

제64회 국어국문학회 전국학술대회

포스트휴먼 담론과 국어국문학의 미래

◆ 일시: 2020년 9월 26일(토) 10:00~17:00

◆ 장소: 온라인 학술대회(ZOOM)

◆ 주최: 국어국문학회

◆ 후원: 한국연구재단

"이 발표논문집은 2020년도 정부재원(교육부)으로 한국연구재단의 지원을 받아 발간되었음." "This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government."

제64회 국어국문학회 전국학술대회

"포스트휴머니즘 담론과 국어국문학의 미래"

• 일시: 2020년 9월 26일(토) 10:00~17:00

• 장소: 온라인 학술대회(ZOOM)

• 주최: 국어국문학회

• 후원: 한국연구재단

09:50~10:00	온라인 학회장 등록 및 접수				
10:00~10:05	개회사 박일용(국어국문학회 대표이사, 홍익대)				
	기조강연				
기조강연 10:05~12:30	포스트휴먼 시대, 언어와 소통 발표: 이찬규(중앙대) 포스트휴머니즘의 조건과 담론의 발표: 신상규(이화여대) 기계학습과 국어학 발표: 박진호(서울대) 비판적 포스트휴머니즘과 문학교 - 타자 '되기'의 경험을 중심으. 발표: 우신영(인천대) 포스트휴먼, 바이러스, 취약성(vu 발표: 노대원·황임경(제주대) 포스트휴먼 담론과 고전서사문학 발표: 오세정(충북대) 종합토론	-육의 관계 로 Ilnerability)	- 좌장: 김성룡 (호서대)		
12:30~13:30	휴식				
	기획발표 1분	과(현대문학)			
현대문학 (1차) 13:30~14:50	포스트휴머니즘 시대와 '노년' 전발표: 김윤정(이화여대) 포스트휴먼 시대의 소설 독법: 멀리서 읽기(distant reading)와 발표: 권은(한국교통대) 포스트휴먼의 시대, 페미니스트 - 김보영을 중심으로	토론: 안미영(건국대) 지도그리기(mapping) 토론: 유승환(부산대)	좌장: 연남경 (이화여대)		

	발표: 허윤(부경대)	토론: 안상원(부산외대)	
	지금-여기의 포스트휴먼과 죄책? - 김혜진과 이종산의 작품을 중		
		토론: 김민옥(부산외대)	
14:50~15:00		휴식	
	인문학의 위기와 서정시의 대안적	부 상상력	
	발표: 유성호(한양대)	토론: 권정우(충북대)	
	포스트휴먼 담론과 '인간-이후'의 - 김혜순의 시를 중심으로	· 한국시	
현대문학	발표: 고봉준(경희대)	토론: 박동억(숭실대)	좌장:
(2차) 15:00~16:20	포스트휴먼 시대 시 교육의 역할 - 공감의 원리와 감정 교육으로		엄경희 (숭실대)
	발표: 이경수(중앙대)	토론: 김종훈(고려대)	
	포스트휴먼시대 〈기술편집예술〉	기 주체와 모방의 문제	
	발표: 이용욱(전주대)	토론: 김명훈(포항공대)	
16:20~17:00		연구윤리교육	
	기획발표 2분고	과(고전문학)	
	〈구운몽〉의 주제와 그 구현 양성 - 포스트 휴머니즘과의 관련성을		
	발표: 엄태웅(고려대)	토론: 엄태식(조선대)	
고전문학	고전 소설의 귀신·신선과 인간의	관계로 본 AI 시대의 인간	좌장:
(1차) 13:30~14:50	발표: 신재홍(가천대)	토론: 이종필(대구대)	이지영 (충북대)
	고전소설의 공간정보를 활용한 나 ; 이론적 방법과 실제	배러티브 구축 방안	(0 / 11)
	발표: 신희경(성신여대)	토론: 오윤선(한국교원대)	
14:50~15:00		휴식	
	캐릭터로 읽는 처용 - 포스트휴	먼 담론과 관련하여	
고전문학	발표: 박애경(연세대)	토론: 고정희(서울대)	좌장:
(2차)	고전적 포스트휴먼으로서의 신선	·이인 서사와 그 의미	염은열
15:00~16:20	발표: 신상필(부산대)	토론: 어강석(충북대)	(청주교대)
	포스트 휴먼 시대 인어 신화의 4	스토리텔링 방향	

Ā	발표: 강민경(한양대)	토론: 윤준섭(서울대)		
_	조선 후기 야담에 나타난 청주의			
년 	발표: 정충권(충북대)	토론: 이승은(한림대)		
16:20~17:00	:	연구윤리교육		
기획발표 3분과(국어학)				
5	국어 차용 어휘 분류에 대한 소고	2(2) - 차용 방식을 기준으로		
국어학	발표: 주지연(영남대)	토론: 김경열(충북대)	좌장:	
(1차) 13:30~14:50			이영제 (고려대)	
	발표: 이수진(전남대)	토론: 이금영(충남대)	(== 1 11)	
14:50~15:00		휴식		
	리과학과 국어 연구 : 언어신경학의 연구들에 대한 메티	· 구전 접근과 구문 이론의 모색	좌장:	
	발표: 오충연(숭실대)	토론: 이정훈(서강대)	장충덕	
15:00~16:20	중부방언 방언구획과 네트워크 분석		(충북대)	
Ä.	발표: 정성훈(목포대)	토론: 김한샘(연세대)		
16:20~17:00		연구윤리교육		
	기획발표 4분과	아(어문교육)		
E T	문학교육의 문학공학적 성격과 그	1 함의		
 '부	발표: 류수열(한양대)	토론: 하윤섭(충북대)		
	기지털 시대의 시민성과 리터러시		좌장:	
(1차) 13:30~14:50	발표: 옥현진(이화여대)	토론: 이인화(교육과정평가원)	정진석 (강원대)	
	'새로운 문법 정보를 읽어내는 인간'과 문법 교육			
Ä	발표: 조진수(전남대)	토론: 이관희(서울교대)		
14:50~15:00		휴식		
3	포스트휴먼 시대의 문식성 교육 '	방향 탐색		
어문교육 별	발표: 최숙기(한국교원대)	토론: 이상아(교육과정평가원)	좌장:	
	포스트휴먼 관점으로 본 청소년의 러시 실천을 구성하는 인간-비인점	의 리터러시 실천: 청소년의 리터 간 행위자 탐색을 중심으로	송지언 (홍익대)	
	발표: 김아미(시청자미디어재단)			
	-	연구윤리교육		

개회사

존경하는 국어국문학회원 여러분 안녕하십니까. 학회 대표이사를 맡고 있는 박일용입니다. 먼저 원격으로 학회에 참여해 주신 회원 여러분께 감사드리며, 발제 토론 사회를 맡아주신 여러 회원님들께 감사드립니다. 그리고 국어국문학회원이 아니면서 기조발제를 해주신 포스트휴먼학회 선생들께 감사드립니다.

올해 국어국문학대회는 지난 5월에 개최하기로 예정되었는데, 코로나19 바이러스 창궐 때문에 날짜를 미루다가 어쩔 수없이 원격 회의 형식으로 대회를 개최하게 되었 습니다. 오늘 학술대회의 주제가 "포스트 휴먼 담론과 국어국문학"인데, 뜻하지 않게 학술대회의 주제와 상통하는 학술대회 형식을 가질 수밖에 없게 되었습니다.

"포스트 휴먼"이란 말을 그대로 풀면 "인간" 이후의 "인간"이라는 뜻으로 이해할 수 있겠습니다. 형용 모순처럼 보이는 "인간" 이후의 "인간"이란 말은 지금 우리가 생각하는 인간 개념과 차원이 다른 새로운 개념의 "인간"이란 말로써, 과학과 기술의 발달로 우리가 생각하는 인간의 한계를 넘어서 질적으로 다른 차원의 인간이 탄생될 것이라는 전제 아래 성립된 말이라 할 수 있습니다.

그런데, 그간 "xxx Sapiens"라는 형태의 용어들이 수다하게 제시되었듯이 그간 인류사에서 "인간"의 개념은 끝없이 변화해 왔다고 할 수 있습니다. 기실 인간이 불을 사용하고, 도구를 사용하고, 언어를 사용하고, 문자를 사용하고, 영상매체를 사용하고, 기계문명을 사용하는 과정들이 "인간"의 한계를 넘는 과정으로서, 그 각각의 단계마다 "인간"의 개념 그리고 인간과 자연, 인간과 절대적 진리, 인간과 등 인간과 인간외적 세계와의 관계에 대한 사유 체계는 변화해 왔습니다.

그렇다면 "포스트 휴먼 담론"이 뜻하는 애토의 의미는 앞으로 전개될 과학기술의 발전에 상응하여 도래할 인간의 모습과 세계에 대한 미래적 상상적 담론을 뜻하는 것이겠지만, 의미를 확대하여 현재 우리가 갖고 있는 "인간"에 대한 관념의 틀을 버리고 인간과 세계에 대한 새로운 차원의 성찰을 요구하는 명제라 이해할 수 있을 것입니다. 이렇게 보면 "포스트 휴먼 담론과 국어국문학"이란 주제는 기존의 인간 개념의한계를 넘어서서 새로운 차원의 인간 개념을 매개로 그간에 이룩해온 국어국문학의성과를 재성찰하고, 미래에 지향할 국어국문학의 성과와 과제를 조망해볼 수 있는 주제라 생각합니다.

오늘 대회에서는 6분의 기조 발표와 24분의 분과별 주제 발표를 합하여 총 30분의 발표가 이루어지겠습니다. 그리고 기조발제는 물론 분과 발표까지도 모두 "포스트 휴

먼 담론과 국어국문학"이라는 주제 아래 종합적인 안내와 더불어 구체적인 연구 성과들이 발표될 것입니다. 이렇게 방대한 규모의 발표를 동일 주제로 묶은 학술대회는 접하기가 쉽지 않습니다. 오늘 학술대회에서 진지한 발표와 토론이 이루어져서 국어국문학 연구의 새로운 장을 여는 지식의 향연이 될 수 있기를 바랍니다. 발표자 토론좌장을 맡아주신 여러 선생님들, 참석해주신 회원님들, 그리고 학술대회를 기획하신 운영위원님들께 다시 한 번 감사드립니다.

2020. 9. 26 국어국문학회 대표이사 박일용

➡ 국어국문학회

제64회 국어국문학회 전국학술대회

기획발표 3분과(국어학) 발표문 및 토론문

주지연(영남	대),	국어	차용	어휘 분	분류에	대한 소고	<u>L(2)</u> –	차용	방스	l을 :	기준으로	<u>-</u> ·····	5
		토-	론: 김	경열(충	북대)			•••••	• • • • • • • •	•••••		•••••	····· 22
이수진(전남	⁻ 대),	4 차	산업에	대응하	는 국	어연구방법	법 고찰		•••••	•••••	•••••	•••••	····· 25
		토.	론: 이·	금영(충	남대)	••••••		•••••		• • • • • • • • •	•••••	•••••	····· 35
오충연(숭실	대),												
		구문	이론의	의 모색	•••••	••••••••••	••••••	••••••	•••••	••••••	•••••	•••••	····· 37
		토-	론: 이	정훈(서	강대)	····	•••••	•••••	•••••	•••••		•••••	····· 53
정성훈(목포	대),	중부	방언 병	방언구호	릭과 네	트워크 분	-석	•••••	•••••				····· 57
		토	론: 김 [·]	한샘(연	세대)	•••••		•••••	•••••	•••••		•••••	····· 79

국어 차용 어휘 분류에 대한 소고(2) - 차용 방식을 기준으로 -

주지연(영남대)

1. 서론

본고의 목적은 국어 어휘 어종의 분류 체계를 재검토하고, 차용방법을 중심으로 국어 어휘 차용 방식의 보편성과 특수성을 살펴보는 것이다. 이를 위하여 본고에서는 다른 언어의 차용 어휘 분석 체계를 살펴보고, 국어 어휘 차용 방법을 구성하는 요소를 해체하고 재조합한 틀을 토대로 국어 차용 어휘 자료를 분석할 것이다. 우리는 고유어, 외래어, 한자어, 차용어 등으로 규정되어 온 범주를 해체하여 살펴봄으로써 그동안 모호했던 부분을 포함한 어휘종의 체계를 더 세밀하게 조망할 수 있을 것으로 기대한다.

2. 차용 어휘1) 분류를 위한 기본적 논의

국어 어휘의 어종은 고유어, 한자어, 외래어로 아래와 같이 나누는 것²⁾이 일반적으로 받아들여지고 있다.

- (1) ㄱ. 고유어: 국어 어휘 중 본래부터 국어에 있었던 어휘
 - ㄴ. 외래어: 외래 요소로 들어와서 원어의 발음을 그대로 차용한 어휘.
 - 다. 한자어: 한지로 표기되며 한국 한자음으로 읽히는 어휘

이러한 구분이 논리적으로 가장 적절한 분류 체계라고 보기는 어려워도³⁾ 국어 어휘에 대한 모어 화자들의 직관에 크게 어긋나지는 않는 '구분'에 해당한다. 다만 위와 같은 구분만으로 국어 어휘 어종을 망라하여 설명하는 '체계'가 될 수 없다는 점은 문제다. 일반적으로 우리가 어종을 분류할 때 고유어 외래어, 한자어로 구별하는 것은 1900년대 초반 일본어학계에서 정립한 분류 체계로 국어학계에서는 대체로 이 체제를 유지한 채로 내부의 경계선들을 좀더 명확하게 하고자 하는 시도⁴⁾가 이루어져 왔고 이 체계를 해체하여 본격적으로 검토한 논의는

¹⁾ 어휘의 범위에 대해서는 다양한 논의들이 있다. 우리가 사용하는 '어휘'의 개념은 김광해(1997)에 준하며 기본적으로 단어를 중심으로 하지만, 단어 구성 요소와 관용어 등을 포함한 집합을 가리킨다. .

²⁾ 심재기(2011, 41-42)의 정의를 정리하여 옮김.

³⁾ 이에 대한 논의는 주지연(2018:279-281) 참조.

⁴⁾ 이희승(1941)에서 새 말을 받아들일 당시에는 '외국어'이고, 일상적으로 사용하게 되며 '외래어'로 구분하여 언급한 이래, 임홍빈(1998)에서는 우리말 문맥에서 한글로 적히거나 한글 자모로 대표되는 우리말의 음운으로 발음되면" 외래어라고 구분한 것, 근대 이전에 외국에서 들어온 용어는 차용어로, 근대 이후 외국에서 들어온 수용어는 외래어로 부르는 관습적 경향을 수용한 심재기(1982)의 논의, 정희원(2004)에서 동화의 정도에 따라 동화 과정이 완료된 경우 차용어, 언중이 외래 어휘임을 인식하는 말을 외래어로 구분한 처리 등이 외래어, 차용어, 한자어 사이의 경계를 다듬어 명확화 하려는 시도에 속한다.

찾아보기 어렵다.

한편, 아래의 예들은 다른 언어의 소리, 의미, 구조, 표기 등을 차용한 결과로 만들어진 국어 어휘들로, 위의 고유어, 외래어, 한자어 개념을 통해서는 설명하기 어렵다.

- (2) ㄱ. 과녁, 썰매, 금실, 장난, 재주, 사냥, 과일
 - L. 핸드폰(mobile phone), 체크카드(debit card), 간직(看直), 졸혼(卒婚)/휴혼(休婚)
 - C. 뜨거운 감자, 필름이 끊기다⁵⁾ 회색지대, 품삯, 옷방
 - 리. 호텔신라, 시그니엘부산, 비방사능/비부피/비체적/비질량/비중량
 - ㅁ. 사돈(査頓), 사발(沙鉢)

(2¬)은 이른 시기 유입된 한자어의 발음이 시간이 지나면서 변화하여 한자어를 구성하는 각각의 한자음과 거리가 멀어진 예다. 대체로 이러한 예들은 '고유어화'한 예로 다루어져 왔으나, 이는 국어 어휘 어종을 고유어, 외래어, 한자어의 3칸 구분 체계 안에서 설명하기 위한무리한 처리이기도 하고, 어떤 어휘에 대해 언중이 친근하게 느끼는지 낯설게 느끼는지와 같이 경계가 명확하지 않은 인식 중심의 판단이 개입된 처리라는 점에서 본고에서 지향하는 설명은 아니다. 주지연(2018:285-286)에서는 차용어의 정의 방법을 다음과 같이 나누고, 사건중심 정의가 차용어의 범위를 확정하여 분류 체계를 정립하는 데 더 유리한 면이 있다고 하였다.

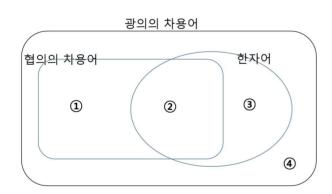
- (3) 기. 인식 중심 정의: 언중이 차용된 사실을 얼마나 '인식'하고 있는가를 고려한 정의로, 언중이 차용 사실을 인식하지 못할 경우 '고유어화'한 것으로 파악한다. 언중의 인식 변화에 따라 고유어와 차용어의 범위가 달라질 수 있다.
 - L. 사건 중심 정의: 차용 여부와 관련하여 확인할 수 있는 사건(언어 요소의 복사, 전이 여부) 자체만을 중심으로 정의. 확인 가능한 차용 사실이 있는 것은 차용어, (아직) 그러한 사실을 발견하지 못한 것을 (잠정적으로) 고유어로 본다.

주지연(2018:285-286)에 의하면 위와 같은 사건 중심 정의의 방법을 택할 경우, 고유어는 우리가 아직 차용의 증거를 발견하지 못한 어휘이고, 차용된 적이 있는 단어가 다시 '고유어화' 한다는 것은 성립할 수 없다. 차용된 단어는 도착어(recipient language)의 체계에 적응 (adaptation)하고 정착(integration)하게 되는 것이지 '고유어'가 되는 것은 아니라는 것이다. 이와 같은 입장6)을 받아들인다면, 우리는 한자어, 고유어, 외래어가 아닌 차용어의 예들도 설명될 수 있는 체계를 마련할 필요가 있으며, 이를 위해 '차용어' 개념을 '고유어'와는 대립되

⁵⁾ 과음에 의한 기억 장애를 의미하는 이 표현이 한국어에 들어온 것은 독일어 표현 "einen Filmriss haben"에서 직역된 것일 가능성이 높기는 하지만, 신견식(2016:28-29)에 의하면, 현재 확인 가능한 자료로는 핀란드어에서 가장 이른 시기에 발견된다고 한다.

⁶⁾ 주지연(2018:283-284)에서는 이러한 관점에서 다음과 같은 도식을 제시하면서 ①은 '버스, 컴퓨터, 커피'등의 외래어류가 대체로 속하는 범주이고 ②는 일본이나 중국으로부터 통째로 들어와서 분석 가능성이 낮은 '공부, 관습, 혁신'등의 2자어류의 한자어가 전형적인 예이며 ③에 속하는 대표적 부류는 '사돈, 복덕방'과 같은 소위 '고유 한자어'류 ④는 이미 차용된 요소들을 합성하여 만든 외래어 단어(핸드폰, 노트북)나 직접적인 언어 형식이 차용된 것은 아니지만, 다른 차용 어휘들로부터 영향을 받아 만들어지거나, 다른 언어의 문법적, 의미적 패턴이 복제된 류가 그 사례가 될 수 있다고 하였다.

고, 외래어와 한자어와 그 밖의 다양한 차용어류를 포함하는 유개념으로 정립할 필요가 있다. (2ㄴ)은 외래어나 한자어처럼 보이지만, 한국에서 만들어진 것으로, 특정한 외국어 표현이 그대로 차용되어 왔다고 보기는 어렵다. (2ㄷ)은 하글로 표기된 국어 어휘이지만, 외국어 표현 의 의미 구조와 형태 구조를 형태소별로 대응시켜 그대로 옮겨왔다는 점에서, 전형적인 고유 어 어휘류와는 그 결이 다르다. (2리)은 한국에서 조합하여 사용하는 표현들이지만, 한국어 단 어 고유의 구조를 가지고 구성된 것이 아니다. 국어의 전형적인 어순에 의하면 '호텔신라'가 아니라 '신라호텔'이라고 할 것이고?), '시그니엘 부산'이 아니라 '부산 시그니엘 호텔'이 라고 할 것이다. 국어의 어순은 수식 받는 대상이 뒤에 나오는 것이 일반적이고, 고유명과 그것이 속한 부류를 동시에 표상할 때는, "삼양라면, 신세계백화점'과 같이 고유명이 앞에 나오는 것 이 일반적이고, 특정 지역명과 지물명의 순서는 '남원 광한루', '서울 광장'처럼 지역명이 선행 하는 것이 일반적이기 때문이다. 따라서 '호텔신라'나 '시그니엘 부산'과 같은 표현의 구조는 국어의 전형적인 구조가 아니라 차용된 구조다. 또 '비방사능/비부피/비체적/비질량/비중량' 의 '비(比)'는 자연과학 및 공학 계 전문용어에 '단위 X에 대한 Y의 비'라는 의미로 마치 접 두사처럼 사용되는데, 이러한 용법은 국어 한자 형태소 비(比)의 가장 전형적인 분포로 보기 어렵다.9) 주지연(2016;389)에서는 의학, 화학 분야의 전문용어에 참여하는 국어 한자 형태소 의 기능을 다루면서 영어 'specific-'이 참여한 전문용어들을 '비(比)-'구성으로 옮길 때 국어 형태소 '비(比)'가 본래는 가지고 있지 않던 분포를 획득하였다고 하였다. '비(比)'는 천 년 전 중국으로부터 차용된 이후 오랜 기간 국어 한자형태소였지만, '비방사능/비부피/비체적/비질 량/비중량'과 같은 단어에서 볼 수 있는 '비X'와 같이 변항을 포함한 구성(construction)의 특정한 '구조'는 근대 이후 차용된 것이다.



위의 도식은 국어 어휘 체계에서 차용어로서 한자어의 범위를 확인하기 위한 특정한 목적으로 만들어진 것이지만 본고의 관심 역시 결국, 차용어라는 유개념 아래, 우리가 지금까지 '한자어, 고유어, 외래어'로 명명해온 전형적이고 한정된 예들을 제외한 어휘들을 단지 설명된 것들의 나머지, 여집합으로 처리하는 것이 아니라 이들을 온전히 분석하고 체계화해 줄 분류 공간이 요구된다는 문제 의식에 있다. 그리고 그분류 공간을 체계화하면서, 만약 필요하다면, 기존의 고유어, 외래어, 한자어로 구별되어 있던 공간을 재배치해볼 수도 있을 것이다.

- 7) 주식회사 공식명도 '신라호텔'이 아니라 '호텔신라'이다 .
- 8) 현재 서울과 부산에 있는 롯데의 호텔 체인으로, 공식명은 '시그니엘 서울', '시그니엘 부산'이다.
- 9) 이러한 양상을 보이는 한자 기반 형태소를 전문용어에서 흔히 볼 수 있다. 천분율을 의미하는 '조출생률, 조사망률, 조혼인률, 조발생률'등의 '조(粗)'는 근대 이전까지 거칠다는 의미로 사용되어 왔지만, 근대 이후 영어 전문용어 중 천분율을 의미하는 'crude-'의 축자적 의미에 그대로 대응하여 사용되면서 천분율과 관련된 새로운 용법을 가지게 된 것으로 보인다. 국어의 한자 기반 형태소 '조-'의 차용 과정 역시 한 출발어로부터, 한 번에 일어난 것이 아니다.

(2ㄴ)~(2ㄹ)의 모든 예들은 출발어¹0)의 언어가 도착어의 언어에 차용되는 과정에서 복사, 전이되는 요소들(소리, 문자, 의미, 구조 등)이 항상 종합적으로 동시에 차용되는 것이 아니며, 각 요소가 때로는 연동되고, 때로는 분리되면서 다양한 방식으로 차용될 수 있음을 보여준다. (2ㅁ)의 예들은 위와 같은 양상이 더 극단적으로 드러난 예다. '사돈'과 '사발'은 중국어나 일본어에서 사용되지 않지만, 우리에게는 '사돈(査頓)', '사발(沙鉢)'과 같이 한자어로 표기될 수 있는 한자어다. 이 단어들은 만주어에서 소리와 의미가 차용된 것으로 알려진 어휘로, 한자는 중국에서 차용된 표기 체계인 한자의 한국 한자음으로 취음 표기한 것이다. 이러한 예들은 소리, 의미 차용과 문자 차용이 심지어 출발어가 다른 사례다. 이러한 사례들은, 우리가 차용 현상을 분석할 때도, 차용되는 요소들을 분석적으로 볼 필요가 있음을 시사한다.

지금까지 살펴본 바를 중심으로 우리는 추후의 언어 차용 논의와 관련하여 참고할 만한 다음과 같은 실마리를 얻을 수 있다.

- (4) ㄱ. 한자어, 고유어, 외래어로 규정된 영역 밖에서 논의될 수밖에 없는 차용 표현들의 예들도 설명될 수 있는 체계를 마련할 필요가 있으며, 이를 위해 '차용 어휘' 개념을 '고유어'와는 대립되고, 외래어와 한자어와 그 밖의 다양한 차용어류를 포함하는 유개념으로 정립할 필요가 있다.
 - L. 차용되는 과정에서 복사, 전이되는 요소들(소리, 문자, 의미, 구조 등)이 항상 종 합적으로 차용되는 것이 아니며, 각 요소가 분리되어 차용될 수 있으므로 우리 가 차용 현상을 분석할 때도, 차용되는 요소들을 분석적으로 볼 필요가 있다.

본고는 위와 같은 실마리를 토대로 차용방식을 중심으로 차용 어휘를 살펴볼 것이다. 그동 안 차용어 분류 논의는 대체로 개별 언어학계에서 각각 개체 차용된 어휘들의 어원 중심으로 수행되었다¹¹⁾. 영어 차용어, 프랑스어 차용어, 중국어 차용어, 일본어 차용어를 나누고 확인하는 작업이 주를 이루면서 차용 방식에 기반한 차용 양상의 다양성은 상대적으로 주목받지 못하였다. 다음 장에서 국어 어휘 차용 방식의 개요를 파악하기 위한 본격적인 논의를 진행하기로 한다.

3. 국어 어휘 차용 방식의 특수성과 보편성

차용 방식을 중심으로 한 차용어 분류의 가장 대표적인 것으로 인용되어 온 체계¹²⁾는 다음 과 같은 베츠(Werner Betz)의 독일어 차용어 분류법이다.

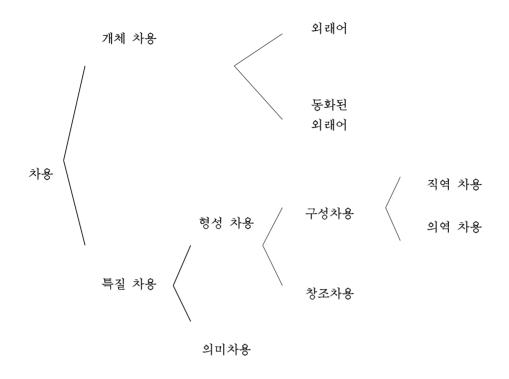
(5) Betz(1974)의 차용 체계¹³⁾

¹⁰⁾ 본고에서는 'doner language, source language'에 해당하는 용어로 '출발어', 'recipient language'에 해당하는 요소로는 '도착어'를 사용하기로 한다.

