 ‘佳(가) 皆(개) 更(갱) 巨(거) 妨(방) 郵(우) 暫(잠) 妥(타)’를 포함하여 전체의 3.7%에 달한다.

한편, 위에서 살펴본 분포 부류 중 RaAB β , WB, WB β , aB, RWA, RWaA, W, WaA부류는 그 해당 한자 형태소의 예가 존재하기는 하지만 2개 이하에 해당하는 소수 유형에 해당한다. 본고에서는 한자 형태소의 분포를 알아보기 위한 한자어의 용례와 분포 부류에 해당하는 한자 형태소의 용례 모두 2개를 초과함을 기준으로 한다.²³⁾ 이는 극소수의 예외적/우연적 용법을 걸러냄으로써 한자 형태소가 가지는 전반적인 분포 특성을 왜곡 없이 파악하기 위함이다.

5. 결론 및 남은 문제: 한자 형태소 분포 조사 결과의 의의와 추가로 고려해야 할 사항

5.1. 결론

본고의 초점은 한국어 한자 형태소의 분포 양상을 살펴보는 데 있으며, 궁극적 목적은 한국어에서 한자 형태소가 운용되는 원리를 밝히고자 하는 것이었다. 한자 형태소의 분포 양상이 직접적으로 한국어 화자의 한자 형태소 운용 원리를 보여주는 것은 아니다. 그러나 한국어 한자 형태소 운용의 현재와 미래를 한자 형태소의 분포를 통해 일부나마 엿볼 수 있다는 것이 본고의 입장이다. ‘再(재)’의 분포를 조사하면 선행 원사(A)와 선행 차사(a)에 제한된 분포가 관찰된다. 이를 통해 한국어 화자가 ‘再(재)’를 포함한 단어를 만들어 쓴다면 ‘再(재)’를 선행 원사(A)나 선행 차사(a)에 배치한 유형의 단어를 가장 거부감 없이 만들어 쓸 가능성이 높을 것으로 예상된다. 이에 비해 ‘再(재)’를 후행 원사(B)나

23) 2회 초과 기준에 대해서는 본고의 3.3에서 언급한 바 있다.

후행 차사(β)에 배치한 단어를 만들어 쓸 가능성은 상당히 낮을 것으로 예측할 수 있다.

이에 더하여 본고의 논의에 추가적인 조사와 수정을 거쳐 교육, 사전, 연구 분야의 자료로써 활용될 수 있을 것으로 기대된다.

5.2. 남은 문제

본고는 한국어 한자 형태소의 분포 양상에 대한 조사의 결과를 기술하는 데 힘을 쏟았으며 이에 대한 본격적 분석에 미치지 못하였다. 따라서 추후 자료에 대한 분석이 이루어질 필요가 있다. 그 과정에서 다음과 같은 점이 더 고려되어야 할 것이다.

첫째, 개인별로 따라 한자형태소의 운용 양상에 차이가 있다는 점을 고려할 필요가 있다. 이는 한자형태소에 대한 지식의 차이/ 특정 한자형태소에 대한 사용 빈도의 차이와 관련이 있다. 이와 관련한 논의로 안소진(2011)을 참조할 수 있다.

둘째, 개별 한자 형태소의 분포 양상은 새로운 단어의 유입과 기존 단어의 소멸 등의 영향으로 항상 변화하고 있다.

셋째, 본 조사에서는 원칙적으로 전문어를 제외하였으나, 전문 영역별로 특히 차사의 목록과 용법에 차이가 있다는 점이 논의될 필요가 있다.

마지막으로 본고는 한자 형태소의 분포를 관찰하기 위한 바탕 자료로 《표준국어대사전》을 절대적으로 참조하였으나, 말뭉치 등으로 그 범위를 적극적으로 넓힐 필요가 있다. 실제로 한국어 화자들의 한자 형태소 운용에 나타나는 유연성을 《표준국어대사전》만으로는 이를 적극적으로 포착하기 어렵기 때문이다. 예를 들어 ‘特(특)’은 《표준국어대사전》만을 참조하여 그 분포 부류를 결정하게 되면 원사의 분포만을 보이는 AB유형으로 분류되는데 이는 실제 현대 한국어 화자의 언어 사용 양상을 정확히 반영하고 있다고 보기 어렵다. 구어에서는 “특 사이즈를 샀다”, “특을 먹을래 보통을 먹을래?”와 같이 원사 용법을 넘어선 용례들이 발견되기 때문이다.

참고문헌

- 강현화(2001), 한국어교육용 기초 한자어에 대한 기초 연구-한국어 교재에 나타난 어휘를 바탕으로, 《한국어 교육》, 12(2), 53-70.
- 강현화·김창구(2001), 연구논문: 어휘력 신장을 위한 기본 한자어의 조어력 조사-한국어 회화 교재에 나타난 한자어를 대상으로. 《외국어로서의 한국어교육》, 25(단일호), 179-201.
- 김지형(2003), 외국인 학습자를 위한 교육용 기본한자의 선정-초·중급 한자를 중심으로, 《어문연구》, 31(2), 377-402.
- 김창섭(1999), 국어 어휘 자료 처리를 위한 한자어의 형태·통사론적 연구, 연구보고서 국립국어원.
- 김창섭(2001), 한자어 형성과 고유어 문법의 제약, 《국어학》 37, 177-195.
- 김창섭(2013), ‘-的’의 두음 경음화와 2자어 3자어론, 《국어학》 68, 167-188.
- 남윤진(1999), 국어사전 표제어의 한자 빈도, 《새국어생활》 9-1, 국립국어원, 49-67.
- 노명희(2005), 《현대국어 한자어 연구》, 태학사.
- 노명희(2006), 한자어 문법 단위와 한자어 교육, 《국어국문학》 142, 465-489.
- 노명희(2014), 한자어 형성과 기능 단위, 《한국어의미학》 43, 159-185.
- 안소진(2011), 심리어휘부에 기반한 한자어 연구, 서울대학교 박사학위논문.
- 안소진(2014), 한자어 형태론의 제 문제와 어휘부, 《한국어학》 62, 373-394.
- 이영희(2007), ‘外國人을 위한 韓國語 學習 辭典’의 漢字 빈도 연구, 《어문연구》 35(4), 455-478.
- Baayen, R. H., & A. Renouf(1996), Chronicling the Times: productive lexical innovations in an English newspaper, *Language*, 69-96.
- [06601 서울시 서초구 서초동 서초중앙로200, 삼풍아파트 13동 1209호]
 전화: 02-6160-4245
 E-mail: panda0705@gmail.com
 투고 일자: 2015. 10. 22.
 심사 일자: 2015. 12. 4.
 게재 확정 일자: 2015. 12. 11.

A Research on the Distribution of Sino-Korean Morphemes

[Ju, Ji-Youn]

한국 한자어의 형태소 분포 조사 [주지연]

We studied the distribution of Sino-Korean morphemes and this research is based on "2Character Word And 3Character Word Theory." Sino-Korean morphemes can appear as an independent word(W), preceding Chasa(affix-like one) character(α), preceding Wonsa(root-like one) character(A), following Wonsa character(B), following Chasa character(β), and preceding element of verb 'ha-'(R). Logically, possible combinations of these 6 categories are 64, but 20 kinds of distribution classes were found in reality. 'AB' type was the class to which the most Sino-korean morphemes belong and 'AB β ' type held the second rank. It is 'RW α AB β ' type which is the maximum distribution of Sino-Korean morpheme.

Key words: Sino-Korean morpheme, Sino-Korean words, distribution, statistics