¹¹⁾ 이러한 경향은, 중국어학계에서도 적용될 수 있는 것으로 보인다. 김태은(2011, 643-644)에서는 중국어 차용어 연구에서 순수 언어학적 고찰은 적고 어원 고증이 많은 수를 차지한다고 지적하였다.

¹²⁾ 장기성(1991)에 의하면, Betz(1936) 이후, 독일어학계에서 언어 차용에 대한 논의가 매우 활발하게 이루어져 왔으며 이는 다른 언어학계의 언어 차용 논의에도 상당한 영향을 끼쳤다. 1990년까지의 독일어 차용어 논의에 대한 설명은 장기성(1991)을 참조.

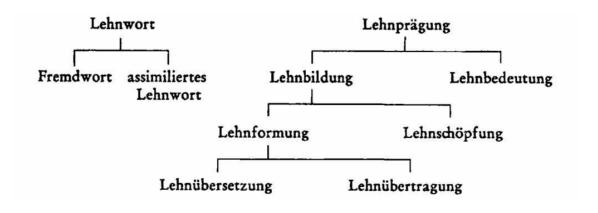
¹³⁾ 용어 번역 필자. 대체로 원문의 용어를 직역하였다. 장기성(1991), 김원(2003) 등 독일어학계의 논의



Betz(1974)에서는 우선 차용을 두 가지로 대분류하였는데, 이 중, 특질 차용(Lehnprägung)은 앞서 언급한 것처럼, 소리와 의미의 쌍으로 구성된 표현 자체가 차용되는 일종의 개체 차용 외의 차용 현상으로 처음으로 제안된 것이라는 데 큰 의의가 있다.

- (6) ¬. 개체 차용: 소리와 의미로 구성된 출발어의 어휘소가 들어온 것으로 도착어에 동화된 정도를 기준으로 외래어와 동화된 차용어로 나눈다.
 - L. 특질 차용: 음성이나 단어 소재 그 자체가 아니라 단어 소재의 형성과 의미, 형 식 내용에 걸쳐 출발어가 도착어어에 미치는 모든 영향과 관련된 차용 이다.

에서는, 본고와 다른 용어로 번역하였다.(내적 차용재, 외적 차용재 등) 원문의 표와 용어는 다음과 같다.(Betz, 1974:137). 이는 Betz(1936) 이후 몇 차례에 걸쳐 수정 제시된 것이다.



이 중 특질 차용의 하위 부류의 특성을 간략히 보이면 다음과 같다.

- (7) 기. 형성 차용: 출발어의 어휘 모델의 영향 아래 도착어의 언어 재료를 사용하여 새로운 단어를 만드는 차용 방식이다.
 - 기-1 구성차용: 출발어의 어휘 모델의 형식을 참조하여 도착어의 형태소를1대1 대응시키는 직역 차용과 다소 유연하게 대응시킨 의역 차용이 있다.
 - 기.-2 창조차용: 출발어의 모델이 단지 도착어의 번역어 형성에 영향을 주었을 뿐, 형식적 모방이 전혀 없는 차용 방식이다.
 - L. 의미차용: 새로운 형태의 단어가 만들어지는 것이 아니라 도착어의 어휘에 이미 존재하는 단어에 다른 의미를 가지게 한 것

그런데 한자권 언어 간의 차용은 위와 같은 서양 언어를 대상으로 한 차용 방식만으로 포착하기 어려운 국면이 광범위하게 발견되는데, 특히 표기(한자)와 관련이 있는 차용 양상이 두드러진다. 이러한 특성은 일본어와 중국어 간의 차용어 연구에서 먼저 지적되었다.

沈国威(1994, 이한섭 역, 2012:.37)에서는 일본어로부터 중국어에 차용된 한자어의 특성을 다음과 같이 제시하면서 '음역차용어=외래어'라는 기준을 적용한다면 일어 차용어의 대부분은 외래어에서 제외된다고 하였다.

- (8) 기. 음을 차용한 일본어 어휘는 그 수가 극히 적다.
 - し. 중국어에 일본어가 들어올 때 일반적으로는 한자 형태가 그대로 들어오게 되며, 手續(수속手続てつづき.), 引渡(인도引ひき渡わたし) (일본어에는 가나가 섞여 있 는 단어) 처럼 가나 표기가 기재된 것은 가나를 삭제하는 과정을 거치게 된다.
 - C. 한자 또는 한자로 환원된 형태로 차용된 일본어는 당연히 재래 중국어와 같이 한자 문자열로 인지되고 발음된다. 즉 일본어 차용어는 다른 외국어 차용처럼 의식적인 음전사 현상이 일어나지 않는다.

그렇다면 우리가 위와 같은 한자권 언어 간 어휘 차용 방식의 특수성을 차용어 체계에서 포착하기 위해서는, Betz(1974) 이하 서양 언어의 언어 차용 이론의 틀을 따르되, "단, 한자어는 문자 차용이다" 한 줄을 추가하기만 하면 되는 것일까? 결론부터 말하면 불가능하다. 동아시아 한자권 언어들에서 한자가 관여하는 어휘들이 모두 같은 차용 방식에 의해, 일시에 일괄차용된 것이 아니기 때문에 그렇게 단순하게 처리하기 어렵다. 특히 한국어의 경우에는 더욱그러하다.14)

우선 한자어가 처음으로 국어 체계에 들어온 시기¹⁵⁾에는 문자와 소리와 의미가 차용되면서 한자어의 의미가 고정되고, 한자음이 형성되었을 것이다. 이후 한국 한자음은 고정된 채로, 문 자와 의미의 쌍이 주요 차용 방식이었을 것이다.

¹⁴⁾ 이와 관련해서는 이 장 후반부에 중국어의 차용어 분류를 살펴보면서 다시 검토하기로 한다.

¹⁵⁾ 이 시기에 대해 다양한 논의가 있으나 이와 관련된 논증은 우리의 주요 관심사가 아니므로 생략하기로 한다.

엄익상(2015:9-10)에서 정리한 중국어 한자음을 기준으로 한 분기를 옮기면 다음과 같다.

(9) ㄱ. 상고음: ?~2세기(동한 말)

ㄴ. 중고음: 3-10세기(위진남북조, 수, 당)

다. 근대음: 10-19세기(송,원,명,청)

ㄹ. 현대음: 20세기~현재

엄익상(2015:17-18)에서는 위의 분기를 토대로, 한자가 기원전 2세기에 전래되었다고 보면 당연히 상고음 시기에 속하지만 현대 한자음을 중국의 고음과 비교하면 상고음보다는 중고음에 가깝다고 보고 있다. 이는 고대 한자음이 중세한자음에 의하여 대체되었기 때문이라는 것이다. 따라서 현대 한자음은 중세한자음에 바탕을 두고 있고, 이 중세한자음은 중국중고음에 바탕하고 있다고 하였다. 이에 의하면, 중국한자음 시기를 기준으로 최소한 상고음 시기에서 중고음 시기까지는 문자(한자)-의미 차용 뿐 아니라, 문자-의미-소리의 차용이 지속적으로 일어나면서 한자음이 형성되고 수정되어 왔다는 것이다.

위의 한자음 시대 분류에 근거해 어휘 차용 방식의 변화를 거칠게 잡아 보면 한자음의 중고음시기까지는 소리-의미-표기 차용이 대규모로 이루어졌고¹⁶⁾, 근대음시기 이후에는 사실상 의미-표기 차용이 대세를 이루면서 더 이상 한자권에서 소리차용 방식이 주류가 되지 못했을 것으로 볼 수 있다.

그러나 중고음을 바탕으로 정착된 한국 한자음이 예외 없이 그대로 남은 것은 아니다. 중국 근대음 시기에 일부 중국 어휘가 문어가 아닌 구어를 매개로 차용되면서, 이전 시기까지 중국 한자음의 변화에 영향을 적게 받고 비교적 고정된 상태로 유지되었던 일부 한국 한자음에 변 화를 일으키기도 하였다. 이러한 양상은 문자-의미 차용이 주를 이루는 상황에서 일부 소리-의미-문자 차용이 이루어지면서 한국에 형성되어 있던 해당 문자와 소리의 조합에 영향을 준 것으로 볼 수 있으나 그 수가 많지는 않았던 것으로 보인다. 예를 들어 중국 한자음의 변화에 영향을 받아서 '탕수육'과 '사탕'의 '糖'의 음이 '당'(중고음)이 아니라 '탕'(근대음)으로 읽히는

사례 등이 있기는 하지만, 현대 한국 한자음의 절대 다수가, 중고음 시기 이후에는 중국 한자음의 변화에 큰 영향을 받지 않았다¹⁷⁾는 점을 부인할 수 없다.¹⁸⁾

¹⁶⁾ 齋藤 希史(2014, 허지향 역, 2018:100-103)에서는 이러한 과정과 관련하여 한자는 표어문자인데(이는 한자가 중국말과 대응하는 측면에서 중국에서도 한자는 표어문자라 할 수 있는 것과 같다) 일본어나 한국어와 같은 언어로 차용된 뒤에는 각각의 고유 언어에 대해 획득한 표어적 기능을 매개로 표음기능을 획득한다는 점을 언급하였다.

¹⁷⁾ 동아시아 한자권의 차용 방식이 일정한 시점 이후로는 사실상 소리 차용과 관계없이 이루어져 왔음 과 관련된 언급으로 황호덕·이상현(2012: 458-460)의 서술도 음미할만하다.

[&]quot;분명 한문에서 근대 한자어까지의 공유의 역사란 '동아시아'를 이야기할 때 반드시 이야기되어야만 할 공동의 원천임에 틀림없다. 하지만 근대라는 만국 체제, 만국어의 질서는 한자를 저고인들의 말씀, 신성한 묵어로부터 떼어내어 저마다 그 소리를 달리하는 민족적 음성 아래 놓이도록 했다. ……물론 그 전에도 한자는 분명히 음성을 가지고 있었다. 하지만 이제 아무도 그 공유성을 지켜내기 위해 홍무정운을 본받는 동국정운을 만들자는 따위의 생각을 하지 않게되었던 것이다 …… 오히려 지금 우리가 음성적 질서 안에서 공유하고 있는 것은 세 나라 밖에 그 원천을 가진 언어들, 즉 외래어인 신어들일지도 모른다."

¹⁸⁾ 박병채(1989), 이돈주(1995)등에서는 이러한 면을 강조하여 한국 한자음이 중고음 시기까지 유입되어 한국어 음운 체계에 수용된 음가를 독자적으로 유지, 발전시켜 왔다고 설명하고 있다.

근대 이후 중국한자어와 일본한자어가 두 언어-중국어와 일본어-의 차이에도 불구하고 사실상 차별 없이 대량 유입되었던 것도, 한국어에 차용될 때는 의미표기 차용 방식(소리가 배제된 방식)을 통했던 점 때문이다.

이와 관련된 일본의 한자음과 관련된 언급으로 沈国威(1994, 이한섭 역, 2012: 47)을 함께 살펴보면 일본어의 경우에도 유사한 상황에 있었던 것을 볼 수 있다.

(10) "일본의 경우 근세 한자어의 수용이 상대 시기와 분명히 다른 일면이 있다. 중세 이후 한자어를 수용할 때에 '오음'에서 '한음'으로 라는 식의 중국어 원음에 대응하기 위한 음성적 모방을 시도하지 않게 되었다.(소수의 당음어는 예외이다) 한자는 일본에서 음훈이 모두 고정되어 일본어의 조어 성분으로 정착되었다."

한편, 한자권 언어 중에서는 중국어학계의 차용 논의에서 차용 방식을 고려한 논의가 비교적 많이 이루어져 온 것으로 보인다. 먼저, 대표적인 중국어의 차용 어휘 논의라고 할 수 있는 Masini(1993)의 분류를 먼저 간략히 살펴보기로 한다. Masini(1993)의 근대 중국어를 대상으로 한 차용어 분류 체계를 간략히 정리해서 제시하면 다음과 같다.¹⁹⁾

- (11) ㄱ. 차음어: 외국어 단어의 의미와 음소 형태를 취하여 해당 언어의 음운 체계에 적 응시킨 것이다.
 - L. 차형어: 외국어 단어의 의미와 문자 형식(writing form)이 들어온 것으로 사실상 근대 이후 일본에서 중국으로 들어온 차용 어휘를 분류하기 위해 만든 항목으로 최초 차형어와 회귀 차형어로 구별한다.
 - L-1: 최초차형어: 일본어 차형어의 분류. 일본어 단어이거나, 이전에 중국 어 서적에 존재했더라도 그 의미가 일본에서 변화한 단어이거나 일 본인이 서양 단어에 대한 차용어로서 만들어낸 것. '幹事'는 본래 중국어에서 일을 처리한다는 뜻이고 명사로 쓰이는 '간사'는 일본어 최초용법이다. '-主意'의 중국어에서 본래 의미는 의를 지킨다는 것 이고 현재 의미는 일본에서 1870년대 생겨났다.
 - L-2: 회귀차형어: 이전의 중국 서적 등에 이미 존재하였고, 여전히 같은 의미로 일본에서 활발하게 사용되다가 그대로 중국에서도 쓰이게 된 것. 예를 들어 敎育, 規則, 意見, 資本, 自由, 進步 등.
 - 다. 차의어: 고유어로서 이미 존재했지만 외국어 모델 어휘에 근거하여 새로운 의미를 가지게 된 것 예를 들어 交際는 맹자에서 '우정'의 의미로 사용되었다가 이후 외부 자극을 통해 새로운 의미 영역을 가지게 되었다, 權利의 원래 의미는 힘과 돈이라는 뜻이었지만, 1864년 만국공법에서 현재 의미로 사용되었다. 保險의 원래 의미는 험준한 곳을 지킨다는 뜻이었다.
 - 리. 차역어: 외국어 모델 어휘의 형태통사론적 구조에 근거하여 만들어진 중국어 단어나 구로, 國債(national debt), 馬力(horse power) 등이 있다.

Masini(1993)에서는 위와 같은 어휘 차용어 체계에 더하여 외국어 단어의 자극을 받아 새로 만들어졌더라도 출발어에 존재하는 특정한 모델 어휘에 근거하지 않고 만들어진 단어로서

¹⁹⁾ 요약 제시 필자.

'고유신어' 범주를 따로 두었다. 신어의 두 유형은 각각 조합신어와 의미신어인데, 조합신어는 새로 글자를 조합²⁰⁾하여 신어를 만든 것이고 의미신어는 기존의 글자가 가진 의미에 (연관이 있는) 새 의미를 첨가²¹⁾하여 만들어내는 것이다.

한편 현대 중국어 차용어에 대한 비교적 최근의 논의로 김태은(2014, 2019)에 정리된 내용을 종합하여 요약, 제시하면 다음과 같다.

(12) ㄱ. 의미 번역

- ¬-1. 의역: 외국에서 새롭게 들여오는 개념이나 문물의 의미를 중국어로 풀어쓴 경우²²⁾
- ¬-2. 차용번역: 출발어 단어를 그 의미와 상관없이 구성 형태소별로 축자적인 번역을 하는 경우²³).
- L. 음역: 외국어 음을 차용한 경우.²⁴⁾
- C. 음의겸역: 원어의 발음 체계를 우선으로 중국어 음역어를 생산할 때 그 한자의 조합이 나타내는 뜻을 고려한 것이다.²⁵⁾
- 리. 형역: 한자는 오랜 시간 동아시아의 공통 문자로 기능해 왔기 때문에, 다른 나라에서 생성된 한자어가 중국어로 들어오는 유형이 이에 속한다.²⁶⁾
- 마. 자모어/자모사: 서양의 라틴 알파벳 자모를 그대로 들여와 사용한다는 점에서 형역의 한 종류로 보아야 한다는 관점이 존재하지만 형역의 경우와는 달리자모어는 발음까지도 외국어의 발음을 그대로 사용하고 있어 차용어 부류로 보지 않기도 한다.²⁷⁾

위와 같은 중국어학계의 차용 어휘 관련 논의는, 국어 차용 어휘 논의에 상당한 시사점을 던져주지만, 아래와 같이 우리의 큰 관심을 받기는 어려운 주제에 논의가 집중된 부분도 있고, 우리가 추가로 고려해야 할 요인들도 있다. 첫째, 중국 차용 어휘 논의에서 상당히 중요하게 다루어지는 최초차용과 회귀차용 구분 논쟁²⁸⁾, 음의겸역 등은 한국어 어휘 차용 논의에서 주요 관심사가 아니며²⁹⁾, 분류 체계의 큰 줄기가 되기 어렵다. 둘째, 중국어의 경우에는, 표기

²⁰⁾ 잡지(雜誌), 화학(化學) 등이 이에 속한다.

²¹⁾ 전기는 본래 중국어로 번개를 뜻하는 電을 사용하여 표상, 洋의 본 의미는 '바다', 이후 외국 물건을 가리키는 기능이 만들어졌다가 현재는 이러한 용법이 거의 소멸하였다.

²²⁾ laser 激光

²³⁾ dark horse 黑马

²⁴⁾ coffee 咖啡 [kha.fei]

²⁵⁾ 김태은(2014:44-53)에서는 "이전에는 이 경우를 한자의 표의 기능을 고려했다는 점에서 의역으로 보는 학자가 있기도 했지만 사실 이 경우 한자들 조합의 뜻은 그 단어 본래의 뜻과 무관한 경우가 대부분이며 우선 발음을 먼저 고려한 후에 한자의 선택에서 뜻을 고려했다는 점에서 음역어의 한 종류로보는 것이 지금의 일반적인 학자들 견해"라고 설명하면서, Coca-cola 可口可乐 kěkǒukělè, TOEFL托福 tuōfú [tʰwofu], Nike 耐克 nàikè [naikʰɣ], 香波(shampoo), 黑客(hacker)등이 그 예라고하였다.

²⁶⁾ 일본 한자어 芸能 げいのう(예능) 이 중국어에 차용되면서 芝能 yìnéng(예능)이 된 예

^{27) &#}x27;HSK'(이 경우는 한어병음 축약형임), 'IBM公司', 'AIDS', 'WTO'

²⁸⁾ 일본제 한자어인지 중국제 한자어인지를 구별하는 데는 조어법 비교, 용자법 비교, 최초 출전 확인 등의 다양한 방법이 동원되어 왔고 양국 언어학계의 언어 차용 논의에서 중요한 내용 중 하나로 다루 어져 왔던 것으로 보인다.

²⁹⁾ 이와 유사한 유형이 아주 없는 것은 아니다. 예를 들어 네이버 서비스 '지식in'은 한자어 '知識人'과 지식+영어 전치사 'in'을 조합한 구성으로 중의적으로 해석될 것을 의도한 것이다. 행사 홍보 어구에 서 흔히 볼 수 있는 "가자go", "떠나자go", "보자go"와 같은 언어 유희 표현들도 발견되는데, 매우

차용으로 들어온 절대 다수의 어휘가 근대 일본 한자어에 해당하기 때문에, 한국어의 그것에 비하면, 문자 차용 양상에 대한 설명이 단순해질 가능성이 있다. 그러나 중국과 일본으로부터 오랜 시간에 걸쳐, 지속적으로, 다양한 방식의 언어 차용이 이루어진 한국의 한자 관련 차용 양상은 한자어는 문자 차용이라고 단순화하기 어렵다. 실제로 용어와 관련해서도 한국어의 차용 현상과 관련해서는 고려할 요인이 더 많다. 중국어학의 일부 논저에서는 '차형어, 차역어 차의어' 같은 용어를 사용하기도 하는데 국어의 경우, 특정 언어로부터 한자를 통한 '차형'(표기 차용)이라는 과정이 성립하기 위해서는 이미 이전에 소리의미표기의 체계적 차용이 이루어진 바 있어야 하고, 차형어가 정말 표기만으로 차용한 것이 아니라 문자(표기)와 의미가 차용된 것이라는 점에서 이러한 개념을 직접 쓰기 어렵다.

이제 우리는 지금까지 다른 언어의 차용 현상를 기반으로 만들어진 차용 어휘 분류 체계와 국어 차용 현상을 견주어 파악한 국어 차용 현상의 보편성과 특수성의 국면들을 참조하여, 차용 방식을 중심으로 차용 어휘를 파악하기 위한 본격적 작업으로서 차용 방식을 구성하는 요소들을 해체하여 살펴볼 것이다. 이는 도착어로서의 한국어에 존재하는 차용 어휘들을 파악하는 데 어떤 틀이 필요할지를 가장 기초적인 단계에서 고민하기 위한 것이다.

4. 차용 방식을 구성하는 요소들의 해체와 조합

국어의 차용 어휘를 이해하기 위한 목적으로, 차용 방식을 구성하는 요소를 해체하면 다음과 같다. 우리는 이들이 어떠한 조합으로 차용 방식을 구성하는지 가능한 모든 조합을 고려하여 살펴보기로 한다.

(13) ㄱ)청각적 기호(소리)-ㄴ) 의미- ㄷ) (언어 형식의)구조, 규칙-ㄹ)시각적 기호(표기)

차용 방식을 구성하는 위의 요소들을 조합한 모든 경우는 다음과 같다.

(14)

①소리	②의미	③구조	④표기
차용	차용	차용	차용
⑤소리·의미	⑥소리·구조	⑦소리·표기	8의미·구조
차용	차용	차용	차용
⑨의미·표기	⑩구조·표기	⑪소리·의미·구조	⑩ 의미·구조·표기
차용	차용	차용	차용
⑬소리·구조·표기 차용	⑭소리·표기·의미 차용	⑮소리·의미·구조·표 기 차용	

제한된 영역에서만 쓰이기 때문에 아직 일반화하기 어렵다.

지금부터는 위에서 목록화한 항목들에 해당하는 사례와 특성을 차례로 살펴보기로 한다.

①소리차용: 유희적 효과나 외국어의 분위기를 내는 효과를 위해 특정 발음을 쓰는 경우가 있으나, 소리만 차용하는 경우를 일반적인 차용의 유형으로 보기 어렵다.

② 의미차용: 소위 의역어라고 불리어 온 유형이 여기 속한다. 출발어에 존재하는 개념이 도착어의 표현 형성의 동기가 되지만, 언어 표현 구조의 모방이 일어나지는 않은 유형이다. 예를 들어 노견(road shoulder)30)를 '갓길'로 차용한 것, red zone을 '청소년 금지 구역'으로 차용한 것 등이 이에 속한다고 할 것이다. 이 차용 유형은 의미구조차용 유형과 혼동되기도 하는데, 의미구조차용의 경우에는 출발어에 있는 언어 표현의 구조(예: hot potato)가 형태소 단위로 대응하여 차용(예: 뜨거운 감자)되면서, 언어 표현의 구조까지 동시에 차용되는 유형에 해당하기 때문에, 출발어에 존재하는 개념(의미)만이 차용되는 의미 차용과는 구별할 필요가 있다.

③ 구조차용: 처음부터 구조만 차용되는 경우는 찾기 어렵다.31) 그러나, 개체 차용이 일어난 뒤에, 특정한 개체들의 구조가 새 어휘의 형성에 영향을 미칠 수 있다. 그래서 일부 언어 표현은, 구조만 차용된 것처럼 보이는 예로 보인다. 따라서 이 유형은 기본적인 차용 유형으로 보기는 어렵고 다른 개체 차용 유형들의 흔적으로서 존재하는 차용 유형으로 볼 수 있다. 예를 들어 '내로남불, 복세편살32)' 같이 사자성어 구조로 만들어진 4자어들, [내로남불, 복세편살…], 일부 상황에서 이름-성 순서로 쓰는 경우 등. 근대 이후 한국 한자음을 가진 한자기반 형태소를 사용하여 한국에서 한자 기반 형태소를 조합하여 만들어 낸 한자어(예: 간직, 감기, 식수, 졸혼, 휴혼)들도 이 유형으로 볼 여지가 있는데, 이미 중국 중고음을 바탕으로 한한국 한자음 정립이 이루어진 뒤에, 국어의 형태소로서 성립한 한자 기반 형태소를 폭넓게 인정할 때 이러한 처리가 가능하다.

④ 표기 차용: 시각적 유희 효과를 위한 디자인적 요소로 차용하는 경우가 있으나³³⁾ 언어 차용이라고 보기 어렵다. 역사적으로는 차자 표기가 표기 차용이라고 할 수 있지만, 현대 국어어휘와 관련된 부분은 미미하다, 단어 단위로 표기만 차용되는 경우는 거의 관찰하기 어렵다.³⁴⁾

⑤ 소리·의미 차용: 가장 주류의 차용 유형으로 지금까지 일반적으로 외래어로 다루어져 온 부류이다. 앞서 살펴본 '개체차용'의 가장 전형적인 유형이라고 할 수 있다. 버스, 컴퓨터, 유

³⁰⁾ road shoulder를 일본어에서 '路脣'으로 번역한 것은 의미구조차용에 속하고, 이것이 한국어에 차용된 것은 의미표기차용, 이후에 우리가 노견'을 우리말 '갓길'로 받아들인 것은 의미차용에 속한다고 할 수 있다.

³¹⁾ 이러한 유형들의 존재는 차용 양상을 평면적으로 파악하기 어려움을 보여준다.

³²⁾ 각각 '내가 하면 로맨스, 남이 하면 불륜', '복잡한 세상 편하게 살자'의 의미이다. 주지연 (2019:68-69)에서는 위와 같은 구성이 굳이 사자성어의 틀을 갖추고 만들어져 사용되는 현상에 대해 "형식성어 [XXXX]_{사자성이}와 같은 틀도 변항을 가진 형식성어"라고 하였다. 이와 같은 관점을 활용하여 설명한다면, 위의 예들은 이미 차용되어 국어에 존재하는 구조 "[XXXX]_{사자성이}에 내용이 채워진 것으로 볼 수 있다.

³³⁾ 유명 가수 브리트니스피어스가 '호남향우회'라는 글자가 디자인 요소로 들어간 옷을 입고 찍은 사진이 화제가 되기도 하였다.

- 커, 쿵푸, 스시 와 같은 예들.
- ⑥ 소리·구조 차용: 해당 사례 아직 못 찾음.
- ⑦ 소리·표기 차용: SK³⁵⁾, CJ³⁶⁾, GS³⁷⁾, SM³⁸⁾,와 같은 기업명, YS, DJ³⁹⁾ 등 주로 정치인 인명에서 흔히 볼 수 있다. 이 유형은 출발어의 특정 어휘 모델을 바탕으로 하지 않고, 의미를 차용한 것도 아니며, 표기와 소리를 차용하는 방식으로, SK를 '선경', YS를 '영삼'으로 읽지 않고 '에스케이', '와이에스'로 읽는다는 점에서 표기만 차용된 경우와 구별된다.
- ⑧ 의미·구조 차용: 주로 'calque' 범주에서 논의되어 온 예들이 이에 속한다. 출발어의 언어 표현이 가지는 의미와 구조가 그대로 들어온 것이다.⁴⁰⁾ 예를 들어 hot potato→뜨거운 감자, gray zone→회색지대, 勞賃→품삯, dress room(옷방), digital signature→전자서명, life cycle→생애주기, leverage effect→지렛대효과, real time → 실시간 등이 있다.
- ⑨ 의미·표기 차용: 이 유형의 차용은, 이미 중국 중고음을 바탕으로 한 한국 한자음 정립이 이루어진 뒤에 성립 가능하다. 들어온 중국 한자어(종결자,홍위병,동북공정, 춘절, 소황제). 일본한자어(변호사, 동산, 부동산, 유치원 등) 등이 이에 속한다.
- * 어휘 차원은 아니지만, 의미표기 차용의 한 양상으로 근대계몽기 한자훈독식 표기41)를 들



- 34) 파편적인 사례들까지 고려할 경우, 모 로펌에서 외국 변호사 등을 고려하여 영문이 섞인 내부 메일을 발신하는 경우, '변호사님'을 관습적으로 'BHSN'으로 표기하는 특수한 예 등이 있지만, 일반적인 국어 사용 맥락의 사례라고 보기 어렵다.
- 35) 예전 선경그룹의 새 기업명으로 1998년부터 선경에서 SK로 바뀌었다.이제 아무도 이 기업을 '선경'이라고 부르지 않고, '에스케이'라고 부르고 있는데, 이는 문자와 해당 문자의 이름을 발음한 소리를 차용하여 고유명을 만든 것이다.
- 36) 제일제당의 새 기업명으로 SK의 경우와 마찬가지로 문자와 해당 문자의 이름을 발음한 소리를 차용 하여 고유명을 만든 것이다.
- 37) 기업명 GS와 관련해서는 LG에서 분리된 그룹이니 GoldStar(금성)의 약자라고 생각되는 경우가 많은데, 실제로는 그룹 차원에서 공식적으로는 무엇의 약자라고 고정하여 밝히지 않고 상황에 따라 다양한 의미를 부여하여 Good Service, Global Spirit, Grow with uS 등으로 말해왔다는 점은 흥미롭다. 소리와 표기만을 차용하여 그룹명을 짓고, 의미는 상황에 따라 붙이는 식이다.
- 38) 이수만의 이름을 딴 것으로 알려져 있다.
- 39) 김영삼, 김대중
- 40) 영어 어휘가 중국어에 직역 차용된 예로는 '下载(download)', '联邦快递(FedEx)', '微软(Microsoft)', '代沟(generation gap)', '马力(horsepower)', '黑板(blackboard)', '黑马(dark horse)', '热线 (hotline)', '热狗(hotdog)', '软件(software)', '硬件(hardware)', '公共关系(public relations)', '超人 (superman)', '超级明星(superstar)', '冷战(cold war)', '情商(emotional quotient)' 등이 있다(김태은, 2011; 660-661), 본고의 분류에 의하면 이 유형은 의미구조 차용에 속한다고 할 수 있다.

수 있다.

- ⑩ 구조·표기 차용: 해당 사례 아직 못 찾음.
- ① 소리·의미·구조 차용: "슈퍼맨, 배트맨, 닥터마틴, 닥터포스터"와 같은 유명 상품/책/영화이름 등이 들어오거나 유사한 구조의 차용이 고빈도로 발생하면서 특정한 구조가 언중에게 각인되면 변항을 포함한 구성이 만들어지기도 한다. 예를 들어 '노실책⁴²) '노양심'의 [노X]나[미스 X⁴³)], [미스터X], [닥터X] [X스Y⁴⁴)], [-맨]같은 구성들이 이에 속한다.
- ② 의미·구조·표기 차용: 특정한 구성과 해당 구성의 의미가 차용되어서 도착어에서 새로운 단어를 만들어내는 데 사용되는 유형으로 '무관중', '무자식'과 같은 어휘를 만들어내는 [무X], '비대면''과 같은 어휘를 '[비X],와 같은 구성들이 이에 속한다. 이 유형의 주요 사례들은 이미중국 중고음을 바탕으로 한 한국 한자음 정립이 이루어진 뒤에 성립 가능하다.
- ③ 소리·구조·표기 차용: 많은 예를 찾기 어렵다. JYP⁴⁵⁾ 와 같은 예가 이에 속하는데 이러한 유형이 추후 증가할 가능성이 있기는 하지만 현재 국어 차용 유형으로 일반화하기 어렵다.
- ④ 소리·표기·의미 차용: 한문, 한자어 유입 초기의 개체 차용이 여기 속한다고 할 수 있다. 소리-표기-의미가 차용되면서 한자음-문자(한자)-의미가 고정되기 시작하였을 것이다.⁴⁶⁾

⁴¹⁾ 훈독식 표기에도 여러 가지 유형이 있으나, 김병문(2014)에서 제시한 유길준의 "노동야학본"의 훈독식 용언 표기를 일부 옮기면 다음과 같다.

入 ^들 어〈36 ⁴⁰⁾ 〉	入 ^들 기를〈77〉	入 ^들 나〈77〉
成 ^{니루} 나니〈30〉	成 ^{니루} 는〈30〉	成 ^{이루} 지〈63〉
守 ^{직히} 여야< (4)	守 ^{작하} 기눈〈5〉	守 ^{덕히} 고〈67〉
死 ^퓩 드라도〈16〉	死命을〈75〉	死命는〈77〉
遊 ^上 지〈8〉	遊 ^골 기만〈24〉	遊 ^音 고〈34〉
生 ⁴ 는〈8〉	生 ^살 기를〈14〉	生 ^사 니보다〈75〉
生나매(4)	生 ^{나이} 매(9)	生 ^날 졔마다〈16〉
立 ^셔 셔〈17〉	立 ^셧 시되〈76〉	立 ^特 지(84)
立 ^세 울지오〈14〉	立세우시니〈16〉	立 ^세 움이〈69〉
見 ^보 시오〈23〉	見 ^보 지 마시고〈43〉	見 ^{보이} 는〈77〉
爭 ^{다토} 는〈74〉	爭 ^{다툼} 이〈83〉	爭 ^{叶토} ol<85〉
謂 ^{일格} 이뇨〈1〉	謂 ^{일음} 이니라〈14〉	謂 ^{일을} 지니라〈53〉
活 ^小 는〈1〉	活 ^ル 을〈25〉	活些라고〈44〉
失 ^일 헛다는〈46〉	失 ^{을 흐} 면〈68〉	失 ^일 치〈79〉

- 42) "박찬호 선수는 3회까지 노실책입니다"
- 43) X는 성명 중 성
- 44) X는 성, Y는 주로 상점 이름이다. 소유를 나타내는 구성(출발어에서는 어퍼스트로피 기호를 사용하여 "'s"로 표기)의 구조가 들어온 것이다. 예를 들어 '킴스피부과, 킴스약국, 강스뷰티, 최스키친' 등.
- 45) 박진영의 이름을 딴 것으로 알려져 있는데, '진영 박'과 같은 이름-성 표상 구조를 차용하고 'JYP'로 적고 '제이와이피'라고 읽는다는 점에서, 소리·구조·표기 차용의 예라고 할 수 있다.
- 46) 齋藤 希史(2014, 허지향 역, 2018: 2018:100-101)에서는 이와 관련하여 "(한자에 대한) 현지어에 의한 해석이 한 글자 한 글자에 대해 고정되면 이것이 의미가 되는 동시에 현지어로 읽을 때의 소리가 된다"고 하였는데, 한자 학습서 등에 나타나는 '하늘 천' '검을 현'과 같은 고정된 훈음이 이러한 예해당한다.

⑤소리·의미·구조·표기 차용: 한국 문장에서 원어를 그대로 노출시켜 표기하는 것(예: 'happy 와 용법으로 사용되어 원어의 표기로 나타나는 경우⁴⁷⁾), 중국어의 경우에는 일부 자모사⁴⁸⁾ 등이 이러한 유형에 속한다.

위의 논의를 종합하면 언어학적 관점에서 특히 현대 한국어 차용 어휘 논의를 전제로 할 때 주요 유형으로 다룰 가능성이 있는 영역은 현재로서는 아래의 음영 영역의 10개 유형 정도에 해당한다고 여겨진다. 단, 아래 차용 방식들의 조합 목록이 동일한 층위의 차용 과정에서 추출된 것은 아니다.

(15)

①소리	②의미	③구조	④표기
차용	차용	차용	차용
⑤소리의미	⑥소리구조	⑦소리표기	®의미구조
차용	차용	차용	차용
⑨의미표기	⑩구조표기	⑪소리의미구조	⑩ 의미구조표기
차용	차용	차용	차용
③소리구조표기	⑭소리표기의미	⑤소리의미구조표기	
차용	차용	차용	

5. 차용 방식을 중심으로 한 (평면적)분류의 의의와 과제

본고에서 차용 어휘를 파악할 수 있는 몇 관점을 검토하고, 차용 방식을 구성하는 요소들을 해체하여 살펴본 것은 도착어로서의 한국어에 존재하는 차용 어휘들을 파악하는 데 어떤 틀이 필요할지를 가장 기초적인 단계에서 고민하기 위한 것이다. 그 중간 결과물로서 차용 방식을 중심으로 한 평면적인 분류 체계를 살펴보았는데 이를 통해 다음과 같은 양상을 이해하는 데

'돌아온 游客' 작년 항공 여객 1억1753만명, 전년比 7.5%↑...국제 여객 이 주도



생일 축하합니다! 'Happy 김혜수's Day~'[포토엔HD화보]

48) 중국어의 자모사에 대해서는, 중국어 범주에서 논의할 것인가, 외국어 표현인 것으로 볼 것인가를 양 극단으로 한 다양한 스펙트럼의 논의가 있다. 우리도 이 유형에 속하는 경우를 어떻게 볼 것인가에 따라 이 범주를 차용에서 제외할 수도 있다. 여기서는 명확한 결론을 내리지 않고 국어 문장 속에서 사용되는 모든 범위를 포괄하여 다룬다.

⁴⁷⁾ 다음은 일간지 제목에서 사용된 사례다.

도움이 되었다.

- (16) ㄱ. 차용 방식의 다양한 양상을 더 세밀하게 이해할 수 있다.
 - 모든 언어의 차용 어휘를 설명하는 데 유사한 틀이 유효한 것이 아니라는 점을 확인하고 국어 차용 현상의 보편성과 특수성에 대한 이해의 실마리를 얻을 수 있다.

그러나 위와 같은 평면적인 관점은 다음과 같은 한계를 가진다.

- (17) ㄱ. 모든 차용 방식의 층위가 같지 않고 특정 어휘가 만들어지는 과정에서 복수 유형의 차용 방식이 시간 차를 두고 관여될 수 있다. (49) 모든 차용 현상이 같은 방식으로, 같은 출발어로부터, 한 번에 일어나고 마는 것이 아니고, 출발어와 도착어의 변화하는 언어 상황 속에서 끊임없이 다양한 방식과 수량과 빈도로 발생한다는 점을 고려해야 한다.
 - L. 단어 이상 단위의 차용과 형태소 단위 이하 차원의 차용을 구별하여 볼 필요도 있는데50) 위와 같은 틀만을 가지고는 어렵다.

⁴⁹⁾ 예를 들어 '부동산, 유치원'과 같은 단어가 일본어로부터 차용되고, '종결자, 소황제'와 같은 단어가 중국어로부터 차용된 것은 해당 단어의 의미와 표기 조합의 차용이 일어난 것으로 볼 수 있는데, 이러한 유형의 차용이 일어나기 위해서는 한국 한자음 정립이 이미 이루어진 상황이라는 전제가 필요하다. 한국 한자음의 형성은, 더 이른 시기 중국으로부터 광범위한 대량 차용이 발생하면서 한자어들을 구성하는 각 한자의 음과 뜻이 국어의 요소로 자리 잡았기 때문에 가능한 것이다.

⁵⁰⁾ 근대 중국 어휘와 일본 어휘 차용과 관련해서도 차자적(형태소) 차용과 차어적(단어) 차용의 차원을 구별하는 관점이 있다. 예를 들어 沈国威(1994, 이하섭 역, 2012:.47-48)은 중일 간의 어휘 교류를 번역어 차용 관계가 아니라 일종의 차용자의 관계로 보는 견해를 소개하였다. 일본의 경우 근세 한자 어의 수용이 상대 시기와 분명히 다른 일면이 있는데 중세 이후 한자어를 수용할 때에 오음에서 한음 으로 라는 식의 중국어 원음에 대응하기 위한 음성적 모방을 시도하지 않게 되었다는 것이다. 즉 한 자는 일본에서 음훈이 모두 고정되어 일본어의 조어 성분으로 정착되었고 알지 못하는 한자어의 경 우 각 글자의 훈에서 음으로 라는 과정을 거쳐 단어 전체의 의미가 파악된다. 이러한 의미에서 차자 적 요소가 커져왔다고 할 수 있다는 것이다. 한편 중국인이 일본어 어휘를 받아들일 때는 중국인의 한자에 대한 전통적인 이해가 일본어의 수용에 어느 정도인지는 명백하지 않지만 영향을 미쳤다고 보 고 있다. 한자어의 차용과 한국어 한자 기반 형태소의 형성과 관련해서는 주지연(2018)에서 자세히 다루고 있다. 주지연(2018:298-300)에서는 "다른 언어와의 접촉에 의해 차용된 단어는 국어 안에서 사용되기 위해 적응(adaptation) 단계를 거치며, 그 중 일부는 국어 단어의 일부로 정착(integration) 한다. 그런데 차용, 적응, 정착 과정에서 동일한 출발어로부터 대량의 어휘가 광범위하게 유입될 경 우, 대량의 차용어군과 그것의 운용 원리가 해당 언어 체계에 정착하기도 한다"고 설명하면서 이를 특정 단어가 개별적으로 차용, 정착하는 현상과 구별하여 체계적 차용, 체계적 정착 현상으로 다루었 다. 즉 국어의 한자어는 광범위한 대량의 차용 과정에서 한국식 한자음이 확립되어 한국 한자어에 사 용되는 한자의 독음에 혼란이 없고 한자로 표기 가능한 새로운 단어가 차용될 경우, 출발어의 음운론 적 특성과 관계없이 확립된 한국 한자음으로 독음한 것이 어휘부에 수용되며, 한자어 구성 요소(한자 형태소)가 분석되고, 일부 구성 요소는 국어 체계에서 생산적인 단어 형성 요소로 발달하기도 하고, 한자어 형성 원리가 국어 형태론적 원리로서 정립되었다는 것이다. 이는 한자권 언어 간의 한자(어) 차용 관계에 대해 차자인가 차어인가를 평면적인 대립항으로 두는 것이 아니라 차용 과정의 특성을 고려하여 볼 때 광범위한 대량의 한자 어휘 차용이 결과적으로 그 내부 구조에 대한 분석을 촉진하여 한국 한자 형태소 발생으로 이어지면서 결과적으로 차자 양상과 유사한 결과에 이른 것으로 보는 입 장이다.

결론적으로 우리 논의에서는 차용 방식의 구성요소를 해체하여 살펴보는 과정에서 몇 가지 국어 차용 현상과 관련된 중요한 실마리들을 포착할 수 있었지만, 언어 차용을 이해하기 위해 고려해야 할 층위들이 더 존재하기 때문에, 위와 같은 평면적 차용 방식 체계가 차용과 관련된 내용을 본격적으로 이해하는 최종적인 틀이 되기는 어렵다. 특히 차용 어휘들이 하나의 출발어에서만, 직접, 단일한 방식으로 들어오는 것이 아니라 여러 경로를 거쳐(특히 한국어의 경우 중역이 근대화 과정의 문화 수입에 상당히 큰 비중을 차지하였다는 점을 고려하면 더욱 그렇다) 들어왔다는 점을 고려할 필요가 있는데 그러한 점은 위와 같은 방식을 중심으로 한 평면적 틀에서 충분히 다루기 어렵다. 이는 차용 어휘를 충분히 이해하기 위한 접근의 경로가하나로 충분하지 않으며 다양한 관점에서 입체적으로 차용 어휘를 정의하고 분류하는 지속적인 시도가 필요함을 보여준다.

참고문헌

- 김광해(1997), 국어어휘론의 지평,
- 김춘월(2016), 동일 어원 차용어들의 의미 차이에 대하여. 한국학 (구 정신문화연구), 39(2), 129-152.
- 김태은(2011), 어휘 계층화 (Lexical stratification) 에 따른 중국어 차용어의 범위에 대한 고 찰. 중국어문학지, 37, 641-670.
- 김태은(2014), 중국어의 외래어 수용과 대응: 음역어를 통해 본 중국의 언어와 문화. 인문과학 연구논총, (39), 39-73.
- 김태은(2019), 중국의 외국어 수용과 중국화에 대한 논의-차용어 적응 방식에 대한 고찰을 중심으로. 中國學, 66, 1-20.
- 박병채(1989), 국어발달사, 세영사.
- 신견식(2016), 콩글리시 찬가, 뿌리와이파리.
- 심재기(1982), 國語語彙論, 집문당.
- 심재기(1989), 한자어 수용에 관한 통시적 연구, 국어학18, 89-109.
- 심재기, 조항범, 문금현, 조남호, 노명희, & 이선영(2011), 국어 어휘론 개설, 지식과 교양.
- 엄익상(2015), 한국한자음, 중국식으로 다시보기, 한국문화사.
- 이돈주(1995), 한자음운학의 이해, 탑출판사.
- 이희승(1941), 외래어 이야기. 춘추, 2(3), 163-193.
- 임홍빈(1998), 국어문법의심층: 어휘범주의통사와의미. 태학사.
- 장기성(1991), 어학: 언어차용에 관한 연구 (2)-그 유형을 중심으로. 어문학연구, 4, 27-52.
- 정희원(2004), 외래어의 개념과 범위. 새국어생활, 14(2), 5-22.
- 조세용(1999), 한자어계 차용어의 개주· 귀화 현상에 관한 연구-15 세기 중기 이후의 귀화어를 중심으로. 한글, (243), 83-108.
- 주지연(2016), 전문용어 (專門用語) 에 사용된 한자 (漢字) 형태소 (形態素) 의 분포 (分布) 와 전문 분야별 한자 (漢字) 교육 (敎育)-화학 (化學), 의학 (醫學) 전문용어 (專門用語) 에 사용된 차사를 중심으로. 어문연구, 44(1), 373-401.
- 주지연(2018), 국어 차용어의 분류에 대한 소고: 한자어의 지위를 중심으로. 언어와 정보사회,

34. 277-303.

- 주지연(2019), 한국어 한자성어 변이 구성 연구. 어문연구 (語文研究), 47(1), 45-74.
- 황호덕·이상현 편(2012), 『개념과 역사, 근대 한국의 이중어사전(영인편)』, 박문사.
- 沈国威(1994), 近代日中語彙交流史: 新漢語の生成と受容 (Vol. 271), 笠間書院. 번역 이한섭 (2012), 근대중일어휘교류사, 이한섭역, 고려대학교출판부.
- 齋藤 希史(2014), 漢字世界の地平: 私たちにとって文字とは何か (新潮選書) (單行本), 허지향역(2018), 한자권의 성립, 글항아리.
- Betz, W(1974), Lehnwörter und Lehnprägungen im Vor-und Frühdeutschen. de Gruyter.
- Haspelmath, M., & Tadmor, U(Eds.)(2009), Loanwords in the world's languages: a comparative handbook. Walter de Gruyter.
- Masini, F(1993), The formation of modern Chinese lexicon and its evolution toward a national language: the period from 1840 to 1898. Journal of Chinese Linguistics monograph series, (6), i-295. 이정재 역(2005), "근대중국의 언어와 역사", 소명출판.

"국어 차용 어휘 분류에 대한 소고(2) - 차용 방식을 기준으로"에 대한 토론문

김경열(충북대)

안녕하십니까.

이번 토론을 맡은 충북대학교 김경열입니다.

먼저 선생님의 논문 잘 보았습니다. 제가 이 분야에 전공자가 아니어서 깊이 있는 토론이되지 못한 점 죄송하다는 말씀을 드립니다. 또한 선생님의 논문을 읽으면서 그동안 공부에 소홀했던 제게 공부할 수 있는 기회를 주신 것에 감사드립니다. 이에 이 논문을 읽으면서 궁금했던 점을 질문하는 것으로 토론을 대신하고자 합니다.

- 1. 표준국어대사전에서 '차용어'는 '외국에서 들어온 말로 국어에서 널리 쓰이는 단어'로 정의하고 있습니다. 그러나 이 논문에서는 차용어의 범위를 넓게 보고 있는 듯 합니다. 필자께서는 국어 어휘의 세 가지 분류, 즉 고유어, 외래어, 한자어로 분류할 수 없는 것으로 (2)의예문을 제시하고 있는데, (2ㄱ)은 (3ㄴ)의 정의에 따르면 차용어로 볼 수 있습니다. 그런데 (2ㄴ, ㄷ, ㄹ)에서 제시한 예들은 어휘가 차용된 이후 조어법이나 번역상의 문제로 볼 수 있을 듯합니다. 이들을 차용어로 처리하는 것은 어떤 장점이 있는지 궁금합니다.
- 2. 차용 어휘 분류에 있어서 차용 방식을 기준으로 살펴본다고 하셨는데 '호텔신라'는 그전체가 차용된 것이 아니라 '호텔'과 '신라'가 각각 차용된 이후 국어의 일반적인 조어법과 다르게 실현된 것이고, '뜨거운 감자'는 'Hot Potato'를 차용한 것이라고 하기보다는 그것을 번역한 것이라고 할 수 있습니다. 그렇다면 이들을 차용어로 볼 수 있는지 궁금합니다. 즉, 전자는 어휘가 차용된 이후 국어화되는 과정에서 조어법에 따라 분류한 것이고, 후자는 번역을 어떻게 하느냐에 따른 것이라고 할 수 있을 듯합니다.
- 3. 4장의 제목을 '차용 방식을 구성하는 요소들의 해체와 조합'이라고 하셨는데, 위 첫 번째 와 두 번째 질문에 따르면 '차용 방식을 구성하는 요소'가 아니라 '차용 요소'들이라고 해야 하는 것이 아닐까 합니다. 즉, 어떠한 어휘가 차용될 때 '소리, 의미' 등의 요소가 차용된 이후 그 요소들이 결합하는 방식이 다른 것이 아닐까 합니다. 이는 차용 방식의 문제가 아니라 차용된 어휘가 국어화되는 과정에서 어떻게 조어되는지에 대한 것이라고 생각이 됩니다.
- 4. 마지막으로 세부적인 몇 가지를 더 살펴보면 4장 ②에서 '노견을 차용한 것'이라고 하셨는데, '차용'보다는 '번역(의역)'으로 볼 수는 없는지 궁금합니다. 또한 ③에서 '한국에서 한자기반 형태소를 조합하여 만들어 낸 한자어'라고 한 부분을 보면 이는 한자어 차용 이후 조어법과 관련된 문제로 보입니다. ⑦에서도 'SK, CJ' 등을 차용 방식이라고 보기보다는 국어에서 축약어를 만드는 조어법의 관점에서 논의할 수는 없는지 궁금합니다.
 - 5. 위와 같은 궁금증을 종합해 보면 이 논문에서 차용 어휘 분류를 위해 기준으로 '차용 방

식'을 제시하고 있는데 이보다는 어휘가 국어에 차용된 이후 국어의 조어법이나 신어 형성, 번역 등에 따라 차용어를 분류하는 것이라고 생각이 됩니다. 이는 원론적인 의문에 해당하는 것이라고 할 수 있는데, 차용어의 범위를 어디까지로 보아야 하는지가 먼저 정의된 이후, 차 용어 분류 기준을 제시하고, 나아가 그 기준에 따라 분류하는 것이 좋을 듯합니다. 이 점에 대해서 필자께서는 어떻게 생각하고 계시는지 궁금합니다.

4차산업에 대응하는 국어연구방법 고찰

이수진(전남대)

- 1. 서론
- 2. 4차 산업혁명과 언어
- 2.1. 4차 산업혁명의 언어학적 의의
- 2.2. 딥러닝과 언어처리
- 3. 국어의 언어정보화의 현황과 연구방법
- 3.1. 새로운 말뭉치 구축
- 3.2. 지역어 말뭉치 구축
- 3.3. 특수목적 말뭉치 구축
- 3.4. 인공지능과 국어연구방법
- 4. 결론

1. 서론

4차 산업혁명이 점차 일상생활에 변화를 일으키고 있다. 이 글은 4차 산업혁명을 맞아 대규모로 이루어지고 있는 말뭉치 구축 및 국어연구방법을 고찰하는 데에 목적이 있다. 현재 국립국어원에서는 한국어 정보화 사업 재개로 말뭉치 빅데이터를 구축하고 있으며, 국어 연구자들도 여러분석 프로그램을 이용하여 다방면의 연구를 진행 중이다. 최근에는 음성인식 기능 구현을 통해 인공지능과의 자연스러운 소통을 위한 새로운 말뭉치 구축 및 구현 등 생성된 자료를 새롭게 살펴볼측면이 있다.

현 시점의 자료 구축 현황과 현재까지의 언어 처리 기술이 국어 연구에 어떻게 적용되고 있는 지를 검토하여 앞으로의 국어연구의 방향 및 인공지능 기술 발전이 국어 연구의 어떤 측면에 이용될 수 있는지 논의하고 인공지능 발달을 기반으로한 4차산업 시대에 대응하는 국어연구방법은 무엇인지 함께 살펴보고자 한다.

2. 4차 산업혁명과 언어

2016년 다보스 포럼에서 세계 경제가 직면한 문제의 대안으로 4차 산업혁명(The Fourth I ndustrial Revolution)을 언급한 이래로 4차 산업혁명은 전 학문 영역의 영향을 미치는 중이다. 3차 산업혁명까지의 정보기술과는 확연한 차이를 보이며 디지털, 물리적, 생물학적 경계가 없어지고 모든 영역에 기술이 융합되는 양상을 띠고 있다. 4차 산업혁명의 대표적인 기술인인공지능, 로봇, 사물인터넷(IoT), 무인자동차 등에 적용되는 음성인식 기술은 4차 산업을 이해하고 발전시키는 데에 큰 역할을 할 것으로 전망된다. 이 장에서는 4차 산업혁명과 언어의상관관계에 대해 살펴보고자 한다.

2.1. 4차 산업혁명의 언어학적 의의

인공지능 기술이 빈번하게 사용되고 있는 분야 중 하나가 번역기이다. 단순 기계번역이 아닌 딥러닝 방식에 기반하여 이전 세대보다 발전된 번역 기술을 보여주고 있다. 1) 인간의 언어지식에 기대는 것이 아닌 기존에 구축된 말뭉치를 기반으로 딥러닝을 적용한 번역 기술은 이전 세대의 기계 번역과는 다른 양상을 보여주고 있다. 또한 단순 번역에서 나아가 애플의 Siri나 마이크로소프트의 Cortana, Amazon의 Alexa, Google의 Google Assistant, 삼성전자의 Bixby 등의 챗봇이 인간이 요청한 기능을 수행하는 데에 더 정밀화 되고 있다.

2.2. 딥러닝과 언어처리

여러 챗봇이 딥러닝을 통해 학습하고 요청 기능을 수행한다. 이는 모바일 또는 컴퓨터 시스템을 통해 명령과 수행이 진행되는데 현행 챗봇의 수행도에서는 ① 음성인식 정확도 ② 연속 대화 불가 ③ 특정 주제 외 대화가 어려움 등의 문제점이 드러나고 있다. 딥러닝을 통해 학습한 언어정보도 아직까지 자연스러운 결과 도출 단계는 아니며 끊임없이 인간의 언어와 언어의 의미를 학습하는 과정에 있다고 할 수 있다.

3. 국어의 언어정보화의 현황과 연구방법

3.1. 새로운 말뭉치 구축

국가 차원의 대규모 말뭉치 구축 프로젝트가 2018년에 재가동되기 전까지 학계 및 관련 산업에서 이용할 수 있는 주요 한국어 말뭉치는 '21세기 세종계획' 말뭉치였다. 약 2억여 어절단위로 구축된 '21세기 세종계획' 종료 이후 후속 사업이 진행되지 않다가 2018년부터 다시국가 차원의 대규모 말뭉치 구축 프로젝트가 시작된 것이다. 이는 4차 산업혁명의 주요 기술인 인공지능의 발달 및 상용화와도 밀접한 관련이 있다고 할 수 있다.²⁾

언어자원의 중요성을 인식하고 1998년부터 2007년까지 대규모로 진행된 '21세기 세종계획'은 국립국어원이 추진한 말뭉치로 소강춘(2016:117)에 제시한 24,648.9만 어절로 대부분의 연구자들은 문자 자료 말뭉치로서 구성된 세종계획 말뭉치를 활용하여 연구를 진행해 왔다. 해당 규모는 인공지능 학습자료로 제공하기에는 양적으로 부족함이 있어 국가 차원의 대규모 말뭉치 구축 사업이 요청되었다. 소강춘(2019)에 따르면 21세기 세종계획의 24,648.9만어절과 새로운 말뭉치 구축 사업을 진행하여 국어 기본 말뭉치 10억 어절을 구축하는 것이라고 하였

¹⁾ 구글 신경 기계 번역(GNMT) 시스템이나 네이버 파파고(papago) 등은 데이터 부족으로 아직까지 만족스러운 번역 문장을 제시하지 못하는 측면이 있기도 하지만 구축된 말뭉치의 규모에 따라 해당 문제는 점차 해결될 것으로 보인다.

²⁾ 음석인식 기술을 앞세운 인공지능 기반 정보제공 및 제어 서비스는 생활 전반과 밀접한 연관을 맺게 되었다. 미국, 일본에서 꾸준히 구축한 말뭉치 규모와는 큰 차이가 있지만 우리말 인공지능 서비스의 개발에 이용된 국립국어원이 구축한 '21세기 세종계획'의 말뭉치는 인공지능 학습에 큰 역할을 했다는 것이 자명한 사실이다. 4차 산업혁명의 대표적인 분야인 인공지능은 기계의 언어 능력 처리 기술에 집약되고 있고, 인공지능의 우리말 학습에 사용되는 말뭉치는 양적 질적으로 대규모화될 필요가 있는 것이다. 말뭉치가 국어 연구에만 국한되는 것이 아니라 학계가 쌓아온 지식의 적극적인 산업화에 기반이 되는 시점이 지금이라고 할 수 있다.

다.

<표 1 새로운 국어 말뭉치 구축 사업 계획>

1998~2007 (21세기 세종계획)	24,648.9만 어절
2018년 구축 말뭉치	3천만 어절
2019년3)	7억 7천만 어절 말뭉치
2020년~	원시, 주석 말뭉치와 새로운 말뭉치 개발 추진 예정

2020년 8월 25일 국립국어원에서는 기구축한 한국어 빅데이터 '모두의 말뭉치(https://corpus.korean.go.kr/)'를 온라인에 공개하였다. 2018년 대규모 한국어 학습 자료 구축 사업의 1차 결과물로서 18억 어절 규모의 말뭉치이다. 특이점이라면 최근 10년간의 신문기사, 서적 2만118종, 일상생활의 음성 대화 및 메신저 대화, 방송 자료, 대본 및 블로그, 게시판 등의 자료 210만 건도 포함된 다양한 말뭉치라고 할 수 있다.

<표 2 국립국어원 '모두의 말뭉치' 자료 목록>

자료명	자료 설명
	- 2009년부터 2018년까지 10년 동안 생산된 신문 기사 연
신문 말뭉치	1억여 어절
	- 기사 총 353만여 건
문어 말뭉치	- 책, 잡지, 보고서 등 저작물 20,188종의 문어 원시 말뭉치
	- 총 20,188건
	- 방송, 강연, 일상 대화 등의 구어 자료, 드라마 대본 등의
	준구어 자료로 구성한 구어 말뭉치
구어 말뭉치	- 구어 자료는 한글로 전사하였고 실제 발음대로 표기함
	- 공적 독백 2,490건/ 공적 대화 19,104건/ 사적 대화
	2,224건/준구어-대본 4,102건(드라마 4,102회 분량)
	- 두 명 이상의 대화 참여자가 메신저로 나눈 메신저 대화
메신저 말뭉치	자료
	- 총 4,203건(대화 메시지 2,174,506개)
	- 블로그, 게시판, 누리 소통망 등에서 수집한 웹 언어 자료
웹 말뭉치	- 리뷰 성격의 웹 언어 자료 포함
면 도오시	- 블로그 11,521건/게시판 9,089건/누리 소통망 1,989,656
	건/리뷰 96,710건
문서 요약 말뭉치	- 국립국어원 신문 말뭉치에서 추출한 기사 4,389건을 대상

^{3) 2019}년에 진행되는 사업에서는 신문 기사 원문 자료, 문어 말뭉치, 구어 자료 원시 말뭉치, 메신저 대화 자료, 일상 대화 원시 말뭉치, 웹 말뭉치(SNS, 블로그, 게시판, 리뷰 등) 원시 말뭉치를 구축할 계획을 세웠다. 성격에 따라 다양한 자료가 구축된다는 점에서 매우 반길 만하다.

	으로 기사에서 출한 주제 문장과 기사를 요약하여 작성한									
	문장으로 구성된 말뭉치									
	- 기사 추출 주제 문장 13,167개/기사 요약 작성 문장									
	13,167개									
형태 분석 말뭉치	- 총 300만 어절 규모의 형태 분석 말뭉치									
경네 군격 글장시	- 총 300만 어절(문어 200만 어절, 구어 100만 어절)									
어휘 의미 분석 말뭉치	- 총 300만 어절 규모의 어휘 분석 말뭉치									
에게 커의 판구 교상시	- 총 300만 어절(문어 200만 어절, 구어 100만 어절)									
	- 개체명의 경계를 인식, 15개의 의미 분류 체계에 따른 표									
개체명 분석 말뭉치	지 부착 말뭉치									
	- 총 300만 어절(문어 200만 어절, 구어 100만 어절)									
구문 분석 말뭉치	- 문장 구조를 파악하여 단어 간 관계를 분석한 말뭉치									
1	- 총 200만 어절									
	- 학술 간행물의 문법성 판단을 연령별로 조사한 말뭉치									
문법성 판단 말뭉치	- 총 19,940개 문장(문법적 문장 9,970개, 비문법적 문장									
	9,970개)									
	- 국립국어원 신문 말뭉치, 구어 말뭉치에서 추출한 기초 문									
유사 문장 말뭉치	장을 대상으로 컴퓨터가 만든 유사 문장과 사람이 작성한									
11 1 2 0 2 0 1	유사 문장으로 구성된 말뭉치									
	- 179,589개 문장									
	- <우리말샘>에 등록된 비슷한말, 반대말, 상위어, 하위어 어									
	휘 쌍을 대상으로 어휘 관계 강도를 평가한 후 결과를 통									
어휘 관계 자료:NIKLex	계 정보로 제시한 자료									
	- 어휘 관계 기초 자료 20만 쌍(비슷한말 60,000쌍, 반대말									
	10,000쌍, 상위어 70,000쌍, 하위어 60,000쌍)									

<표 2>4)에 정리된 자료는 연구의 목적에 따라 선택될 수 있으며, 구어 및 현재 생성되고 있는 웹상의 자료들도 수집되어 다양하게 이용할 수 있다. 소강춘(2019:79-80)에서 밝힌 바와같이 수요자의 요구를 반영한 말뭉치 구축이라는 점에서 보다 활발한 국어 논의 및 응용에 적용할 수 있을 것이다. 기초 국어 말뭉치 구축, 기초와 응용 국어 말뭉치 상시 구축, 국어 말뭉치 공유와 확산 체계 구축과 운영 등을 목표로 진행하는 국립국어원의 국어 말뭉치 구축 사업은 4차 산업혁명 시대의 국가 전략과 산업발전에 기여한다는 점에서 큰 의의를 갖는다.

소강춘(2019:89)에서는 ①원시 말뭉치, 신어 등의 언어자료, ②분석 말뭉치(형태, 구문, 의미, 감성, 개체명 등), ③유의어 자료(사전), ④유의어 자료, 유사 표현 자료, 의미역 격틀 자료, 뉘앙스 자료, 개체명(고유명사) 자료, 전문용어 자료, ⑤구어(대화체) 자료, 인터넷(웹, 채팅 등) 자료, 누리소통망(SNS) 자료 등을 모두 모아 우리말 빅데이터(말뭉치) 구축 지원, 공

⁴⁾ 이는 국립국어원 '모두의 말뭉치'에 게시된 각 말뭉치의 자료 소개를 정리한 것이다. 자료명의 정식명 칭은 제시한 명칭 앞에 '국립국어원'이 포함되어 있으나 제시한 표에서는 편집상의 문제로 생략하였다. 자료 소개 내에는 자료명, 공개일, 자료 유형, 관련 사업, 자료 설명, 예시 등이 균일하게 적용되어 있어 연구자 및 이용자가 이용 목적에 맞는 자료를 요청하여 이용할 수 있으며 약정서 제출 승인후 이용이 가능하다. 말뭉치는 JSON(UTF-8 인코딩) 형식으로 구축되어 있으며 이 중 '국립국어원 신문 말뭉치', '국립국어원 문어 말뭉치', '국립국어원 구어 말뭉치', '국립국어원 웹 말뭉치'는 기가바이트(GB) 단위의 대형 자료이다.

유, 유통의 기반을 마련하고자 하였다.

3.2. 지역어 말뭉치 구축

지역어 말뭉치는 상대적으로 규모가 크지 않은 편이다. 『한국어 지식 대사전』의 100만 여 어휘 중 방언은 약 9만 여개였을 뿐이다. 이태영(2016:53)에서 지적한 바와 같이 『우리말큰사전』, 『표준국어대사전』, 『한국어 지식 대사전』 등에 지역어 표제어 등록 방식이 표준어 대응형해설이어서 최선으로 보기는 어렵다. 이러한 현상의 원인이 방언 말뭉치 구축의 미비, 검색및 활용이 활발하지 않다는 점 등이 지적되었다. 주로 지역어 및 역사자료는 활용하는 사람이적다는 점 때문에 대규모로 사업화되어 결과물이 축적되기보다는 개인 연구자에 의해 구축되는 측면이 커 접근과 활용이 용이하지 않았다는 한계가 있었다.5) 하지만 현재 단계에서 지역어 보완 조사 등을 꾸준히 이루고 있어서 지역어 말뭉치 역시 양적인 축적을 이루고 있다.

<표 4 지역어조사 지역 목록>

도	지역어조사(2004-2011)	보완조사(2015-2019)									
경기	용인 처인/백암, 화성, 포천, 파주, 양평, 이천, 강화	연천, 안성, 시흥, 양주, 평택, 가평, 여주									
강원	삼척, 원주, 양양, 홍천, 평창, 인제, 정선	강릉, 고성, 양구, 철원, 춘천, 화천, 횡성, 영월									
충남	공주, 대전, 논산, 서천, 예산, 서산, 천안	금단, 당진, 보령, 부여, 아산, 연기, 홍성, 청양, 태안									
충북	제천, 청주, 충주, 옥천, 보은, 영동	괴산, 단양, 음성, 진천									
전남	함평, 곡성, 진도, 영광, 보성, 영암, 신안, 광양	강진, 고흥, 구례, 담양, 무안, 승주/순천, 여수, 완도, 장흥, 해남, 광산, 장성, 나주, 화순									
전북	완주, 남원, 무주, 군산, 고창, 임실	김제, 부안, 순창, 익산, 장수, 정읍, 진안									
경남	고성, 창원, 창녕, 산청, 남해, 울산, 하동	거제, 거창, 김해, 밀양, 사천, 진양/진주, 통영, 함양, 합천, 양산, 의령, 함안									
경북	경주, 상주, 청송, 고령, 청도, 의성, 봉화	군위, 김천, 문경, 안동, 영덕, 영양, 영일/ 포함, 영천, 예천, 울진, 선산, 영주, 성주, 칠곡									

⁵⁾ 지역어 자료 구축의 국가 사업의 진행 결과로 『한국구비문학대계』85권(1980-1992), 『한국방언자료 집』1-9권(1987-1995), 『지역어 조사 보고서』음성전사파일(2005-2012), 『21세기 세종계획 한민족언 어정보화 통합 검색 프로그램』 방언 말뭉치(1998-2007), 『민족생활어 조사 보고서』(2007-2013), 2015년 전국 방언 지도 제작 결과로 지역어 종합 정보 사이트(dialect.korean.go.kr)를 운영 중이며 『지역어 보완 조사 보고서』(2017~2019)가 가장 최근의 성과라 할 수 있다.

제주시 건입, 북제주군 구좌/한경, 서귀포시 호근/표선/색달/대정

<표 4>에서 보듯이 지역별 조사 지역 확보 역시 2020년 현재 많은 지역이 보완된 상태이다. 이러한 조사의 결과 역시 지역어 종합 정보(dialect.korean.go.kr)에서 공유 중이며 지역어 검색, 지역어 지도, 문학 속 지역어, 사진으로 보는 생활어 등의 서비스를 제공하고 있다.

3.3. 특수목적 말뭉치 구축

3.1.에서 지적되었던 것과 같이 수요자의 요구는 다양하고 이를 모두 수용한 말뭉치 구축은 요원한 상황이다. 국가 단위 말뭉치 구축에서 세부적인 요구를 다 수용하지 못하고 있는 상황에서 기업은 필요에 따라 자체 말뭉치를 구축하여 기술에 적용하기도 한다. 특수목적을 위해 구축한 말뭉치는 특정 맥락에 맞춰 집중적으로 구축된 말뭉치이기 때문에 대화맥락 분석 및통계적으로 이용하기에 장점이 존재한다.

아래 자료는 인공지능 학습용 대화 말뭉치 구축자료의 일부이다. 6) 음성인식 서비스를 사용하는 사용자의 요구에 따라 구성된 말뭉치는 대화문을 기반으로 한다. 인공지능 학습용으로 구축되었으나 연구용으로는 가공을 통해 txt 파일 등으로 변환 사용이 가능하며, 기존에 연구자들이 이용하는 'UTagger'(울산대 한국어처리연구실 배포), '지능형 형태 분석기'(국립국어원배포) 등으로 통계적 처리가 가능하다. 이와 같이 특수목적에 의해 구축되는 말뭉치는 AI 음성인식 서비스를 제공하는 기업의 요청에 의해 구축되는 측면이 강해 다양한 주제 말뭉치로확장하는 방안도 모색되어야 할 것으로 보인다.

SPEAKER			▼ Action																
	Utterance	Intent	title	singer	singer _gend er	song writer	lyricis	album	runnin g_tim e	genre	mood	speed	releas ed_ye ar	tpo	modi	i volum e	lyric	choic e	ord
U	십센치 음악 찾아 줘	request_search(singer 심선	!치)	십센치					-										
S	어떤 음악을 찾으시나요?	request_constraint																	
U	십센치 노래야	add_constraint(singer 십센	치)	십센치															
U	그 중에서 십센치가 첸이랑 콜라보 한 곡을 찾아 줘	add_constraint(singer 십센	치, 첸)	십센치	, 첸														
S	찾으시는 곡이 Byebabe맞나요?	search(title Byebabe - SM	Byeba	be - SN	STATIO	ON													
U	00	yes_answer																	
U	이 노래 언제 나온 건지 알려줄 수 있어?	ask_released_vear(title <b< td=""><td><byet< td=""><td>abe - S</td><td>M STAT</td><td>ION></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></byet<></td></b<>	<byet< td=""><td>abe - S</td><td>M STAT</td><td>ION></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></byet<>	abe - S	M STAT	ION>													
S	이 노래는 2017년에 나온 곡입니다	inform_released_year(title	<byet< td=""><td>abe - S</td><td>M STAT</td><td>ION></td><td></td><td></td><td></td><td></td><td></td><td></td><td>2017년</td><td></td><td></td><td></td><td></td><td></td><td></td></byet<>	abe - S	M STAT	ION>							2017년						
U	음 그래?	backchannel																	
U	십센치가 또 다른 사람이랑 콜라보 한 곡 있으면 알려 줘	add_constraint(singer 십센	치)	십센치															
S	너를 보네가 있습니다	search(title 너를 보네)	너를 !	보네															
S	재생활까요?	suggest_play																	
U	그걸로 재생해 줘	request_play(title <너를 보	<너를	보네>															
S	노래를 재생한니다	play																	
U	고마워	thank																	
S	서비스를 중료합니다	closing																	
U	하이 빅스비	greet																	т
S	네 무엇을 도와드릴까요??	opening																	
U	멜론에서 유명한 신곡 추천해 쥐	request_recommend(releas	ed vea	r 2019	₫. mod	ifier 유	명한)						2019년		유명한	1			
S	어떻게이별까지사랑하겠어널사랑하는거지를 추천합니다		recommend(title 어떻게이팅어떻게이별까지사랑하겠어널사랑하는거지																
U	이 곡 별로야	no_answer(title <어떻게이																	
U	다시 추천해 줘	request_recommend																	
S	흔들리는꽃들속에서네샴푸향이느껴진거야를 추천합니다	recommend(title 흔들리는) 흔들리는꽃들속에서네샴푸향이느껴진거야																	
U	좋아 이걸로 들어 줘	request_play(title <흔들리																	
S	재생을 시작합니다	play																	
U	수고했어	thank																	
S	이용해 주셔서 감사합니다	closing																	
U	노래 좀 찾아 줘	request_search																	т
U	내가 어제 들은 노래야	add_constraint(modifier 어	제 들은)											어제	ie.			
S	어제 들으신 10곡을 찾았습니다.	search(choice 10)																10	
U	전부 재생해 쥐	request_play																	
U	볼륨은 10으로	set_volume(volume 10)														10			
S	노래를 재생합니다	play																	
S	볼륨을 10으로 올립니다	volume_up(volume_10)														10			

⁶⁾ 해당 자료는 조경순(전남대) 교수가 ㈜폴투윈코리아와 협업하여 구축하였으며 음악과 영화 검색, 식당과 택시 예약 상황이 6.4만여 문장, 22.5만여 어절 규모로 구축되었다.

3.4. 인공지능과 국어연구방법

언어정보처리 기술의 발달로 국어 연구자들은 다양한 언어처리 프로그램을 개발하여 분석에 적용하였다. 김한샘·최정도(2020:9-10)에서도 언어 연구를 위한 '기본적인' 프로그램으로 '텍스트에디터', '(말뭉치) 분석 도구'로 '지능형 형태소 분석기', '유태거(Utagger)', '검색 도구'로 '글잡이 II(직접)', '글잡이 II(색인', '한마루 1.0', '한마루2.0' 등을 제시하고 있다. 만약 연구에 사용할 원시 말뭉치가 구축되어 있지 않다면 이러한 기본 정보를 먼저 구축해야 하지만, 앞서 살펴본 것처럼 대규모 원시 말뭉치가 구축되어 있는 상황이라면 이 단계는 생략이 가능할 것이다. 그러나 살펴본 것과 같이 국어연구에서 사용하는 연구방법은 말뭉치를 분석하여 이용 가능한 원시형태로 만들고, 목적에 따라 검색하는 방식을 사용하고 있다.

3.4.1. 다양한 범주의 말뭉치 구축

이용 가능한 말뭉치의 구축 한계가 있으며, 아직 만족할 만한 말뭉치가 구축되어 있지 않다면 접근성은 더욱 낮을 수밖에 없는 것이다. 한가지 더 고민할 것이 있다면 엄청난 속도로 생성중인 인터넷 상의 언어자료들을 어떻게 가공하고 처리할 것인가 하는 점이다. 그간의 연구는 문어와 구어를 일정부분 나누어서 처리하였던 측면이 있다. 하지만 '국립국어원 구어 말뭉치(버전1.0)', '국립국어원 메신저 말뭉치(버전 1.0)', '국립국어원 웹 말뭉치(버전1.0)'에서 수집한 자료의 성격을 어떻게 이해할 것인가 하는 점 역시 앞으로 대규모 말뭉치를 이용한 연구에서 고민하지 않을 수 없다. 남길임(2018:25)에서 확인한 웹 말뭉치에 대한 평가는 다음과같다.

<표 4 웹 말뭉치 영역의 평가(남길임(2018:25) 수정 인용>

	ㄱ. 기존의 균형 말뭉치와 다른 엄청난 용량 제공
장점	ㄴ. 즉각적인 현실 언어 반영
	ㄷ. 자유롭게 활용 가능
	ㄱ. 웹 자료의 대표성 및 균형성 문제
문제점	ㄴ. 웹 언어 자원의 시기에 대한 언어 연구 수용의 방법 문제
	ㄷ. 웹의 장르, 레지스터 등이 언어학 연구에 가지는 함의

문어를 대상으로 한 말뭉치의 분석은 어느정도 정제된 모습을 보여주고 있으나 문자유희 등이 자유로운 웹 및 메신저에서는 특정 어휘나 구문을 연구하기에는 적정하지 않을 수 있다. 전통적인 문자언어 활동을 하는 것이 아닌 '댕댕이' 등이 '멍멍이'로 쓰이는 것을 어떻게 처리할 것인가 하는 문제도 생겨나는 것이다. 더불어 메신저의 언어는 구어인가 문어인가 하는 점등 역시 연구자들이 고려해야 하는 다양한 말뭉치 품질과 가공 기준의 고려대상이 될 것이다.

기존의 문어 말뭉치, 구어 말뭉치 외 매체 변화에 따른 언어 사용 양상이 차이나는 점은 말뭉치를 이용하는 국어학 및 응용국어학 및 전산언어학 분야 연구자들은 새로운 논의를 시작할때이다. 국어학 분야에서는 주로 말뭉치의 국어학적 분석이, 전산언어학 분야에서는 언어의기계적 인식 오류 감소 방안이나 자동 태깅 등에 대한 연구가 진행되어 왔다.

이렇게 대규모로 구축된 말뭉치를 이용한다 하더라도 말뭉치는 계속해서 추가 구축되어야 하고, 또한 구축한 말뭉치의 질적 향상을 도모해야 한다. 아쉬운 점이라면 현대국어 위주로 구축된 말뭉치에 비해 지역어이나 국어사 자료는 말뭉치를 이용하여 균일한 처리나 통계적 분석이 요원하다는 것이다. 국어사의 경우 현재 훈민정음 창제 이후 시기의 문헌 자료 말뭉치가 구축되어 있으나 연구자들의 검토에 의해 계속해서 수정 보완되고 있는 상황이다. 훈민정음 창제 이전의 차자표기 자료들도 꾸준히 발굴됨과 동시에 언어 정보화 하여 국어의 역사적 정보 이해에 더욱 기여할 수 있을 것으로 보인다.

대규모 현대어 말뭉치가 구축되고 있으나 현대어 및 표준어에 국한된 것이 많고, 지역어의 자료적 지위는 높지 않았다. 하지만 지역어가 가지는 특징을 표준어와 경쟁하여 배척되어야 하는 것이 아니라 보완 관계에 있어 국어의 언어 사실을 선명하게 파악할 수 있게 한다. 기존에 한글고전소설이나 판소리 사설, 구비문학, 문학 작품 말뭉치와 방언 어휘 말뭉치 등 지역어를 기반으로 한 말뭉치 구축이 꾸준히 이루어져 왔으며, 연구자의 조사가 이루어져 왔지만이를 체계적으로 관리하고 정제하는 작업들 역시 더욱 확장되어야 할 것으로 보인다.

3.4.2. 인공지능과 도메인 지식을 이용한 말뭉치 구축

지금까지의 말뭉치 구축 현황을 보면 자료입력 및 정보 제공자의 데이터 수집을 통해 이루 어져 왔다. 웹 데이터의 경우 크롤링(crawling)을 통해 자료를 수집하는데 웹에 존재하는 데 이터 전체가 유의미한 것이 아니라 그 중 자료를 선별하는 과정이 선별되어야 할 것이다.7)

박진호(2019:138-139)에서 공학적인 기계학습에서 도메인 지식의 역할이 축소되었다는 인식이 있지만 딥러닝 기법에 의한 언어처리에 분명한 한계가 있음을 지적하고 있다. 형태소 분석기 등을 통해 분석한 결과에서 분절의 문제나 분절의 문제가 있을 때 품사 태깅이나 형태소 분석을 어떻게 할 것인가 하는 문제가 생긴다는 것이다. 현재 단계의 기계학습으로 이를 해결하고자 하더라고 ①분류 ②회귀 ③생성/변환 중 하나의 방식을 선택해야 한다는 것인데, 한국어 형태소 분절을 분류로 처리하고 모델링 한 결과 기존의 형태분석기보다 높은 성능을 보였다는 결론이다.8) 정밀하고 고품질의 말뭉치를 구축하기 위해서는 딥러닝 등의 신경망 모형으로 대표되는 처리 방식에 언어학자의 도메인 지식을 통한 검토가 필수적이다.

3.4.3. 산업 연계 말뭉치 구축 및 활용

기존의 말뭉치는 텍스트 말뭉치였다면, 텍스트화 되지 않은 자료의 데이터화가 계속해서 병행되어야 할 것이다. 이 경우 인공지능의 이미지 인식 기능을 통해 기본적인 처리가 가능한 영역이라는 충분히 활용할 수 있을 것이다. 앞서 챗봇 등이 기능이 명령-수행에 국한되었다면 4차 산업혁명 시대의 발전 방향은 감성적 영역의 요청 역시 가능한 인공지능 개발이 목표가될 것이다. 특히 딥러닝을 이용한 음성의 감정 분류 연구 등은 세부적인 구어 연구 영역에도 적용할 것으로 기대된다. 동일한 구문이라고 하더라도 감정에 따라 표현하는 의미 차이가 발

⁷⁾ 웹을 통해 제공되는 현행 기사문의 경우 보도자료를 제공하여 원자료가 존재하는 경우 ①보도자료 그 대로 송고 ② 보도자료 취사 선택 송고 ③ 보도자료 기반 취재 후 기사 송고 등으로 동일한 내용이 양적으로 늘어날 수 있는 환경이 있으므로 자료 처리에 따라 어휘나 문장 등의 출현 빈도 문제 등이 차이가 날 수 있다.

⁸⁾ 박진호(2019)에서는 세종 형태의미 분석 말뭉치를 대상으로 1200만 어절 규모를 검토, 200개의 범주를 확인하여 스크립트 상의 오류를 체크하여 신경망 모델 구성과 훈련에 적용하였다. 해당 결과 어절 단위 정확도 96.25%, 형태소 단위 재현율 98.17%, 형태소 단위 정확률 97.84% 의 성능을 보여 UTagger 등의 형태소분석기보다 월등히 높은 성능을 확인했다고 한다. 언어학자의 도메인 지식이 딥러닝 등의 공학적 한계를 보완할 수 있음을 실증하였다.

생하므로 음성 감정어 DB 구축과 음성 분석 시스템 성능의 향상은 4차 산업시대에 인공지능의 활용도를 더욱 높일 수 있는 방향이 될 것이다. 이를 위해서는 국어의 감정연구 등의 연구결과가 반영되어 데이터 수집과 분석이 필요할 것이다. 매스미디어 시대의 다양하게 확보 가능한 음성·음향 데이터와 나아가 이미지 처리는 국어연구에서 새로운 방향을 제시할 것이다.

이러한 기술을 이용하여 구축한 지역어 말뭉치를 산업적으로도 활용할 수 있는 방향이 확장될 것으로 보인다. 음성 인식 및 대화시스템의 고도화를 이루고 표준어와 지역어와의 대응, 의미 대응 및 공백에 대한 시각화를 가능케 할 것이다. 더불어 전사 외의 음성자료 등의 확보가 산업적으로 이용가능한 지점을 발견한다. 지역어 자연발화의 대화 맥락의 특징이나 사용자의 억양에 따른 의미 분석 등은 지역어 기반 말뭉치의 활용도를 높이고 자연발화(지역어)의음성인식 기술 등의 상용화에 학문적 의의를 제공할 것으로 기대된다.

4. 결론

앞서 살펴본 바와 같이 4차 산업시대는 인공지능으로 대표되고, 인공지능은 대규모의 말문치를 기반으로 한 언어처리 및 이미지, 음성 처리의 단계까지 와 있다. 그렇다면 기계분석과차이는 어디에 있는 것이며, 인간과 국어연구자들은 어떤 역할을 담당할 수 있을 것인가 묻지않을 수 없다. 인공지능이 끊임없이 정답을 찾아내게 하고 신경망 학습을 진행한다고 하더라도 자유 문맥 및 대화맥락 안에서 생성되는 의미의 변별과 새로운 표현의 생산은 아직까지 인공지능이 도달하지 못한 영역이라고 할 수 있다. 이러한 고유한 인간 언어 영역 및 지식을 바탕으로 빠른 속도로 변화하는 언어 환경에 대한 분석이 필요하며 그것은 현재의 언어자원에 대한 이해와 수집에서부터 가능할 것으로 판단한다.

참고문헌

강범모(2011), 『언어, 컴퓨터, 코퍼스 언어학』, 고려대학교출판부.

권혁승·정채관(2012), 『코퍼스 언어학』, 한국문화사.

김한샘(2017), 「한국어 학습자 말뭉치 주석의 요건과 실제-형태 분석 말뭉치 구축을 중심으로」, 『배달말』 61, 배달말학회, 149-173쪽.

김의중(2016), 『알고리즘으로 배우는 인공지능, 머신러닝, 딥러닝 입문』, 위키북스.

남길임(2018), 「웹 말뭉치를 활용한 언어 연구의 현황과 쟁점」, 『한국어 의미학』 60, 한국어의 미학회, 23-49쪽.

박준혁·이성욱·임윤섭·최종석(2019), 「대화 말뭉치 구축을 위한 반자동 의미표지 태깅 시스템」, 『소프트웨어 및 데이터 공학』8-5, 정보처리학회, 213-222쪽.

박진호(2019), 「딥러닝 기반 자연어 처리에서 도메인 지식의 역할」, 『인공지능인문학연구』 4, 중앙대학교 인문콘텐츠연구소, 135-166쪽.

이태영(2003), 「방언 말뭉치의 전산화와 활용」, 『한국어학』 21, 한국어학회, 65-104쪽.

이태영(2016), 「국어사전과 방언의 정보화」, 『국어국문학』61, 국어국문학회, 51-73쪽.

장경준, 허인영(2019), 「<우리말샘> 어휘 역사 정보의 현황과 개선 방안」, 『한국사전학』 33,

- 한국사전학회, 116-152쪽.
- 장필성(2016), 「[EU] 2016 다보스포럼: 다가오는 4차 산업혁명에 대한 우리의 전략은?」, 『과학기술정책』 26(2), 과학기술정책연구원, 12-15쪽.
- 정승철 외(2019), 『2019년 지역어 보완 조사』, 국립국어원.
- 조진수, 최선희, 김범진, 강지영, 나상수, 구본관(2020), 「말뭉치를 활용한 공공용어의 난도 공식 개선 연구」, 『국어국문학』 191, 국어국문학회, 193-226쪽.
- 최병서, 이익훈, 이상구(2020), 「신조어 및 띄어쓰기 오류에 강인한 시퀀스-투시퀀스 기반 한국어 형태소 분석기」, 『정보과학회논문지』, 47(1), 한국정보과학회, 70-77쪽.
- 홍정하·김주영·강범모(2008), 「세종 구문분석 말뭉치의 구축과 통사 범주 및 기능의 통계적 분포」, 『민족문화연구』 49, 고려대학교 민족문화연구원, 285-331쪽.
- Eddington, D.(2017), 『언어학자를 위한 통계학』, 홍성훈 옮김, 한국외국어대학교 지식출판 원.

"4차산업에 대응하는 국어연구방법 고찰"에 대한 토론문

이금영(충남대)

이 논문은 다양한 말뭉치 구축 현황을 살펴보고 4차 산업혁명시대에 대응하는 국어 연구의 방향을 제시한 점에서 의의가 있으며, 특히 국어정보학 분야의 연구에 흥미를 가지고 있는 많은 연구자들에게 기본적인 길라잡이가 될 수 있을 것으로 판단됩니다. 다만 이 분야에 대해 깊이 있는 지식을 가지고 있지 못한 토론자는 다음과 같은 몇 가지 질문을 드리고자 합니다.

- 1. 서론에 의하면 이 논문의 목적은 1) "현 시점의 자료 구축 현황과 현재까지의 언어 처리 기술이 국어 연구에 어떻게 적용되고 있는지를 검토하여 2) 앞으로의 국어 연구의 방향 및 인공지능 기술 발전이 국어 연구의 어떤 측면에 이용될 수 있는지 논의하고 3) 인공지능 발달을 기반으로 한 4차산업 시대에 대응하는 국어연구방법은 무엇인지"고찰하는 것인데, 2)와 3)이 명확히 구분되어 기술되고 있지 않으며, 특히 2)의 밑줄 친 부분에 대한 논의가 명확히 기술되고 있지 않은 것으로 보입니다. 즉 3.4절에서는 현재까지 대부분 현대국어 그리고 표준어 위주의 말뭉치 구축에서 국어사 자료와 지역어에 대한 말뭉치 구축이 보다 확대될 필요가 있으며 아울러 다양한 범주 및 주제의 말뭉치 구축이 필요한 점, 구어 말뭉치와 문어 말뭉치의 구분 등 말뭉치 자료의 성격 규명이 필요한 점, 도메인 지식을 이용한 말뭉치구축, 텍스화 되지 않은 자료의 데이터화, 음성 감정어 DB 구축 등이 '인공지능과 국어연구방법'으로 논의되고 있는데, 이는 대부분 앞으로의 국어 정보학 분야의 연구 방향으로 볼수 있는 것이 아닌가 합니다.
- 2. 토론자 역시 이 논문에서 주장한 대로 구어 말뭉치와 문어 말뭉치, 메신저 말뭉치와 신문 말뭉치의 구분 등 다양한 말뭉치 자료의 성격 규명이 필요하다고 보는데 이에 대한 발표자의 견해는 어떠한지 묻습니다.
- 3. 3.1 등에서 '수요자의 요구를 반영한 말뭉치'와 3.3에서 '다양한 주제 말뭉치'를 언급하고 있는데 그와 같은 말뭉치의 예들을 좀 더 구체적으로 제시하고, <표 2>에 제시된 말뭉치와 어떻게 다른지 비교하는 논의가 있으면 좋겠습니다. 그리고 이러한 다양한 말뭉치의 발달이 국어학 전반의 연구에 어떠한 영향과 의의가 있는지 고려해 볼 수 있으며, 그 반대 방향도 거론될 수 있겠습니다.
- 4. 11쪽에서 "나아가 이미지 처리는 국어 연구에서 새로운 방향을 제시할 것"이라고 기술하였는데 이미지 처리가 어떤 면에서 국어 연구에 새로운 방향을 제시한다는 것인지, 그리고 "의미 대응 및 공백에 대한 시각화"는 구체적으로 어떤 의미를 갖는지 좀 더 부연 설명이 제시되면 좋겠습니다.

뇌과학과 국어 연구 : 언어신경학의 연구들에 대한 메타적 접근과 구문 이론의 모색

Neuroscience and Linguistics of Korean:

A study for the theory of syntax with meta-research to linguistic neurology

오충연(숭실대)

1. 머릿말

이 연구는 문장 이해 및 산출에 대한 언어신경학(linguistic neurology)의 연구 결과들에¹⁾ 정합하는 통사론의 연구방법을 모색하는 데 목적이 있다.

구문처리(syntactic process)에 관련된 주요한 뇌신경 영역은 기존에 알려진 바와 같이 브로카 영역으로 보인다. 그러나 그 과정이 문법형식의 처리와 동일화되거나 직접 대응하지는 않으며 자체적이고 자율적인 연산체계도 아니다. 구문처리의 신경학적 내용은 브로카 영역이 뇌의 거의 전 영역에 걸친 상호 신경교류라고 할 수 있다. 신경학적 범주는, 형태론이나 통사론, 그리고 의미론의 경계를 설정하는 등의 기존 언어학의 학술 범주와 동일하게 구분하기 어려워서, 언어신경학의 연구를 이론 언어학에 적용하는 일은 쉽지 않다.

구문생성과 관련된 신경처리 내용에 대한 필자의 생각을 단도직입적으로 말하자면, 의미기억으로서의²⁾ 어휘항들은 베르니케 영역을 통해서³⁾ 브로카 영역의 좌측삼각부(left pas triangularis, LPT)에 처리되어야 할 작업기억항목으로서 유입되는데, 어휘항들의 조합이 구 (Phrase) 또는 기타 유의미한 통사론적 단위가 되기 위해서는 좌측판개부(left pas opercularis, LPO)의 역할이 주요한 것으로 보인다.⁴⁾ 그런데 LPO가 구문처리를 하는 방식에 대해서는 흔히 언어학에서 제시되는 통사론적인 연산으로 가정되기는 어려운데, 특히 자체적인 처리를 가정하는 일은 현실성이 없다. LPO는 시각, 체감각, 운동 영역뿐만 아니라 전두엽에 이르기까지 전되(whole brain)상의 연결선(wires)를 가지고 있다. 뇌의 각 구역은 구문처리 과정중에 브로카 영역과 상호적인 활성화 패턴을 갖는 것으로 보이는데, 이는 어휘항들이

¹⁾ 언어신경학은 신경심리학(psycho-neurology)의 하위분야라고도 할 수 있는데, 가시화할 수 있는 신 경학적 내용들을 심리적 과정으로 해독하는 방법을 사용함으로써 언어 과정에 대해 보다 객관적인 근 거로 형식화할 수 있을 것으로 기대된다.

²⁾ 의미기억(semantic memory)는 장기기억(long-term memory)의 일종으로, 표시체와 표시대상(음상과 개념)이 자의적인 관계의 상징능력이라고 할 수 있다. 장기기억은 의미기억 외에도 절차기억과 일화기억을 포함한다. 이때의 의미는 뇌의 한 섹터를 차지하는 기록된 창고 방식이 아니라, 신경활동의 재구에 해당하므로 기억시기마다 일치하지 않는다. 따라서 베르니케 영역은 그것을 저장한 곳은 물론아니고, 다중 영역의 활성화에 의한 의미기억을 음상과 결부하여 언어적 처리를 유도하는 통로로 보인다.

³⁾ 베르니케 영역을 '통한다'는 데에 유의하기 바란다. 어휘 의미기억은 베르니케 영역에 저장된 것이 아니라, 의미에 따른 각 뇌구역의 연결이 베르니케 영역에서 음상적 연결로 이루어짐으로써 가능한 것으로 보인다. 따라서 베르니케 영역은 LPT에 신경 데이터가 유입되는 경로로서 이해하는 것이 더 적확하다.

⁴⁾ 브로카 영역(Broca's area)은 브로드만 영역(Brodmann Area) 44, 45(BA 44, 45)에 해당하며 이들은 각각 LPO와 LPT에 해당한다.

각 영역의 인지적 통합과 재인을 통해서 유의미한 조합으로 변환되는 과정으로 해석할 수 있다. 예컨대 여러개의 어휘들로 이루어진 조합은 다른 영역들의 섭동에 의해 비로소 구 (phrase)에 해당하는 하나의 심리적 표상으로 전환되는 것으로 보이며, 이것이 문장처리 과정에서 일시적인 단일항으로 축소된 작업기억으로 변환되는 효과를 얻음으로써, 어휘항들의 나열이 구 단위로서의 재편성되는 통사론적 효과를 주는 것으로 보인다. 여기까지는 본문에서 언어신경학적 연구의 결과들을 토대로 고찰할 수 있으리라 기대한다.

그러나 기존 언어학의 통사론적인 연구가 이에 정합하기 위해서는, 방법론적이 전환이 필요하다. 기존의 생성문법에서 뇌의 모듈적 특성을 통사론이나 의미론과 같은 문법 층위의 분립성의 근거로 생각하던 방식과는 거리가 있기 때문이다. 그렇다고 해서 브로카 영역이 통사론적 핵심기능이 없다고 단언하기도 어려운데, 다만 이것은 신경학적 기능이기 때문에 언어학의 통사론적 형식으로 환원하여, 순수한 통사부가 있느냐 없느냐로 결부시키에는 난해하다. 이러한 문제들은 통사론이 이전과 다른 방식으로 전환되어야 함을 시사한다.

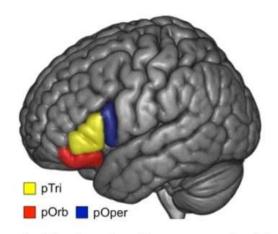
2. 전제와 배경

2.1. 신경학적 쟁점

2000년대 이후에는 EEG나 MRI 등, 각종 신경 영상화 기기가 발달하면서 언어신경학에서도 뇌의 단원성(modularity)에 근거하여 통사부를 별도의 연산체계로 설정한 생성문법에 대하여 검증하려는 시도들이 있었다. 그 핵심은 브로카 영역의 기능이다. 간략히 추리면, Grodzinsky를 중심으로 브로카 영역의 활성화가 구문의 복잡성보다는 어휘의 수와 관련된 읽기 폭, 즉작업기억 부담에 비례하여 활성화된다는 결과와, 구문구조의 계층적 복잡성에 비례한다는 Makkuchi(2009)와 같은 연구 결과가 대립한다.

브로카 영역은 좌측두판개부(left pas opercularis, LPO)로 불리는 브로드만 영역 44(Brodmann Area 44, BA44)와, 좌측두삼각부(left pars triangularis, LPT)라 불리는 BA45영역으로 구성되어 있다. 선행연구에 대한 필자의 검토로는 어휘수, 즉 읽기 폭과 비례하는 영역은 주로 LPT이며, 계층적 구조에 비례하는 양상은 주로 LPO에서 보이는 것으로 파악된다.5) 그런데 LPO가 자체적으로 문법 처리 수준의 기능을 한다기보다는, 통사론적 단위의 재인 및 산출을 위해 뇌의 각 부위들의 다중적 활성화를 조율하는 신경처리 기능을 하는 것으로 보인다.

⁵⁾ Makkuchi(2008), Matchin(2014) 참조.



The left inferior frontal gyrus (LIFG), divided into standard anatomical subregions. pOper = pars opercularis, pTri = pars triangularis, pOrb = pars orbitalis

<그림1> 브로카 영역의 구성(Matchin 2018:1683)6)

따라서 앞으로의 신경학적 쟁점은, 언어처리의 층위별 기능에 대응하는 영역을 찾아내는 일보다는 연결 패턴을 검출함으로써 패턴의 분석을 바탕으로 한 언어과정 설정으로 변화될 것으로 생각된다. 이런 연구에는 연결선을 효과적으로 검출하는 방법과 영역별 활성화 시차 등을계산할 수 있는, 공간 및 시간 해상도가 높은 기기 및 기술의 개발이 필요하다.

이러한 분야의 연구에서 구문생성에 대한 문법학의 역할은 문법적 처리의 과정을 설정하는 신경심리학에 형식적 내용을 제공하는 일이기도 하지만, 보다 주도적으로는 언어과정에 대한 이론적 제안으로 보다 효율적인 실험 목표를 정하고 실험 설계에 참여하는 일이다.

2.2. 모듈화 특성에 대한 접근법

되는 하나의 신경 또는 신경군의 단위로 신경 처리를 하지 않는다. 우리가 하나의 인지 또는 신체 기능이라고 개념화하여 명명한 어떤 일을 수행하기 위해서는 여러 뇌구역의 상호작용이 필요하다. 그것은 아무리 단순한 의식 활동이라도 마찬가지다. 따라서 뇌가 모듈화되어 있다는 것이 곧 의식기능이나 신체적 기능이 각각의 영역에 분화되어 있음을 뜻하지는 않는다. 그러므로 오랫동안 인지적, 신체적인 기능 손상과 결부된 뇌병변의 대응 연구를 직접적으로 언어처리에 적용한 일은 오류가 있어 보인다. 예컨대 'A기능'으로 명명된 의식 또는 신체 기능에 장애가 발생하여 검사한 결과가 뇌의 a구역의 손상에 대응한다고 했을 때, a구역이 A기능을 담당한다는 식의 추론이다. 대부분은 A기능의 수행에 참여하는 뇌의 신경 요소들은 그보다 훨씬 복잡하게 얽혀 있다. 비록 a구역이 A기능에 참여하기 때문에 병변이 되기는 하지만 그것만으로 A기능이 발휘되는 것도 아니다. 즉 a구역의 역할은 A 기능의 필요조건이지만 충분조건은 아닌 것이다. 반대로 a구역이 A기능에만 참여한다고 제한을 두어서도 안 된다.

예를 들어 베르니케 영역(BA23)에 손상이 발생하면 잘 알려진 바와 같이 어휘인출에 문제

⁶⁾ 브로카 영역은 좌측전두판개부(left inferior pars opercularis, LPO:pOper)와 삼각부(left inferior pars triangularis, LPT:pTri)로 구성되며, 안와부(pas orbitalis :pOrb)도 문법형식과 관련하여 중요한 역할을 하는 것으로 보인다.

가 생긴다. 정확히는 어휘 인출을 못한다기보다는 잘못된 어휘를 인출하는 것이다. 그 이유는 이 부위에 사전처럼 어휘들이 기록되어 있어서가 아니다. 어휘 의미는 뇌의 각 영역들이 의미특성에 따라서 음상과 결부하여 활성화되면서 재생되는 것으로 보이는데, 베르니케 영역은 이때 음상과 각 영역에서 구성하는 의미를 연결하는 핵심적 부위로 보인다. 기 물론 이때의 '의미'도 음상에 연결되는 특정 부호, 즉 소쉬르의 시니피에(signifie)와 같은 단일 관념이 아니다. 우리가 어떤 단어에 대한 의미표상이라고 느끼는 것마저도 복합 신경활동에 의해 구성된 것이다. 언어학에서 가정하듯이 어휘의미는 한 단위가 아닐뿐더러, 이것을 다시 의미자질로 분해할 수 있는 내용의 것도 아니다.

구문처리도 마찬가지이다. 예컨대 BA44(LPO)가 구문구조의 복잡성에 비례하여 활성도가 높아지는 경향을 보인다고 해서, 우리 문법학자가 생각하는 수준 -예컨대 구구조 형성이나 이동조작 등- 의 문법적 처리가 BA44에서 이루어진다고 결론짓는 일은 위험하다.

뇌의 모듈화를 문법 층위와 직결시키는 토대 위에서 신경학적 내용을 검토하고자 하는 언어 신경학의 연구 방향도 수정되어야 한다.

2.3. 해석의 문제

근래의 뇌신경 연구는 MRI를 주로 사용하는데, fMRI는 MRI로 초당 수백 회 이상 연속촬영하여 동영상화할 수 있는 기법이다. MRI는 신체에 자기를 형성하여 반송되어 오는 자기 공명신호를 영상신호로 처리하는 기기와 기술을 동원하게 된다. 자기 반응은 산소포화도에 따라이루어지는데, 동맥에서 공급되는 신선한 혈액일수록 산소가 많이 녹아 있다. 결과적으로 뇌의 혈류량을 측정하여 영상으로 처리하는 일이다.

이러한 실험에서는 기초적인 해석의 문제가 완전히 해소되지 않는다. 극단적인 경우에는 동물의 죽은 뇌에서도 종종 MRI 반응이 발생하기도 하는데, 이것을 곧바로 의식 활동이라고 간주하기는 어렵다. 기본적인 해석의 방법은 혈류량 자체를 데이터의 양으로 간주하는 일인데, 혈류량이 많은 것은 에너지의 양과 직접적인 관계가 있겠지만 이것이 반드시 처리하는 정보량이 더 많다고 확신할 수는 없다. 예컨대 자동적 처리는 적은 에너지로도 많은 데이터를 처리할 수 있고 비자동적 처리는 그 반대일 것이다. 또한 자동성과 비자동성은 뇌구조 자체에 기인할 수도 있고, 반복적 경험에 의해 강화된 인지도식(schema)에 의한 것일 수도 있다. 에너지가 많이 든다는 것은 언어 외적인 처리를 병합해야 하는 것으로 해석할 수도 있어서 언어적처리의 증가로 바로 평가되어서도 곤란하다.

또한 앞 절에서 말했듯이, 특정 구역의 활성도가 그 구역의 역할로 알려진 기능과는 반드시 일치하지 않을 수도 있다. 모든 신경 영역은 연결체계로 이루어져 있기 때문이다.

EEG는 뇌파검출기라고 할 수 있다. 머리의 각 부위별로 산출되는 뇌파를 검출하는데, 단말의 위치와 파의 강도를 사용하면 신호의 조합을 이미지로 구성할 수 있다. 물론 공간해상도는 MRI에 비할 바가 아니지만, 측정 대상이 직접적인 뇌활동이라는 점과 무엇보다 시간해상도가훨씬 높다는 장점이 있다. 이 때문에 근래에는 MRI와 EEG를 병합한 연구들이 늘어나고 있는데, 이때 전자기파 신호 간섭으로 왜곡되는 현상을 상쇄하기 위해 기술적인 방법을 동원해야한다.

^{7) 2016}년 4월 캘리포니아 버클리 대학(UC Berkeley)의 연구팀이 발표한 뇌 의미지도(semantic map)는 대중에게 오해를 줄 수 있는 소지가 있다. 그것은 특정 담화상에 나타나는 어휘의미에 따라 활성화되는 뇌 영역을 보여주는 것이지, 해당 단어가 들어 있는 사전항목의 위치로 평가해서는 안 된다.

3. 신경학적 고찰

3.1. 작업기억과 통사처리

브로카 영역(Broca's area)의 기능에 대한 꽤 긴 논쟁도 앞절의 문제에서 비롯되는 것으로 보인다. 잘 알려진 바와 같이 브로카 영역의 손상은 통사론적인 능력을 상실케 하는 실문법증 의 원인이다. 수형도로 표현되는 구문구조와 해당 구조에서의 문법단위가 이동 등의 연산에 이루어지는 일이 신경학적인 실체로서 존재하는가에 대한 회의(懷疑)는 언어심리학자에게서 언어신경학자에게까지 이어지고 있다. 이에 대해 회의에 대표적인 연구자는 Grodzinsky이다.

Grodzinsky(2000)는 언어사용에서 좌전두피질(left anterior cortex)의⁸⁾ 기능적 역할에 대한 새로운 견해가 제안된다고 하면서, 전반적인 인간 언어능력 국한된 이 구역에서 구현되는 것은 아니라고 한다. 즉 전반적인 구문처리는 브로카 영역과 주변부위-예컨대 판개 (operculum)⁹⁾나 섬엽(insula)¹⁰⁾, 또는 하부백질(subjacent white matter)- 등에서 이루어지는 게 아니라는 주장이다. 그는 브로카 영역이 구문 처리에 중요한 역할을 하지만 '매우 한정적인 일부(higly specific one)'라는 것이다. 브로카 영역은 이동 변형과 흔적-삭제 가설로 선형화를 이루도록(in line with the Trace-Deletion Hypothesis) 위치를 추도(extraction)하는 일 사이의 관계 연산을 포함한 수용적 메커니즘을 하는 신경 중심부(neural home)라고할 수 있지만, 반대로 구조 형성이나 어휘삽입과 같은 기본적인 조합기능은 이 구역에서 지원되지 않으며 의미 조합도 지원하지 않는다고 한다.

Grodzinsky(2000)는 브로카 영역을 처리해야 할 알고리즘에 대한 역할을 기계적으로 배분하는 곳으로 평가하는 것으로 생각된다. 이러한 그의 생각은 브로카 영역의 활성화가 구문구조의 복잡성보다는, 거리효과와 같이 결합해야 할 단어와 단어 사이에 다른 단어가 많이 들어가 있을수록 더 많이 아루어진다는 연구들에 근거를 두고 있다. 이는 앞서 말했듯이 브로카 영역의 손상으로 인한 실문법증은 문법을 구성하는 한 고리가 고장난 것에 기인하는 현상으로서, 브로카 영역이 문법 전반을 담당해서가 아니라는 추론과 결부된다.

3.2. 작업기억과 분리된 핵심 연산처리

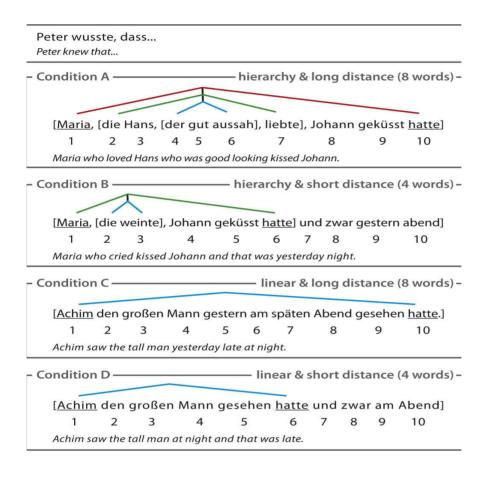
브로카 영역의 활성화가 작업기억부담, 즉 어휘항의 수와 비례한다는 다수의 연구 결과에 대해서, Makuuchi외(2009)는 다소 상반된 연구결과를 제시한다. 이들의 연구에서는 어휘수가 같게 조정한 단문과 문 구조의 문장에 대해, LPO가 구조적 복잡성과 비례하여 활성화한다는 MRI 실험을 한 결과를 제시하였다. 이들의 주장은 연구의 제목 그대로 작업기억으로부터 핵

⁸⁾ 좌측 전측두엽은 브로카 영역을 포함한 구역이다.

⁹⁾ 판개(operculum)은 브로카 영역 주변의 측전두엽 전반을 일컬으며, 브로카 상부의 두정판개(parietal operculum), 앞쪽의 전두판개(frontal operculum)와 회판개(orbital operculum), 그리고 측두부의 측두판개(temporal operculum)으로 구성된다.

¹⁰⁾ 섬엽(insula)은 브로카 영역 안쪽의 피질부위에 해당하며 실비우스 틈새를 열어서 확인된다(필자주). 감정 및 감각과 관련이 된다고 알려져 있으며, 내뇌인 변연계(limbic system)에 상대적으로 가깝다. 변연계는 비교적 낮은 밀도의 신경조밀도를 이루지만 처리속도가 빠르고 대뇌의 척수와 연결되어 감각의 중간통로에 연결된 것으로 보인다. 변연계의 핵심인 시상(thalamus)은 거의 모든 뇌신경활동에 작용하는 것으로 보이며, 언어과정(language process)에서도 예외가 아니다.

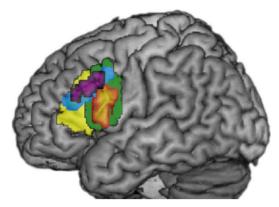
심적인 연산처리 능력이 분리될 수 있다는 것이다.



<그림 2> Makuuchi외(2009:8363)에서 실험과제로 제시된 문장의 구조들11)

<그림 2>에 대한 실험은 각각 구조효과와 거리효과를 측정하기 위한 것이었다. 결과는 LPO와 좌측하전두골(left inferior frontal sulcus, LIFS)에 있어서 구조효과는 A와 B간에 LPO활성화가 두드러지게 높게 나타났지만, C와 D 간에는 미약하나 D가 약간 높았다. 물론 A와B가 C와 D에 비해서 확실히 높았다. 즉 A>B>C, C<D이다. LIFS의 활성화는 A>C>B>D로 계층성과 정비례하지는 않았과 다만, C와 D의 주어-술어 거리에 따라 비례함으로써 거리효과를 보여주었다. 이때 A와 B의 주어-술어 거리와도 맞물린다. B보다 C가 활성화된 것도 거리에비례한 것이다. 12)

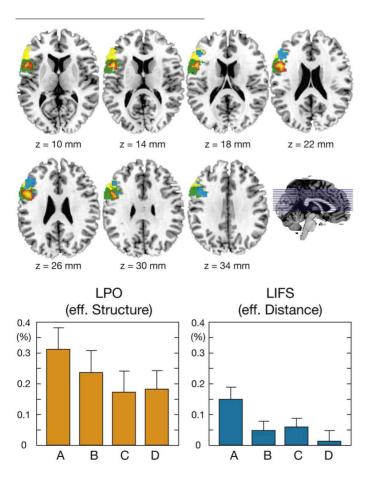
¹¹⁾ 이 문장들은 각각 10개의 단어로 이루어져 있다. A와 B는 계층적이고, C와 D는 그보다는 선형적이라고 한다. A와 B에서는 선의 색과 밑줄로 짝을 보여준다. C와 D에서는 서술어의 위치를 바꾼 것이다. 이들 문장은 각각 주어, 술어, 기타요소 등이 재배치되면서 문장의 구조가 형성된다.



<그림 3> 과제별 주요 활성화 영역에 의한 심리생리학적 분석(Makuuchi외 2009)

<그림 3>에서 주황색은 구조 효과(녹색을 포함하여 BA44), 파란색은 거리 효과(보라색을 포함하여 LIFS)를 보이는 구역이다. 노란색은 BA45(LPT에 해당)구역이다.

이러한 결과에 대한 Makuuchi외(2009)의 설명은 LPO가 작업기억 항목수에 의한 거리효과 보다는 계층적 처리에 의해 활성화됨을 보여주므로, 이것이 기존에 알려진 것처럼 브로카 영 역이 단순히 작업기억 처리가 아니라 문장의 계층적 처리를 담당한다고 함으로써, 이것이 작 업기억과 구분되는 핵심적인 언어능력이 있음을 보여주는 것이라고 주장한다.



3.3. 형태-통사론적 지표

이론 언어학 내부에서 고전적인 논쟁이 되어버린 통사론과 의미론의 독립성은 신경학적 관점에서는 무의미할 수도 있겠지만, 언어신경학자들은 여전히 이 문제에 매달리는 경향이 있다. 과연 의미와 무관한 통사부가 신경학적으로도 존재하느냐의 문제이다.

Shinri Ohta 외(2013)의 연구에서는 좌측하전두이랑(left inferior frontal gyrus)¹³⁾의 기능에 대해서 좀더 섬세하게 밝히고자 한 연구라고 할 수 있다. 실험은 일본어로 이루어졌는데, 의미가 없는 유사어로 이루어졌으나 통사구조상으로는 명확한 문장, 즉 재버워키文 (Jabberwocky)을 피실험자에게 읽히고 MRI 검사를 하는 것이었다. 이들의 통사구조는 조사나 어미로 파악된다.

(1) 가. 복합문 ざざが よよが ききると せえある。 (?자자가 요요가 키키한다고 세사다.) 나. 단문 ざざの よよが ひひ てとる。 (??자자의 도도-가 히히어 테토다.)

즉 (1가)와 같이 병합도(Dgree of Merger)가 높은 것이 활성도에 영향을 미치는 변수라고 가정하는 데에서 연구를 시작한 것이다. 이들의 연구에서는 우선 형태-통사론적 정보만 가지고 있고 의미는 알 수 없는 의사단어(pseudowords)로 이루어진 구성으로 재귀적으로 병합된 문장을 피실험자에게 제시하여 뇌활성화 검사를 한다. 결과는 좌측하전두이랑(left inferior frontal gyrus)의 활성화를 확인했다. 또한 이들의 결론은 좌하전두이랑과 상모이랑(SMG supramarginal gyrus)의 14) 네트워크가 재귀적으로 병합된 문장에서 계산을 보조한다는 것이다.

위 실험은 일본어로 시행된 것이지만, 한국어에도 시사하는 바가 있다. 문법요소가 구조적 골격을 예측하게 하는 것은, 해당 형식과 구문구조 사이에 후천적으로 형성된 도식이라고 보인다. 이것은 생애 초기에 형성되는 것으로 통사론적 지표가 되는 형태들에 대한 개별문법의 사상(寫像 mapping)이라고 보아야 할 것으로 보인다. 이는 보편적이고 추상적인 기저 구조에 의해 문법형식이 자리잡는 방식이 아니고, 반대로 개별 언어와 개별 문형에 의해 추상화된 형태-통사적 구조화에 의한 것이라고 생각된다.

한국어로 말하자면, 조사나 어미와 같은 요소는 어휘삽입과 같은 구체적인 의미해석 요인이 없더라도 구조를 예측하게 하지만, 이것은 구조-의미의 병합체인 개별 문장으로부터 추출된 요소라고 할 수 있다. 예컨대 관형형 '-ㄴ'은 선행하는 술어를 수식어로 바꾸는 기능 외에도 후행하는 명사핵을 예측하게 하고 전체적으로는 명사구와 격을 수반하게 하는 요소로 무의식적인 파악이 이루어질 수 있다. 따라서 '-ㄴ'은 관형어구 또는 관형절의 핵으로서만이 아니라, 더 상위요소인 명사구와 격을 예측하게 한다.15) 이는 LPO가 전뇌적인 연결을 통해서 어휘적의미요소들을 통사론적 단위로 구성하는 과정이, 형태론적인 자료와 결부되면서 이루어지는

¹³⁾ 브로카 영역(LPO, LPT)를 포함하여 안와부(pars orbitalis)까지의 전측두엽의 하부 구역을 이른다.

¹⁴⁾ 상모이랑은 체감각 영역의 일부로서 촉감과 공간인지 등과 결부되며 사람의 동작을 식별하는 데 관여하는 거울 뉴런 시스템의 일부이다. 이 영역의 손상은 수용성 실어증(베르니케 실어증)을 불러온다고 한다.(인터넷 Wkikpeda 정리 참조) 이 영역은 베르니케 영역과 상당 부분 겹친다.

¹⁵⁾ 이는 기존의 통사 이론에서 핵계층이나, 위-아래(Top-Down) 또는 아래위(Bottom-Up) 방식의 구성 원리가 실제로는 한정적인 면을 설명하고 있음을 말한다.

문법 사상으로 해석할 수 있다.

4. 브로카의 연산방식

4.1. BA44의 연결선들(wires)과 연결 부위

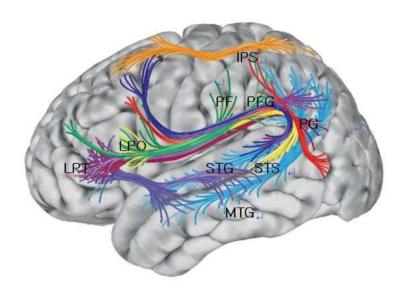
LPO가 구문의 계층적 처리에 밀접한 상관성을 보인다는 Macuuchi외(2009)의 연구는 고무적이지만, 언어학에서 형식화할 만한 체계를 이 영역이 수행하는 것으로 추정되지 않는다. 실제의 계층적 처리는 LPO 내부에서 이루어지는 것이 아니라, 다른 영역과의 활발한 신경 연결망에 의해서 이루어진다. 물론 이 과정에서 문법형식들에 대한 처리가 결정적인 역할을 하기도 한다.

Bernal 외(2015)는 57개 논문의 84개 실험 연구들로부터 데이터 풀링(pooling-data)을 통해 브로카 영역의 망(network)을 모형화하여, 16개의 상호 활성화 다발(cluster of activation)을 도출하였다. 주요 다발은 전두판개(frontal operculum), 좌측후측두 구역(left posterior region), 보조운동영역(supplementary motor area), 두정엽(parietal robe)이다. 그 다음으로는 왼쪽 시상(left thalamus)을 포함한 피질하 구조체(sub-cortical structures), 좌측 조가비핵(left putamen), 이차시각영역(secondary visual areas), 그리고 우측 소뇌 (right cerebellum) 등에서 발견되었다(Bernal 외 2015:1).

Bernal 외(2015)의 결과를 앞서 소개한 Grodzinsky(2000), Macuuchi 외(2009), 그리고 Otha(2013) 등의 연구와 결부하여 종합하면, 브로카 영역 주변의 전측두 구역은 문법적 형식에 대한 처리와 결부되고, 구단위의 구성을 위해 의미적 내용을 구성하는¹⁶⁾ 영역은 브로카 영역에서 상대적으로 먼 체감각 영역 및 시각 영역, 운동 및 전두엽에 이르기까지 광범위한 것으로 보인다.

브로카 영역을 중심으로 한 뇌의 영역간 연결선들은 Petrides(2014)에서 잘 종합해주고 있다. Petrides(2014:171)은 브로카 영역과 베르니케 영역을 중심으로 연결선(wire)의 연결체계를 비교적 총체적으로 보여준다. 브로카 영역을 이루는 구역은 좌측삼각부(left pars triangularis)와 좌측판개부(left pars opercularis)로 이루어져 있는데, 좌측삼각부는 상.중측두이랑(sup.mid temporal gyrus)과 집중적으로 연결되어 있다. 이들은 베르니케 및 청각 영역이다. 좌측판개부는 보다 많은 영역과 연결되어 있다. 구체적으로는 베르니케 영역의 상측두골(superior temporal sulcus)을 비롯해서, 이차적으로는 후측두엽(posterior temporal cortex), 후전두피질(lateral frontal cortex), 상.중측두이랑 및 상측두엽과 후상미이랑 및 전치외이랑, 그리고 두정엽내골(intraparietal sulcus) 등의 연결망을 형성하는 것으로 보인다.

¹⁶⁾ 이 연구에서 계속 언급하고 있지만, 여기서의 '의미'는 의미론에서 말하는 의미자질 분해 등의 형식적인 내용이 아니라 심상 등의 심리 및 신경학적 내용을 일컬음에 주의하기 바란다.



<그림 4> Petrides(2014:171) 언어 과정에 있어서 신경전달 경로로 보이는 연결선들의 양상¹⁷⁾

4.2. 연산방식에 대한 추론

Matchin(2014)의 연구에 의거하면 읽기폭에 비례 활성화를 보이는 구역은 LPO보다는 LPT 이다. 즉 이 구역이 작업기억을 담당하는 것으로 보이는데, 구문 이해 과정에서 청각영역 또는 베르니케 하부 영역으로부터 생성 인출된 어휘기억¹⁸⁾ 신호를 전달받아서 이를 LPO에 보내는 것으로 추정된다. LPO는 기능적 형식의 참조와 더불어 의미구성 단위를 만들어 내기 위해서 심상화 또는 추상적 포착에 필요한 다른 구역들 -예컨대 두정엽(운동), 후두엽(체감각, 2-5차 시각), 전두엽 등과 교섭해야 한다. 하나 또는 여러 개의 단어는 여기서 얻어지는 단일심상으로 전환되고 이것은 다시 단일 작업기억항목이 되어 전체 문장 해석에 참여하게 된다. 따라서 하나의 복합문을 이해한다고 했을 때, 작업기억 부담은 증가되었다가 해소되었다가를 반복할 수 있으며, 어휘수가 작업기억용량을 초과하더라도 즉시적으로 문장을 처리할 수 있다.

예컨대, (2)와 같은 문장에서 '승기'는 마지막 술어에 이르기까지 작업기억항목으로서의 부담이 유지되지만, '사과꽃'은 '예쁘다'를 읽는 순간에 부담이 해소되고 '사과꽃이 예쁘다' 전체가 하나의 어휘처럼 심상화되어 상위절의 처리에 참여하게 된다.

(2) 승기는 수지가 사과꽃이 예쁘다고 말했다고 회상했다.

¹⁷⁾ 이 연결선은 Petrides(2014)를 옮긴 것이나, 논의의 편의상 구역명은 필자가 표기하였음을 밝혀둔다. LPT는 좌측두엽과의 연결선들이 강화되어 있고, LPO는 후두 중앙부에 연결선이 집중되어 있다. 이 부위는 다시 전두엽 및 두정부, 그리고 시각 영역에 이르기까지 고르게 2차적 연결선들로 이어진다.

¹⁸⁾ 여기서의 어휘기억도 신경세포에 고정적으로 기록되어 있는 것이 아니라, 재구 과정을 통해서 생성된 것임에 유의하기 바란다. 기억의 재구는 참여 신경구역과 전달 패턴의 재생에 의해서 구현되므로, 베르니케 영역이 어휘인출을 한다는 것은 해당 어휘의미와 관련된 인지적 참여가 이루어지는 구역들의 연결망을 통해서 이루어진다.

이때의 '-는'은 술어를 통해서 부담이 해소되기 위해 대기중이지만, 연달아 나오는 명사구들은 이 해소를 지연시키고 그대로 부담이 남는다. 만일 '수지가'가 '수지이다'는 의미론적으로는 어색하지만, 통사론적으로는 작업기억 부담을 해소시키는 형태론적 지표로 절 재인 절차가 끝난다.

즉 '사과꽃이 예쁘다'는 두 개의 단어가 각각의 어휘기억으로부터 인출되지만, 이것은 의미처리를 위해 LPO가 상모이랑이나 2-5차 시각 영역에 신경 교류를 통해서 단일항으로 전환시킬 수 있다는 것이다. 그러나 '승기는'과 '수지가'의 연쇄는 이것이 어려운데, 의미내용을 단일화하기 위해서 다른 영역에 전달하기 전에 문법형식이 저지되므로 불필요하게 망체계에 신경부담을 주지 않을 것이다.

따라서 대립된 양상을 보이는 Grodzinsky(2000)와 Makuuchi(2009)의 연구는 실제로는 대립되지 않는 하나의 원리로부터 파생된 현상일 수 있다. 작업기억 부담을 유발하는 거리효과는 첫 항목이 마지막 항목을 만남으로서 해소되는 방식인데, 이 해소는 통사론적 항목으로 전환됨을 의미한다. 현재로서는 이러한 가정이 Grodzinsky와 Makuuchi의 비정합성을 정합적으로 설명할 수 있는 길로 생각된다.

이러한 과정은 통사론적 현상에 대해 접근하는 방식에 대해서도 시사하는 바가 크다. 예컨 대 구문 재인에 있어서의 작업기억 부담의 해소는 구문처리상의 일시적이고도 새로운 표상의 탄생으로 가능한데 성분의 이질적인 위치는 결국 표상처리가 되지 않는다는 것을 의미한다. 섬제약이나 국어의 어순제약 등과 같은 문법현상은 이러한 체계와도 맞물린다.

4.3. 재인(recognition)과 산출(product)

구문처리와 관련하여 지금까지의 fMRI 기법을 적용한 연구는 대부분 재인에 관한 것이다. 발화는 두상의 움직임 때문에 촬영하기가 어려움에 따라서 실험설계가 더 어렵기 때문으로 생각된다. 앞으로의 연구는 언어산출에 대해서 보다 본격적인 실험방법이 도입되어야 할 필요가 있다. 다만 문장의 재인을 중심으로 한 연구라 하더라도, 이를 산출로 추정하는 방법도 개발될 필요가 있다.

5. 구문이론(syntactic theory)의 모색

5.1. 구문생성에 대한 형식적 경계의 문제

기존의 문법학은 개별 언어자료에 대해, 문법양식의 함수적 상관성을 통해 추론되는 메타언어적인 형식화를 기반으로 하고 있다. 예를 들어, 이미 형식화된 개별 문장의 구조적 공통성에 대하여 상위적 형식으로 바꾸고 이것을 '심층구조'와 같은 생성기저로 표현하는 방식이다. 그러나 우리는 근원적인 문제에 봉착하게 되는데, 어떤 대상의 진정한 생성기저는 그 대상의질료나 양식이 아니라는 점이다. 단도직입적으로 말하자면, 언어의 생성기저는 언어가 아니며 언어로부터 축출한 형식도 아니다. 이는 하나의 학술 범주 내에서 이루어질 수 있는 논의에 대비해볼 때, 학술적 대상이 되는 범주 자체의 형성을 연구해야 하는 논의로서 범주 밖의 것을 보아야 한다는 것을 의미한다. 그러므로 범주 밖의 것을 범주 내의 형식으로 논의해야 하

는 방법상의 심각한 도전이 기다리고 있다.

통사론에서 언어로부터 축출한 기저양식들의 특성은, 언어의 특성을 그대로 가지고 있다. 가장 큰 특성은 연산과정이 단선적이라는 점이다. 신경학적 측면에서는, 자각할 수 있는 하나의 단선적인 의식을 구현하기 위해서는 복합적이고 병렬적인 처리가 필요하다. 심리적 실재성의 측면에서 본다면, 훨씬 효율적인 병렬적 처리 기제가 있는데, 이를 굳이 단선적인 규칙을 가정할 필요가 없다는 것이다. 단선적으로 설정한 처리 체계는 우리 문법학자들이 이해하기위한 수단이지, 심리적이거나 신경학적인 실체는 아니다.

다른 하나는 그 과정이 일방향적이라는 점이다. 예를 들어, 어휘 삽입이 이뤄지고 어휘-통사적 특성에 의하거나 고유한 구문구조에 따른 심층구조를 형성하고 다시 표면구조와 발화로이어지는 과정이 그러하다. 신경처리는 그와 다른 양상을 보이는 것으로 생각된다. 예컨대 어휘적 수준의 항목이 브로카 영역으로 전송된다고 하여도, 브로카 영역은 다른 영역과의 교류를 통해서 다시 작업 항목을 재설정하고 이것을 다시 상위 수준으로 재처리하는 과정이 있는 것으로 보인다. 본고의 관점은 언어심리학에서 가정하는 연결주의적 구문처리 모형에 좀더 가깝다. 연결주의는 문법처리의 수준이 병렬적이면서 상하향 교차적이다.

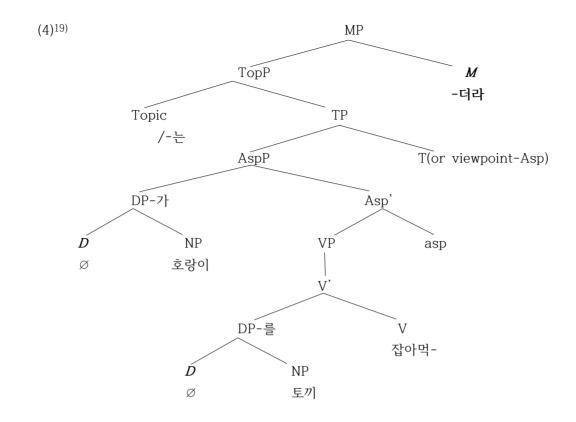
기존의 언어신경학적 연구 결과들에 대한 구문론적인 접근으로서 이 연구에서 제시하게 될 사항은, 문장의 이해나 발화의 과정에서 브로카 영역에 작업기억으로 등록된 복수의 어휘항들은 브로카 영역과 연결된 다중 영역의 망활동으로 구(phrase)에 대응하는 통합된 표상으로의 전환이 이루어지며 이것이 다시 단일 작업기억항목으로 브로카 영역에 재유입됨으로써 성분의계층적 처리가 가능하다는 것이다. 장기기억으로서의 어휘기억은 문장의 이해(recognition)나산출(production)에 있어서 하나의 작업기억(working memory) 항목으로 구문론적인 신경처리에 부담을 주지만, 여러 개의 어휘로 구성된 구 또는 기타의 구문론적 단위가 하나의 단일 작업기억 항목으로 재처리되는 신경과정을 통해서, 계층적 처리가 가능하다는 것이다. 이 과정에서 반드시 필요한 신경처리 내용은 통사론적 구조체를 형성하는 과정에는 언어적 신경내용과 비언어적 신경내용이 순환적으로 조응해야 한다는 것이다.

5.2. 생성과정에 대한 평면적 설명의 한계와 대안

문법형식, 특히 굴절형식은 강력한 통사론적 지표이다. 문법형식은 개별언어에 따라서 사상 mapping된 구조가 다를 수 있어서, 이것을 보편적인 형식으로 변환하는 일은 다소 어려워보이며, 필요해 보이지도 않는다. 그 이상의 보편성은 인간이 세계를 인식하는 방식, 즉 일반 인지체계와 구분되지 않을 것으로 생각된다. 예컨대 논항의 한정성이나 술어의 상은 인식론적인 보편성을 가지고 있다. 이것이 언어마다 표현될 수 있는 문법적 장치는 차이가 난다. 문법양식으로 보자면, 주격이나 완료상과 같은 문법양식을 보편성을 띠고 있는 통사론적 기제라고한다면, 그것은 기본적으로 인식론적인 체계와 구분되지 않는다. 따라서 통사론에서 언어적인 것과 비언어적인 것의 경계 설정은 어려워 보인다.

이러한 상황에서 통사론에서 중요한 근거는 다시 문법형식이 될 것으로 보인다.

그러나 전통적인 논리적 개념으로부터 마련된 통사론적 범주들을 다루는 방식은 수정할 필요가 있다. 예컨대 핵계층이론은 핵이 논항을 끌어들이거나 문법적 자격을 부여하는 방식으로 구성원리를 설명하려고 한다. 투사를 기반으로 한 상향적 구성이 문법원리의 기반이다. 그러나 핵은 해석의 결과이며, 생성의 기저가 아닐 수도 있다. 다음 (4)에 대해 생각해보자.



(4)에서 서법인²⁰⁾ '-더라'는 명제 내부의 구성요소인 논항의 양화성(quantity)과 관련을 맺는다. 이때에 발현되는 양화는 논리형식으로 표현할 수는 있으나 논리형식으로 일치되는 것은 아니며, 서법과 상관적 체계로서 해석된다. 실은 D에 관여하는 해석 요소는 상, 시제, 화제 등으로²¹⁾ 다변화된다. 여기에는 역현상도 있다. 지시사나 조사가 상이나 시제를 결정할 수도 있다. 예컨대 (5가)의 경우에는 엄밀히는 부정상(infinitve aspect)으로써 기존의 개념으로는 상태상(stative)에 속하는 속성상(attributive)이지만, (5나)에 표시된 논항의 지시성은 상이나 시간 해석을 현재진행으로 결정한다.

(5) 가. 호랑이는 토끼를 먹잇감으로 찾는다. 나. 저 호랑이가 토끼를 먹잇감을 찾는다.

이런 문제에 대해서 통사론의 쉬운 접근은 해당 위치의 문법형식이 생략 또는 삭제되었다고 하거나 제로형을 설정하는 방법이었다. 이는 구조주의에서부터 제안된 분포의 역할에 기인한 다. 특정 문법범주는 특정한 위치를 차지해야 한다는 원칙이기도 하다. 즉 해당 위치의 문법 성을 보증해주는 것은 자매항과 같은 결합계층상의 분포다.

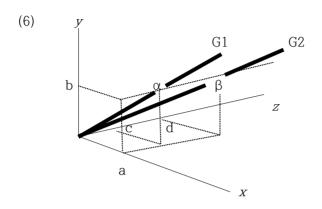
¹⁹⁾ 국어에서 주어의 서술짝은 상구로 보인다. 다만 영어와 같은 언어에서는 화제가 주어로 구현되는 언어이므로, 결과적으로 주어의 짝이 시제가 되어 절을 시제소구로 설정하는 일이 가능해보인다. 이 도해해에서 상이나 시제와 같은 기능적 범주에 대한 위상에 대한 논증은 필자의 다른 연구에 의거한 것이며, 이에 대한 증명은 이 연구에서 논외로 한다. 이러한 처리가 전체 논의와 직접적인 관련이 없다고 보인다. 다만 서법이나 시제 등 자매항이 아닌 범주가 한정소에 관여하는 현상에 대해 이들의 위상 차이를 보이기 위함이다.

²⁰⁾ 근래에는 문체법을 서법과 분리하나, 국어의 문말어미는 서법적 의의를 함께 가지고 있다고 생각된다.

²¹⁾ 예컨대, 완료상이나 과거시제는 논항을 지시적으로 표시한다.

언어의 선형적 특성과 성분의 순차성으로 나타나는 문법성을 확인하기 위해, 구문의 구조를 분석하는 방법에 대한 접근 방식은 계층적 구조를 그려내는 일이다. 그런데 계층구조나 이를 기반으로 한 변형과정은 평면, 즉 2차원적 해석이다.

2차원적 계층성으로 설명하기 어려운 병렬적 문법현상에 대해서, 두 벡터가 창출하는 평면 적 해석의 틀에서 나가 벡터를 더 추가하는 3차원적 방식을 생각해볼 수 있다.



(6)에서 G1과 G2는 다른 문법범주이다. 각각의 범주에 속한 α와 β는 직관으로도 다른 범주로 요소로 생각될 수 있다. 그 직관을 설명하는 데 있어서, 우리가 가지고 있는 문법적 틀은 예를 들어 평면 xz(또는 yz)로써 α와 β가 다른 범주라고 설명하는 데에 아무런 문제가 없다. 그런데 이것이 (4)의 D나 M처럼 일치요소와 같은 상관성을 가지고 서로 문법관계를 이루는 현상이 있다고 했을 때, 이것을 설명하는 방식에 대해 생각해보자. 2차원적 설명으로는 성분통어(C-command)와 같은 위상에 따른 제약을 설정할 수는 있겠으나, 구체적인 상호 점검이나 일치성을 설명하기는 어렵다. 만일 α와 β를 각각 (4)의 D와 M이라고 한다면 이들의 일치성은 xy면의 xa, yb의 위상으로 설명해야 한다. 반대로 우리의 설명 평면이 xy라면, α와 β는동일한 문법형식 또는 문법범주로 보일 수 있지만 다른 것일 수 있다.²²⁾

만일 우리의 현실적인 문법 직관이 α 와 β 의 차이를 포착하면서도 상관성을 현상적으로 알고 있지만, y축이 설명체계 밖에 있다면 문법설명은 실패한다. 다음 (7)을 보자.

- (7) 가. 물이 바닥에 흘러 있다.
 - 나. *시냇물이 강에 흘러 있다.

cf. 그 계곡의 시냇물이 이곳 한강까지 흘러들었구나.

'-어 있-'을 완료 또는 결과상태로 해석하고, 'N에'를 도달점으로 설명하는 문법체계로는 (7가)와 (7나)의 문법성이 같기를 기대하겠지만 직관은 그렇지 않다. 이러한 경우에 통사론에서는 의미론이나 세계지식의 문제로 넘기기도 한다.

5.3. 구문생성 이론과 언어 신경 과정

²²⁾ 필자의 연구로는 국어의 '-었-'이 그러한 특성을 갖는다. 영어에서도 '-ed/en'이 시제로 기능하기도 하지만, was 등의 조동사로 이끌리는 구문에서는 완료상이다. 이는 주격 점검의 문제를 유발하는데, 격과 굴절소의 일치 현상을 설명하는 일은 2차원적인 문법 지평에서 한계를 갖는다.

기존의 구문생성에 대한 관점은 분석을 기반으로 한다. 여전히 언어자체에 대한 분석이 우선이다. 그러나 구문생성에 대한 심리적 과정은 앞서 비유한 것처럼 전체적인 메시지의 형성과 그것을 언어로 형상화하기 위한 순서라고 할 수 있다. 설계이념에 처음부터 벽돌과 창호가구분되지는 않듯이, 특정 언어요소가 다른 언어요소를 끌어들여서 문장을 이룬다는 상향식 형성으로 구문생성이 이루어지는 것은 아니다. 적어도 신경학적 관점에서는 명제와 서법, 핵과논항이 구분되는 체계는 구문생성의 출발이라기보다는 결과를 해석한 일에 해당한다.²³⁾

그러한 점에서 구문생성과정은 발화의 과정과 정합될 필요가 있다. 이미 오래전에 제시된 Fromkin(1971)의 다섯 단계 모형의 과정은 어휘적 수준의 핵이 스스로의 통사자질에 따라서 구문을 형성해 나가는 것이 아니라, 반대로 메시지가 구구조를 형상하고 이 형상된 위치에 어휘를 설정하며 발화로 이어지는 순서이다. 따라서 여기서 말하는 메시지 수준의 내용은 비언어적이라고 할 수 있다. 구조는 물론 어휘 선정에 되지 않은 상태이기 때문이다. 이 과정이(4)의 문장에서는 이차원적 평면에서 장거리에 해당하는 문법요소 간에도 상호점검하는 현상을 일으키는 요인을 설명하는 데 더 유리할 것으로 보인다.

6. 맺음말

이 연구에서는 기존의 언어신경학적 연구들을 토대로 구문처리에 대한 신경학적 과정에 대해 추론하였다. 문장의 재인은 어휘항이 작업기억부담으로 LPT에 유입되지만 이것이 구구성으로 파악하는 과정은 LPO가 주도하는데, LPO의 작용 체계는 전뇌적 연결선을 통해 의미표상을 마련하는 것이다. 이것은 다시 단일 작업항목이 되어 LPT에 유입되며 순환적인 과정을 거친다. 이 과정은 단방형적이지도 않으며 단선적이지도 않다. 양방향으로 점검되며 병렬적처리가 가능한 것으로 보인다. 안와부(pars orbitalis)는 문법적 표지들을 처리하여 구구조를 예측하게 하는 것으로 보인다. 이것은 문장의 재인과정에서 일어나는 일이지만, 몇가지 점에서는 산출에도 적용할 수 있다. 병렬적 처리로 이루어지는 과정이라는 점과 재유입 순환이 가능한 상호처리가 가능하다는 점은 재인과 산출에 공통적이기 때문이다. 특히 언어산출에서는 언어형식이 특정되지 않은 상태가 시작점이기 때문에, 핵과 같은 형식적 요소를 구문형성의 씨앗처럼 설정하는 방식은 수정이 필요해보인다.

본 연구에서는 이러한 특성을 보다 잘 설명하려면 이론적인 언어학에서 보다 입체적인 구문 생성 체계를 고안할 것을 제안하였다.

참고문헌

오충연(2012), 국어 명사의 한정체계, 『어문학』117, 한국어문학회. 오충연(2019), 뇌과학적 의미 탐구의 현황과 과제, 『한국어 의미탐구의 현황과 과제』, 한국문

²³⁾ 다음과 같은 비유를 해보자. 즉 분석을 토대로해서 기둥을 놓고 벽돌로 벽을 쌓고 문틀과 창틀을 내고 그 안에 창호를 내고 벽체 위에 지붕을 얹는 것이 집의 형성과정인가의 문제이다. 그보다는 어떤 집을 지을 것인가에 대한 설계이념과 그에 따른 설계도, 그리고 그에 따른 구조물의 위치설정과 재료의 지정이 순서이다.

화사.

- Byron Bernal, Alfredo Ardila and Monica Rosselli(2015), Broca's area network in language fuction: a pooling-data connectivity study. *Frontiers in Psychology, vol.6.*
- Grodzinsky, Y.(2000), The neurology of syntax: Language use without Broca's area, Behavioral and Brain Sciences 23, 1-71.
- Huth, Alexander G., Heer, Wendy A. de, Griffiths, Thomas L., Theunissen, Frederic E. & Gallant, Jack L.(2016), Natural speech reveals the semantic maps that tile human cerebral cortex, *Nature* 532, 453-458.
- M. Makuucki, J. Bahlmann, A. Anwander, & A. D. Friederici(2009), Segregating the core computational faculty of human language from working memory, *PNAS*.vol 106. 8362-8367.
- Matchin, W. G.(2014), Investigations of the Syntax-Brain Relationship. Ph. D. thesis UC Irvine, UMI.
- _____(2018), A neuronal retuning hypothesis of sentence-specify in Broca's area. *Psynon Bull Rev* 25, 1682-1694.
- Petrides, M.(2014), *Neuroanatomy of Language Regions of the Human Brain*. ELSEVIER.
- Santi, A. & Y. Grodzinsky (2007). Working memory and syntax interact in Broca's area, *NeuroImage 37*, pp.8-17.
- Shinri Ohta, Naoki Fukui, Kuniyoshi L. Sakai(2013), Syntactic Computation in the Human Brain: The Degree of Merger as a Key Factor, *PLOS ONE* vol.8, issue 2. 1-16.

"뇌과학과 국어 연구 : 언어신경학의 연구들에 대한 메타적 접근과 구문 이론의 모색"에 대한 토론문

이정훈(서강대)

오충연 선생님의 발표 잘 들었습니다. 언어와 관련된 뇌 과학 연구 분야에 대해 비판적으로 고찰하고, 통사론에 대한 시사점을 모색해 보는 기회를 가질 수 있었습니다.

발표 내용이 구체적인 현상에 대한 분석, 구체적인 제안보다는 전반적인 상황이니 지향점에 대한 것이므로, 보다 구체적인 연구 사례를 보태면 발표 내용을 더 쉽게 이해할 수 있을 듯합니다. 이에 말씀하신 내용을 이어서 한두 가지 연구의 사례를 소개하고, 이어서 발표 내용을 보다 충실히 이해하기 위한 질문 몇 가지를 드리고자 합니다.

신경 언어학 연구 사례

가. 주격 중출 구문의 유형

- 박명관/최기용/정원일(2019), Differentiating Three Types of Double Nominative in Korean: An ERP-Based Study, 생성문법연구 29-1.
- 실험 자료
 - ② 영이가 마음이 예쁘다. [속성 관계]
 - ④ 영이가 동생이 예쁘다. [가족 관계]
 - 母 서울이/서울에 차가 많다. [주격/여격]
- ERP 실험 결과
 - syntactic anomaly
 - ② 대 에/의: the morpho-syntactically intimate relation
 - 의 대 의 : 의카 more exceptional syntactic licensing
 - All in all, this points to the fine differentiation of the three different types of DNC, albeit of the same form composed of two NOM-marked NPs.

나. 부사어 어순 선호도

- 유제욱(2019), 한국어 다중 부사 구문의 실시간 처리에 나타나는 어순 선호도, 석사학위 논문, 건국대학교/
- {시간, 장소} > 도구

발표에 대한 질문

다. '5.1. 구문 생성에 대한 형식적 경계의 문제'

- '형식적 경계'의 '형식'이란?
- 기존 통사론은 '단선적/일방향적'이고 신경학적 측면은 '병렬적/상하향교차적'이라고 함. 이에 대해 다음과 같은 점은 어떻게 이해하면 될는지? 첫째, 생성문법과 같은 기존 통사론도 병렬적 연산을 인정함. 예를 들어 '나의 친구가 네가 쓴 논문을 읽었다'는 '나의 친구가'와 '네가 쓴 논문을'이 병렬적으로 형성되는 과정을 포함함(cf. workspace & MERGE). 둘째, '일방향적'이라는 판단의 내용은 '어휘삽입 다음에 심층구조 형성, 그 다음에 표면구조' 식의 순서로 이해됨. 그런데 이러한 순서는, 초창기에는 가정되었으나, 90년대부터는 폐기됨. 어휘항목이 결합하며 구조가 형성될 뿐이며, 심층구조와 표면구조의 구분 없음. 셋째, 기존 통사론은 일방향적 도출만으로도 현상을 설명할 수 있으므로 일방향적 설명을 택함. 만약 상하향교차적이라면, 언제 상향적이고, 언제 하향적이며, 또오르락내리락 하는 구체적인 과정을 밝혀야 함. 넷째, 기존 통사론은 두뇌의 작동 방식과 독립적임. 언어 현상이 있고 이 언어 현상을 논리적으로 설명할 수 있는 가설 체계를 고안한 것일 뿐임. 동화책 식으로 얘기하면 이러한 가설 체계가 우연히 인간 종의 두뇌에 구현되어서 신경 언어학적 관찰이 가능한 것고, 인간 종과는 전혀 다른 생명체(?)에 구현되면 신경 언어학적 결과는 나타나지 않을 것임.
- "단선적으로 설정한 처리 체계는 우리 문법학자들이 이해하기 위한 수단이지, 심리적이 거나 신경학적인 실체는 아니다"라고 하였는데, 발표자가 생각하는 심리적 실체, 신경학적 실체가 무엇인지? 참고로 문법학자들이 밝힌 것은 현상을 설명할 수 있는 '가설 체계'인바, 이러한 '가설 체계'가 실체인지(발견) 아니면 그야말로 이해를 위한 수단인지(발명)는 논란거리인 듯함. 이에 대해 '수단'쪽을 지지하는 것으로 판단되는데, 근거는?

라. '5.2. 생성 과정에 대한 평면적 설명의 한계와 대안'

- "병렬적 문법 현상에 대해서 … 3차원적 방식을 생각해 볼 수 있다"라고 하였는데, 3차원적 구조 형성 방식은 기존에도 제안된 바 있음. 이에 대한 요약적 논의는 Carnie, A.(2010), Constituent Structure, Oxford University Press, 10장 참고. 강명윤 (1995)의 '그 자동차가 후면이 범퍼가 끝이 날카롭다'와 같은 격 중출 구문 분석도 같은 맥락으로 판단됨(강명윤(1995), 주격 보어에 관한 소고, 생성문법연구 5-2). 한편 기존 통사론은 그러한 추가적인 장치 없이 해당 현상을 설명하는 방향을 취하고자 함.
- (5)와 관련하여, (5가)가 부정상이라면, '검치호랑이는 토끼를 먹잇감으로 찾았다'도 부정 상? (검치호랑이: 4,000만년 전~만 년 전까지 살았던 고양이과의 육식동물) 또 (5나)가 현재진행이라면, '나는 일어나면 찬물 한 잔을 마신다'도 현재진행?
- (4), (6)과 관련하여, "α와 β를 각각 (4)의 D와 M이라고 한다면 이들의 일치성은 xy면의 xa, yb의 위상으로 설명해야 한다. (즉, xy 평면에서 D와 M의 위치가 같다?) 반대로 우리의 설명 평면이 xy라면, (xz라면 혹은 yz라면?) α와 β는 동일한 문법형식 또는 문법 범주로 보일 수 있지만 다른 것일 수 있다."라고 하였는데, x, y, z는 각각 무엇? 또 (6)과 같은 3차원 공간은 어떻게 나타나는가?
- (7)과 관련하여, '-어 있-'이 성립하려면 '물, 시냇물'이 사건 종결 후 존재해야 하는데 (기존 논의의 '존재 조건'), (7나)는 '강'이 (7가)의 '바다'와 달리 존재 조건에 부응하지 못해서 성립하지 않는 듯. 즉, 존재 조건을 총족하려면 '강'이 용기(container)처럼 해석

되어야 하는데 어휘적 용법상 그러한 해석이 수월하지 않아서 (7나)가 성립하지 않는 듯. 그래서 '강'을 '호수'로 바꾸면 '시냇물이 호수에 흘러 있다'에서 보듯이 성립 가능성이 제고됨. 따라서 '-어 있-'을 결과 상태로 해석하고 '바닥에, 강에, 호수에'를 도달점으로 간주하고 설명할 수 있는 듯.

마. '6. 맺음말'

- 신경 언어학적 연구의 "특성을 보다 잘 설명하려면 이론적인 언어학에서 보다 입체적인 구문 생성 체계를 고안할 것을 제안"함. 이 제안은 두뇌의 신경학적 작동 방식과 '이론 언어학'의 관계가 매우 밀접한 것으로 파악하는 것으로 사료됨. 그런데 과연 그러한가?
- 바. 끝으로, 발표 제목 "뇌과학과 국어 연구"와 관련하여, 구체적인 언어 현상으로 제시된 (5), (7)은 뇌과학과 관련하여 어떻게 이해하면 되는지? 두뇌는 병렬 처리 기제인데 이러한 병렬 처리 기제를 언어 분석에 도입해야 한다?

논지 파악이 미비한 점은 해량하시기 바랍니다. 끝.

구어구무하히

중부방언 방언구획과 네트워크 분석

정 성 훈 (목포대학교 국어국문학과)

국 어 국 문 학 회

국어국문학회

1. 방언

- (1) 정의
 - 표준국어대사전, "한 언어에서 사용 지역 또는 사회 계층에 따라 분화된 말의 체계"
- (2) 분류
 - 지역방언 (regional dialect)
 - : 한 언어의 사용 지역에 따라
 - 사회방언 (social dialect)
 - : 한 언어의 사용 계층에 따라

2. 방언 구획

- (1) 정의
 - 지역방언의 특징을 파악하고, 지역방언들의 특징에 따라 몇 개의 방언권으로 나누는 일련의 작업
- (2) 방언구획 방법
 - 전통적인 방법
 - : 등어선(isogloss) 설정
 - : 등어선이 겹치는 등어선속(bundle of isoglosses) 특정 → 방언의 경계
 - 계량적인 방법

국 어 국 문 학 회 ㅡ

3. 방언권 구획

- (1) 지역 크기에 따라
 - 대방언권
 - 중방언권
 - 소방언권

4. 방언권 구획 연구

- (1) 대방언권
 - 小倉進平(1920, 1944), 이극로(1932), 河野六郞(1945), 이숭녕(1967), 김형규(1974), 김영황(1982), 김병제(1988) 등
- (2) 중/소방언권
 - 이기백(경북 1969), 김영태(경남 1975), 김영송(경남 1977), 이익섭(강원도 1979, 1981), 이기갑(전남 1984), 김택구(경남 1991), 소강춘(전북 1989), 김충회(충북 1990), 전광현(전북 1995), 김덕호(경북 1997), 정철(경상도, 1997), 김정대(경남 2012), 정성훈(경상도 2016), 정성훈(전라도 2017), 정성훈(강원도 2018), 강경원(충남 2018) 등

국 어 국 문 학 회

3. 한국어와 대방언권 구획

- (1) 오구라신페이 (小倉進平, 1920/1944)
 - 1944년, 〈조선어 방언의 연구〉
 - 현대적인 한국어 방언 구획 연구의 시작
 - : 한국어의 계통과 특징을 밝히려는 목적
 - : 한국어의 공시적인 방언자료, 전국적으로 조사
 - 6개의 대방언권으로 구분
 - ① 평안도방언: 평안북도(옛 후창군제외), 평안남도
 - ② 함경도방언: 함경북도, 함경남도(영흥이남 지역 제외)
 - ③ 경기도방언: 경기도, 충청북도, 충청남도,

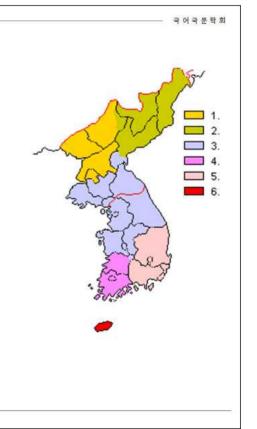
강원도(옛울진군제외),

함경남도(영흥이남지역),

전라북도(무주, 금산지역)

- ④ 전라도방언: 전라북도(무주지역제외), 전라남도
- ⑤ 경상도방언: 경상북도(옛 강원도 울진 포함), 경상남도
- ⑥ 제주도방언: 제주도

국어국문학회 -



국 어국 문학회

3. 한국어와 대방언권 구획

- (2) 고노 로쿠로 (河野六郎, 1945)
 - 오구라신페이 (小倉進平)의 방언 구획과 유사
 - : 한국어의 역사와 계통에 관심
 - : 국어 방언 자료, 한국어 음운사 논의
 - 5개의 대방언권으로 구분
 - ① 서선(서북)방언: 평안북도, 평안남도
 - ② 동선(동북)방언: 함경북도, 함경남도
 - ③ 중선(중부)방언: 경기도, 충청북도, 충청남도, 강원도
 - ④ 남선(남부)방언: 전라북도, 전라남도, 경상북도, 경상남도
 - ⑤ 제주도방언: 제주도

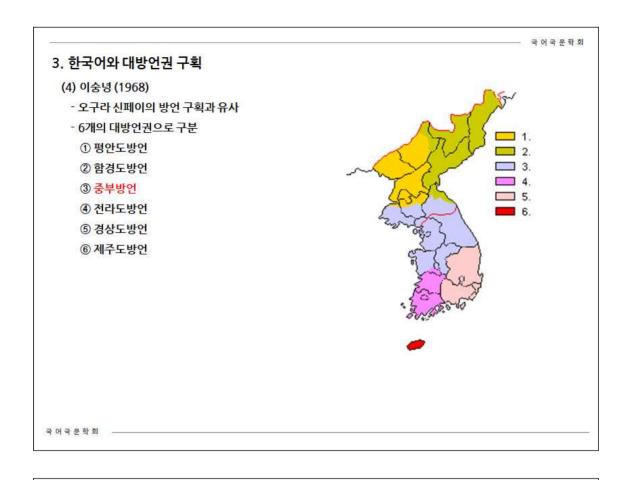
국 어 국 문 학 회

국 어국 문학회

3. 한국어와 대방언권 구획

- (3) 이극로 (1932)
 - 1932년, 〈동광〉 "조선말의 사투리"
 - 한국인에 의한 한국어의 방언 구획
 - 5개의 대방언권으로 구분
 - ① 관서방언: 평안도와 황해도의 일부
 - ② 관북방언: 함경도
 - ③ 중부방언: 경기도와 강원도의 일부
 - ④ 호남방언 : 충청도, 전라도
 - ⑤ 영남방언 : 강원도의 일부, 경상도
 - ※ 제주도 방언에 대한 언급은 없음

국 어국 문학회 -



3. 한국어와 대방언권 구획

(5) 김영황 (1982)

- 1982년, 〈조선어방언학〉

: 북한 학자에 의한 방언 구획 연구

: 음운, 문법, 어휘 고려

- 7개의 대방언권으로 구분

① 서북방언:평안북도,평안남도,자강도 ② 동북방언:함경북도,함경남도,양강도

③ 육진방언: 함경북도 육진지방(경원, 경흥, 부령, 온성, 종성, 회령)

④ 중부방언: 경기도, 강원도, 충청북도, 충청남도, 황해도 일부

⑤ 서남방언: 전라북도, 전라남도 ⑥ 동남방언: 경상북도, 경상남도

⑦ 제주방언: 제주도

국 어 국 문 학 회 _

4. 중부방언의 하위방언구획

- (1) 이병근, 박경래(1988)
 - 경기도 하위방언권 구획
 - ※ 경기도 하위방언구획에 대한 연구는 아직 미흡
 - 어휘 분화를 기준으로 대략적 하위방언권 제시
 - ① 경기 북부
 - ② 경기 남부
 - ③ 서부 도서 지역
 - ※ 박기영(2003)
 - : 역시 개략적이고 인상적인 방언 구획
 - ① 남북 구분
 - → 북부, 서울말과 많은 공통점 / 남부, 충청도 방언과 많은 공통점
 - ② 동서 구분
 - → 동부, 강원도와 충북 방언과 많은 공통점 / 서<mark>부, 서울말과 많은 공통점</mark>
 - ③ 해안도서 지역과 내륙 구분
 - ※ 특수하게 남한산성 일대, 강화, 옹진 등 일부 지역, 고립적인 모습

국 어국 문학회 -

국어국문학회

4. 중부방언의 하위방언구획

- (2) 이익섭 (1981)
 - 강원도 하위방언권 구획
 - 52장의 언어지도(음운, 어휘, 문법)
 - : 등어선마다 차등을 두는 방법에 따라 구획
 - ① 영동 방언권, 성조지역
 - → 북단영동방언(1), 강릉방언(2), 삼척방언(3) 서남영동방언(4)
 - ② 영서방언권(5), 음장지역
 - 경기방언과 유사
 - ※ 김봉국(2001)
 - : 지리적 특성 고려
 - ① 영동 방언권(태백산맥 내부와 그 동쪽)
 - → 성조방언: 삼척, 강릉, 영월 등
 - → 준성조방언: 정선, 평창, 양양, 고성 등 (전이지역)
 - ② 영서 방언권(태백산맥서쪽)
 - → 음장방언: 철원, 화천, 양구, 인제, 춘천, 홍천, 횡성 등

국 어 국 문 학 회 -



4. 중부방언의 하위방언구획

- (3) 강경원(2018)
 - 충청남도하위방언권구획
 - 〈한국언어지도〉(2008)의 어휘자료 활용
 - : 군집분석을 활용한 방언 구획
 - ① 서북방언: 태안, 서산, 당진, 예산, 홍성
 - ② 서남방언: 청양, 보령, 부여, 서천
 - ③ 동북방언: 천안, 아산, 연기
 - ④ 동남방언: 공주, 논산, 대전, 금산
 - ※ 한영목(1999)
 - : 충청남도 방언, 전이지역의 특징이 있으나 독자성 강조
 - ① 1지역: 충청북도와 전라북도에 인접
 - → 천안, 대전, 논산, 공주, 부여 등
 - ② 2지역: 해안지역
 - → 서천, 보령, 홍성, 서산, 태안 등
 - ③ 3지역: 경기도에 인접
 - → 당진, 예산, 아산, 청양 일부 등

국 어국 문학회 -



5. 중부방언 하위방언구획

- (4) 김충회 (1990)
 - 충청북도하위방언권구획
 - 어휘 200개/문법 3개/음운 14개
 - : 등어선마다 차등을 두는 방법에 따라 구획
 - ① 동부 방언권: 강원도와 경북 서북부에 인접
 - → 단양, 제천
 - ② 중부 방언권: 경기도와 충남에 인접
 - → 중원, 괴산, 음성, 진천, 청원
 - ③ 남부 방언권: 충남, 전북, 경북 서남부에 인접
 - → 보은, 옥천, 영동
 - ※ 박경래(1992, 2000)
 - : 김충회(1990)의 방언구획과 유사
 - : 지리적/역사적 조건과의 관련성 강조
 - ① 북부 방언권(고구려): 단양, 제천, 충주, 괴산(연풍면/장연면)
 - ② 남부 방언권(신라): 옥천(군북면제외), 영동, 보은
 - ③ 중부 방언권(백제): 기타 충청북도 지역

국 어 국 문 학 회 -



방언 구획과 네트워크 연구 방법

국어국문학회

1. 방언 구획의 연구 방법

- (1) 최근 방언 구획 연구 흐름
 - 등어선을 대신하여 '거리(distance)' 개념 이용
 - : 언어 변이의 정도를 계량적으로 파악하려는 연구
 - ☞ 방언 간 유사성에 주목
 - 최운호 (2010)
 - : 중국 내에서 사용되고 있는 몽골어파의 16개 방언/언어의 어휘 대응 사전 자료 활용
 - : 레벤시타인 거리 알고리즘 적용, 어휘 간 거리 행렬 측정
 - : 각 방언/언어들 간 유사도 계산, 다변량 분석(multivariate analysis)
 - 정성훈 (2016, 2017, 2018)
 - : 경상도, 전라도, 강원도 하위방언구획
 - → 〈한국언어지도〉(2008)의 언어지도와 어휘항목 153개 이용
 - : 2원 네트워크 분석 적용, 하위방언 간 거리 행렬 측정
 - : 네트워크 최소 단계 거리로 하위방언 간 유사도 계산, 군집 분석(clustering analysis)
 - 강경원(2018)
 - : 충청남도 하위방언구획
 - : 최운호(2010)의 어휘 간 거리 행렬을 이용한 군집 분석

국 어국 문학회 _

2. 방언네트워크 자료와 전처리

- (1) 자료 처리
 - 이익섭외(2008). 〈한국언어지도〉
 - : 북한 지역 제외
 - : 도안형 언어지도
 - → 지도 상, 방언 자료를 부호와 등어선으로 표시
 - : 현지조사 방법, 1985년에 수집 마무리
 - → 군 단위 142개 지점에서 조사
 - : 153개 어휘항목 중 9개 항목 제외 (어휘 X)
 - → 144개 어휘항목, 903개 어휘 수
- (2) 자료의 특징
 - 음성기호로 전사된 자료를 음운 단위로 한글 표시
 - : 한 지역에 2개 이상의 방언형이 사용될때, 모두 반영
 - 해당 방언형이 없거나 조사가 제대로 이루어지지 않은 지역은 'N'으로 표기



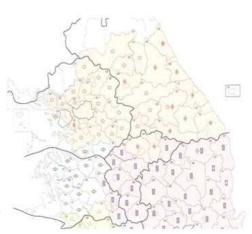
국 어 국 문 학 회

2. 자료 분석

(3) 중부방언의 자료 어휘 수와 조사지역

- 총 903개 어휘 수
- : 베, 나락, …, 세다, 쉬다
- 총 59개 지역(군 단위)
 - : 경기도 지역(총 19개 지역)
 - → 가평, 강화, 고양, 광주, 김포, 남양주, 시흥, 안성, 양주, 양평, 여주, 연천, 옹진, 용인, 이천, 파주, 평택, 포천, 화성
 - : 강원도 지역(총 15개 지역)
 - → 고성, 명주(강릉), 삼척, 양구, 양양, 영월, 원성(원주),인제, 정선, 철원, 춘성(춘천), 평창, 홍성, 화천, 횡성
 - : 충청남도 지역(총 15개 지역)
 - → 공주, 금산, 논산, 당진, 대덕(대전), 보령, 부여, 서산, 서천, 아산, 연기(세종), 예산, 천안, 청양, 홍성
 - : 충청북도 지역(총 10개 지역)
 - → 괴산, 단양, 보은, 영동, 옥천, 음성, 세천, 중원(충주), 진천, 청원(청주)

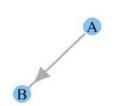
국 어 국 문 학 회 -

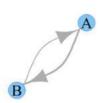


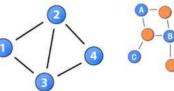
3. 네트워크 분석

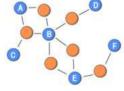
(1) 기본 개념과 네트워크 유형

- 노드(node)와 링크(link)로구성
- 방향성(direction)
- 가중치(weight)
- 1원 네트워크 vs 2원 네트워크
- 전역 네트워크 vs 국부 네트워크
- (2) 네트워크 분석
 - 구조주의에서 시작
- 소쉬르의 언어관 유사
- 네트워크 분석은 연구대상을 구조로 인식
- 개체뿐만 아니라 관계에 주목
- (3) 하위방언 네트워크
 - 방향성 X, 가중치 O, 2원 네트워크









국 어 국 문 학 회

국 어 국 문 학 회

4. 하위방언 네트워크 분석

- (1) 하위방언네트워크분석(sub-dialect network analysis)
 - 심층 하위 방언 구조의 네트워크화
- 2원 네트워크 분석
- (2) 하위방언 네트워크 분석의 장점
 - 방언의 구조를 공간적으로 시각화
- 방언권 체계 안에서 관계적 속성 파악
- 방언권 체계 안에서 하위방언의 위치와 특성을 계량화
- : 하위 방언 안에서도 유사한 정도 측정 가능
- 질적 방법과 양적 방법, 동시 적용 가능
- 하위방언 네트워크, 유사성에 초점
- ※ BUT, 전통적 방언구획, 방언의 차이에 초점
- (3) 네트워크 지표
 - 국부 네트워크, 하위방언의 중심성
 - : 연결(강도) 중심성 ,매개 중심성, 근접 중심성, 고유벡터 중심성, 페이지랭크 중심성
 - 전역 네트워크, 하위방언의 군집 분석
 - : 하위방언간 최소단계거리를 이용한 계층적 군집 분석

국 어국 문학회 -

5. 하위방언 네트워크 분석의 특징

- (1) 방언 간 유사성(similarity)에 중점
 - 등어선을 이용한 방언 구획 → 방언 간 차이에 초점
 - 방언, 언어공동체의 분화
- (2) 네트워크의 관계 속성 중시
 - 인접 지역뿐만 아니라 원거리 방언 간 비교 가능
- 사회 방언 등 다른 언어자료 처리 가능
- (3) 계량화와 시각화 용이
 - 중심 방언과 주변 방언 간 비교
 - 군집 하위방언간 비교
- (4) 다량의 언어자료/네트워크 분석 가능
- (5) 네트워크 추가/삭제 용이
 - 서로 다른 하위방언 네트워크 통합/축소 가능
- (6) 하위방언의 고유성 파악 용이
 - 하위방언 네트워크 간 군집 분석을 통해 중심지/고립지역 등 파악

국 어 국 문 학 회 -

중부방언과 하위방언 네트워크 분석의 실제

- 국어국문학회

1. 네트워크 자료 전처리

- (1) 하위방언 네트워크 자료 작성
 - 지도번호(조사항목), 어휘 수, 하위방언 지역
 - 144개의 어휘항목
 - : 903개의 어휘 수
 - 한 지역에서 중복된 어휘 존재
 - 총 8953개의 방언 자료

순번	지도번호(조사항목)	어휘	하위방언 지역
1	1	베	경기도 연천
2	1	베	경기도 파주
3	1.810	베	경기도 포천
4	11)-	베	경기도 강화
5	1	베	경기도 김포
6	1	볘	경기도 고양
7	1	베	경기도 남양주
8	1	베	경기도 가평
.995		37723	*** 8
8946	153	쉐다	충청북도 단양
8947	153	쉐다	충청북도 청원(청주)
8948	153	쉐다	충청북도 괴산
8949	153	쉐다	충청북도 보온
8950	153	쉐다	충청북도 욕천
8951	153	쉐다	충청북도 영동
8952	153	쉬다	충청남도 보령
8953	153	쉬다	충청남도 금산

국 어국 문학회 -

- 국어국문학회

1. 네트워크 자료 전처리

- (2) 하위방언 지역과 어휘형태수
 - 경기도 지역(총 19개 지역)
 - : 가평-150, 강화-146, 고양-145, 광주-145, *김포-144*, 남양주-153, 시흥-148, 안성-148, 양주-150, 양평-149, 여주-149, *연천-144*, 옹진-145, 용인-148, *이천-154*, 파주-147, 평택-153, 포천-148, *화성-154*
 - 강원도 지역(총 15개 지역)
 - : *고성-149*, 명주(강릉)-157, 삼척-151, 양구-159, 양양-159, 영월-154, *원성(원주)-149*, 인제-159, 정선-150, 철원-155, 춘성(춘천)-152, *평창-172*, 홍성-154, 화천-150, 횡성-157
 - 충청남도지역(총 15개 지역)
 - : 공주-149, *금산-142*, 논산-153, *당진-155*, 대덕(대전)-150, 보령-154, 부여-152, 서산-154, 서천-145, 아산-150, 연기(세종)-152, 예산-154, 천안-149, 청양-152, 홍성-151
 - 충청북도 지역(총 10개 지역)
 - : 괴산-158, 단양-159, 보은-157, *영동-148*, 옥천-152, 음성-152, 제천-151, *중원(충주)-163*, 진천-149, 청원(청주)-155

국 어국 문학회 _

- 국어국문학회

2. 네트워크 행렬 자료

(1) 네트워크 2원 행렬 자료

- 가중치가 없는 2원 행렬 자료 (59*903)

어 <mark>위형태</mark> 하위방언 지역	(나무)함박	(나무)함지	(벳-)뭇단		힌자욱	힌자위	힌창
강원도 고성	0	0	0		0	0	0
강원도 명주	0	0	0	()***()	0	0	0
강원도 삼척	0	0	0	720	0	0	0
강원도 양구	0	0	0	()***()	0	0	0
강원도 양양	0	0	0	2000 at 1 100	0	0	0
강원도 영월	0	0	0	(999)	0	0	0
-	-	: ···			1223		220
충 <mark>청</mark> 북도 옥천	0	0	0	9.000	0	0	0
충청북도 음성	0	0	0	2000.00 C 400 0	0	0	0
충청북도 제천	0	0	0	3275	0	0	0
충청북도 중원	0	0	0	()	0	0	0
충청북도 진천	0	0	0	(27)	0	0	1
충청북도 청원	1	0	0		0	0	1

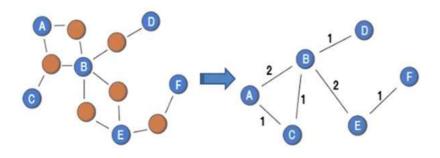
국 어 국 문 학 회

국어국문학회

2. 네트워크 행렬 자료

(2) 네트워크 투사(projection)

- 가중치가 없는 2원 네트워크를 가중치가 있는 1원 네트워크로 변환
- : 2원 네트워크(two-mode network) → 1원 네트워크(one-mode network)
- : 곱행렬 이용
- R 프로그램
 - : tnet package
 - : igraph package



국어국문학회

2. 네트워크 행렬 자료

- (3) 가중치가 있는 1원 네트워크
 - 시각화
 - 중심성 분석 (국부 네트워크 분석)
 - 군집 분석 (전역 네트워크 분석)

국 어 국 문 학 회

3. 네트워크 분석과 시각화
(1) 중부방언의 하위방언 네트워크그림
- 가중치가 있는 1원 네트워크(one-mode network)

국어국무학회

4. 네트워크 중심성 분석

(1) 가중치가 있는 1원 네트워크(one-mode network)

- 가중치(w), 두 방언의 공통된 어휘 수

: 최대값 120 (충청남도 당진-충청남도 홍성)

: 최소값 31 (경기도 고양-강원도 삼척)

번호	하위방언(i)	하위방언(j)	가중치(w)
1	강원도 고성	강원도 명주(강름)	89
2	강원도 고성	강원도 삼척	63
3	강원도 고성	강원도 양구	87
4	강원도 고성	강원도 양양	105
5	강원도 고성	강원도 영월	69
6	강원도 고성	강원도 원성(원주)	63
950	(200)	(ter)	9000
3417	충청북도 청원(청주)	충청북도 영동	67
3418	충청북도 청원(청주)	충청북도 옥천	80
3419	충청북도 청원(청주)	충청북도 음성	85
3420	충청북도 청원(청주)	충청북도 <mark>제</mark> 천	81
3421	충청북도 청원(청주)	충청북도 중원(충주)	81
3422	충청북도 청원(청주)	충청북도 진천	93

국어국문학회

4. 네트워크 중심성 분석

(1) 연결(정도) 중심성

- 노드 간 연결선의 수가 많을수록

: 활동성(activity) 강조

(2) 연결강도 중심성

- 노드 간 연결선과 가중치 고려

: 활동성 강조

(3) (표준) 근접 중심성

- 자신과 다른 모든 노드와의 거리

: 독립성(independence) 강조

(4) 매개 중심성

- 다른 노드 간 연결을 중개/통제하는 역할

: 역으로 네트워크의 연결을 위협하는 역할

(5) 고유벡터 중심성

- 연결된 노드의 연결(정도/강도)까지 고려

: 연결 중심성의 확장판

(6) 페이지랭크 중심성

- 구글 페이지랭크 알고리즘 응용

 $C_d = \sum_{j=1}^g x_{ij}$

 $k^{\alpha}\left(i\right) = \sum_{j=1}^{n} w_{ij}^{\alpha}$

 $C = \frac{\sum\limits_{v_i \in \ V, \, i \, \neq \, j} \frac{1}{d_{ij}}}{N-1}$

 $C_B(i) = \sum_{i \neq v} \sum_{j \neq v_i \neq i} \frac{g_{ivj}}{g_{ij}}$

 $\lambda C_e = AC_e$

 $C_p = \alpha A D^{-1} C_p + \beta 1$

국 어 국 문 학 회

4. 네트워크 중심성 분석

(1) 연결 중심성 분석 결과

하위방언 지역	하위방언 지역 연결 중심성		연결 중심성	하위방언 지역	연결 중심성	
충북 중원(충주)	58	강원 횡성	58	강원 양구	58	
경기 용인	58	충북 청원(청주)	58	강원 화천	58	
경기 화성	58	충남 연기(세종)	58	강원 철원	58	
경기 가평	58	경기 이천	58	강원 인제	58	
충남 아산	58	경기 <mark>양</mark> 주	58	충남 논산	58	
충남 천안	58	경기 시흥	58	충북 옥천	58	
충북 괴산	58	강원 <mark>홍</mark> 성	58	충남보령	58	
경기 광주	58	충남공주	58	충북 보은	58	
충북 제천	58	충남청양	58	강원 양양	58	
충북 음성	58	충남 당진	58	충남 서천	58	
경기 양평	58	충남 부여	58	경기 김포	58	
경기 포천	58	경기 파주	58	충북 단양	58	
충북 진천	58	경기 연천	58	충북 영동	58	
경기 안성	58	충남홍성	58	강원 고성	58	
경기 평택	58	강원 원성(원주)	58	강원 영월	58	
경기 남양주	58	충남 서산	58	충남 금산	58	
경기 여주	58	경기 옹진	58	강원 명주(강릉)	58	
충남 예산	58	강원 평창	58	강원 정선	58	
강원 춘성(춘천)	58	경기 고양	58	강원 삼척	58	
충남 대덕(대전)	58	경기 강화	58			

국 어국 문학회 ㅡ

국어국문학회

4. 네트워크 중심성 분석

(2) 연결강도 중심성 분석 결과

하위방언 지역	연결강도 중심성	하위방언 지역	연결강도 중심성	하위방언 <mark>지</mark> 역	연결강도 중심성
충북 중원(충주)	9098	강원 횜성	8506	강원 양구	7886
경기 용인	8920	충북 청원(청주)	8488	강원 화천	7866
경기 화성	8908	충남 연기(세종)	8462	강원 철원	7854
경기 가평	8906	경기 이천	8448	강원 인제	7830
충남 아산	8752	경기 양주	8340	충남논산	7804
충남 천안	8752	경기 시흥	8324	충북옥천	7748
충북 괴산	8752	강원 홍성	8320	충남보령	7700
경기 광주	8748	충남공주	8196	충북 보은	7654
충북 제천	8726	충남 청양	8158	강원 양양	7598
충북 음성	8714	충남 당진	8114	충남 서천	7442
경기 양평	8654	충남 부여	8104	경기 김포	7406
경기 포천	8638	경기 파주	8068	충북 단양	7272
충북 진천	8616	경기 연천	8042	충북 영동	6942
경기 안성	8610	충남 홍성	8022	강원 고성	6770
경기 평택	8590	강원 원성(원주)	7996	강원 명월	6732
경기 남양주	8578	충남 서산	7996	충남금산	6670
경기 여주	8578	경기 옹진	7968	강원 명주(강릉)	6502
충남 예산	8564	강원 평창	7952	강원 정선	5872
강원 춘성(춘천)	8560	경기 고양	7946	강원 삼척	5618
충남 대덕(대전)	8510	경기 강화	7898		
		250 10/200 (20)			

국 어국 문학회 —

국 어 국 문 학 회

4. 네트워크 중심성 분석

(3) 근접 중심성 분석 결과

하위방언 지역	근접 중심성	하위방언 지역	근접 중심성	하위방언 지역	근 <mark>접</mark> 중심성
강원 삼척	0,000181	경기 시흥	0,000134	충남 대덕(대전)	0.000124
강원 정선	0,000175	경기 파주	0,000133	경기 평택	0,000124
강원 명주(강릉)	0,000157	충남공주	0,000133	경기 광주	0.000124
충남금산	0,000151	강원 양양	0,000132	경기 양평	0,000123
강원 영월	0,000151	충북보은	0,000132	충남 천안	0.000123
강원 고성	0,000149	충북 옥천	0,000131	경기 이천	0,000122
충북 영동	0,000145	경기 양주	0,000129	경기 화성	0,000122
충남 서천	0,000144	강원 철원	0,000128	경기 여주	0.000122
경기 김포	0,000141	경기 포천	0,000128	강원 홍성	0,000121
충남보령	0,000141	충남 예산	0.000128	경기 용인	0,000120
충남 홍성	0,000139	강원 인제	0,000128	충북 진천	0.000119
경기 고양	0.000139	강원 화천	0,000128	충북 청원(청주)	0,000118
충북 단양	0,000138	강원 평 창	0,000128	강원 횡성	0.000118
경기 연천	0,000137	강원 양구	0.000127	강원 춘성(춘천)	0.000117
충남 청양	0,000137	충남 연기(세종)	0,000127	경기 가평	0.000117
충남 서산	0,000136	경기 옹진	0,000127	충북 음성	0.000115
충남 당진	0,000136	경기 안성	0,000126	충북 제천	0,000115
충남 부여	0,000135	충남 아산	0,000126	충북 괴산	0,000114
충남 논산	0.000135	경기 남양주	0,000126	충북 중원(충주)	0,000110
경기 강화	0,000134	강원 원성(원주)	0,000125		

국 어국 문학회 —

국 어국 문학회

4. 네트워크 중심성 분석

(4) 매개 중심성 분석 결과

하위방언 지역	매개 중심성	하위방언 지역	<mark>매개</mark> 중심성	하위방언 지역	매개 중심성
강원 삼척	177.83	충북 <mark>제</mark> 천	0	경기 연천	0
강원 정선	125,83	경기 포천	0	충남부여	0
충남금산	12,50	경기 안성	0	충남 서산	0
충남홍성	5,00	경기 평택	0	경기 옹진	0
경기 김포	4,50	충북 진천	0	강원 원성(원주)	0
경기 고양	3,50	경기 남양주	0	강원 철원	0
충남청양	2,00	경기 여주	0	강원 화천	0
경기 강화	2,00	충남 예산	0	강원 양구	0
충남보령	1.00	강원 춘성(춘천)	0	강원 평창	0
충남 서천	1,00	충남 대덕(대전)	0	충남 논산	0
충북 중원(충주)	0	충북 청원(청주)	0	강원 인제	0
경기 용인	0	경기 이천	0	충북 옥천	0
경기 화성	0	충남 연기(세종)	0	충북보은	0
경기 가평	0	강원 횡성	0	강원 양양	0
경기 광주	0	경기 시흥	0	충북 단양	0
충남 천안	0	경기 양주	0	충북 영동	0
충남 아산	0	강원 홍성	0	강원 고성	0
충북 괴산	0	충남공주	0	강원 영월	0
충북음성	0	충남 당진	0	강원 명주(강름)	0
경기 양평	0	경기 파주	0		

국 어 국 문 학 회 —

국 어 국 문 학 회

4. 네트워크 중심성 분석

(5) 고유벡터 중심성 분석 결과

하위방언 지역	하위방언지역 고유벡터 중심성		고유벡터 중심성	하위방언지역	고유벡터 중심성	
충북 중원(충주)	1,000000	충북 청원(청주)	0,935282	강원 화천	0,86339	
경기 용인	0.988050	경기 이천	0,934879	강원 양구	0,861735	
경기 화성	0,986449	충남 연기	0,932833	강원 평창	0,860502	
경기 가평	0,984317	강원 횡성	0,931872	충남논산	0,856823	
경기 광주	0,970238	경기 시흥	0,925469	강원 인계	0,852042	
충남 천안	0,967112	경기 양주	0,924691	충남보령	0,850491	
충남 아산	0,966877	강원 <mark>홍성</mark>	0,912150	충북옥천	0,844742	
충북 괴산	0,960125	충남공주	0,903137	충북 보은	0.837091	
충북 음성	0,959563	충남 청양	0,900627	강원 양양	0,828443	
경기 양평	0,959289	충남 당진	0,895957	경기 김포	0,821867	
충북 제천	0,958151	경기 파주	0,894007	충남 서천	0.819834	
경기 포천	0,957603	경기 연천	0,892535	충북 단양	0,797104	
경기 안성	0,955710	충남부여	0,891884	충북 영동	0.759112	
경기 평택	0,951392	충남홍성	0,886016	강원 고성	0,738149	
충북 진천	0,951036	충남 서산	0,882998	충남금산	0.731705	
경기 남양주	0,950651	경기 고양	0,882480	강원 명월	0,729378	
경기 여주	0,949829	경기 옹진	0,881350	강원 명주(강름)	0.704198	
충남 예산	0,944488	강원 원성(원주)	0,879810	강원 정선	0,634915	
강원 춘성(춘천)	0,939541	경기 강화	0,876087	강원 삼척	0,608520	
충남 대덕(대전)	0.936449	강원 철원	0,865041			

국 어국 문학회 —

국 어국 문학회

4. 네트워크 중심성 분석

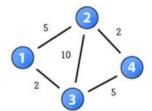
(6) 페이지 랭크 중심성 분석 결과

	하위방언 지역	페이지랭크 중심성	하위방언 지역	페이지랭크 중심성	하위방언 지역	페이지랭크 중심성
E .	충북 중원(충주)	0,01884	충북 청원(청주)	0.01774	강원 화천	0,01664
	경기 용인	0,01850	경기 이천	0,01766	강원 양구	0,01668
	경기 화성	0,01848	충남 연기(세종)	0,01769	강원 평창	0,01683
	경기 가평	0,01848	강원 횡성	0,01779	충남논산	0,01652
	경기 광주	0,01819	경기 시흥	0.01742	강원 인제	0,01659
	충남 천안	0,01820	경기 양주	0.01746	충남보령	0,01632
	충남 아산	0,01820	강원 홍성	0.01745	충북 옥천	0.01644
	충북 괴산	0,01822	충남공주	0.01721	충북보은	0,01626
	충북 음성	0,01815	충남 청양	0.01714	강원 양양	0,01617
	경기 양평	0,01802	충남 당진	0,01706	경기 김포	0,01579
	충북 제천	0,01818	경기 파주	0.01697	충남 서천	0,01587
	경기 포천	0,01799	경기 연천	0,01692	충북 단양	0,01558
	경기 안성	0,01794	충남 부여	0,01705	충북 영동	0,01499
	경기 평택	0.01791	충남홍성	0,01690	강원 고성	0,01469
	충북 진천	0,01796	충남 서산	0.01685	충남금산	0,01449
	경기 남양주	0,01789	경기 고양	0,01675	강원 영월	0,01463
	경기 여주	0,01789	경기 옹진	0,01680	강원 명주(강름)	0,01422
	충남 예산	0.01787	강원 원성(원주)	0,01686	강원 정선	0,01309
	강원 춘성(춘천)	0,01788	경기 강화	0,01667	강원 삼척	0,01263
	충남 대덕(대전)	0,01778	강원 철원	0,01661		.V an an an an

국어국문학회 —

5. 하위방언 네트워크 군집 분석

- (1) 네트워크 군집 분석
 - 노드들 사이에는 상대적 최소 단계 거리를 가정하며, 군집의 수와 내용, 구조 등이 알려져 있지 않은 상태에서 그 특성과 관계를 파악하기 위한 일종의 탐색적인 분석방법
- (2) 최소 단계 거리(shortest paths)
 - 거리(distance) 개념
 - 연결 가중치 역수
- (3) 두 노드 간 최소 단계 거리
 - 두 노드 간 연결 가중치 역수의 총합이 최소가 되는 경로
 - 최단거리 경로: " 1-2-3-4"
 - 최단거리: 1/5 + 1/10 + 1/5 = 0,5



국 어 국 문 학 회

- 국어국문학회

5. 하위방언 네트워크 군집 분석

- (4) 하위방언 네트워크의 최소 단계 거리(59*59)
 - 최대값 15,49193(경기도 고양-강원도 삼척), 최소값 8,00000(충남 당진-충남 홍성)

	강원 고성	강원 명주 (강름)	강원 삼척	강원 양구		충북 음성	충북 제천	충북 중원 (충주)	충북 진천
강원 명주(강릉)	11,31371								
강 <mark>원</mark> 삼척	13,26650	12,24745				3			
강원 양구	11,57584	12,56981	13,56466						
강원 양양	9,89950	12,00000	13,41641	12,16553					
강원 영월	12,84523	11,53256	12,20656	12,68858					
강원 원성(원주)	13,11488	12,96148	13,56466	12,40967		3			
***			222						
충북옥천	14,1775	13,96424	13,52775	13,67479	•••				
충북 음성	13,3791	13,67479	14,03567	13,22876					
충북 제천	13,1909	12,88410	13,63818	12,64911		11,18034			
충북 중원(충주)	13,4164	13,26650	14,14214	13,19091	***	10,72381	10,39230		
충북 진천	13,8564	14,28286	14,69694	13,85641	•••	9,11043	11,66190	11,83216	
충북 청원(청주)	14,3527	14,83240	14,89966	14,00000	•••	11,78983	12,08305	12,72792	10,86278

국어국문학회 _

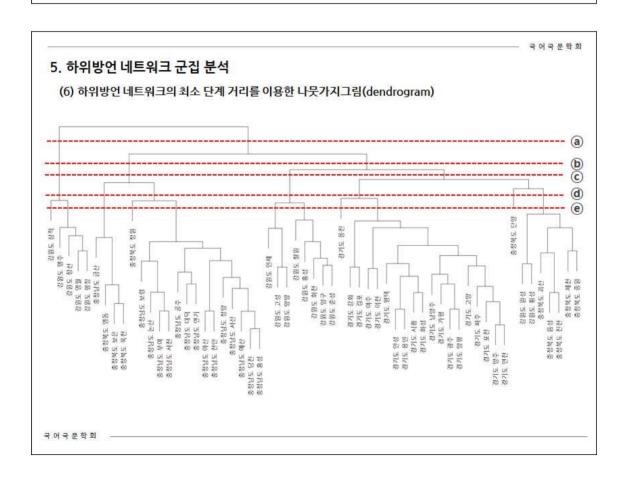
5. 하위방언 네트워크 군집 분석

- (5) 계층적 군집 분석
 - 임의의 노드에서 가까운 노드와 결합하면서 하나의 군집을 형성할 때까지 결합하는 방식
 - 계층적 군집 분석 방법에서 가장 적합한 군집분석 방법
 - : 평균 거리 연결 방법, 최장 거리 연결 방법, 최소 거리 연결 방법, ward 거리 연결 방법 등
 - : '평균 거리 연결 방법'(최운호 2010, 정성훈 2016, 2017, 2018)
 - 방언의 군집 분석, 평균 거리 연결 방법이 가장 적합
 - : 평균 거리 연결 방법, 두 노드(군집) 사이의 거리를 각 군집(노드)에 속하는 모든 개체들의 평균거리로 정의하여 가장 유사성이 큰 군집으로 묶어 나가는 방법

군집 분석 방법	cophenetic 상관계수
평균 거리 연결 방법	0,84184
최장 거리 연결 방법	0,79749
최소 거리 연결 방법	0,78147
WARD 연결 방법	0,63467
중심 연결 방법	0,58639

$$d_{average}\left(X,\,Y\right)\,=\,\frac{1}{\left(N_{\!1}\times N_{\!2}\,\right)\sum_{i}\sum_{j}d_{ij}}$$

국어국문학회



구어구무하히

6. 중부방언 방언구획과 하위방언 군집 분석 결과

- (1) 중부방언의 1차 하위방언(@)
 - 2개의 군집 형성
 - ① 강원도 영동 지역(삼척, 명주(강릉), 정선, 영월, 평창)
 - ② 기타 중부 지역
- (2) 중부방언의 2차 하위방언(b)
 - 3개의 군집 형성
 - ① 강원도 영동 지역 (삼척, 명주(강릉), 정선, 영월, 평창)
 - ② 충청남도 + 충청북도 청원(청주), 옥천, 보은, 영동
 - ③ 경기도 + 강원도 영서 지역 + 충청북도(청원(청주), 옥천, 보은, 영동 제외)
- (3) 중부방언의 3차 하위방언(©)
 - 4개의 군집 형성
 - ① 강원도 영동 지역 (삼척, 명주(강릉), 정선, 영월, 평창)
 - ② 충청남도 + 충청북도 청원(청주), 옥천, 보은, 영동
 - ③ 강원도 영서 지역(원성(원주), 횡성제외)
 - ④ 경기도 + 충청북도(청원(청주), 옥천, 보은, 영동 제외/원성(원주), 횡성 포함)

국 어국 문학회 -

국 어 국 문 학 회

6. 중부방언 방언구획과 하위방언 군집 분석 결과

- (4) 중부방언의 4차 하위방언(@)
 - 7개의 군집 형성
 - ① 강원도 영동 지역 (삼척, 명주(강릉), 정선, 영월, 평창)
 - ② 충청남도 금산 + 충청북도 옥천, 보은, 영동
 - ③ 충청남도(금산 제외) + 충청북도 청원(청주)
 - ④ 강원도 영서 지역(원성(원주), 횡성 제외)
 - ⑤ 경기도
 - ⑥ 충청북도 단양
 - ⑦ 강원도 원성(원주), 횡성 + 충청북도 괴산, 음성, 중원(충주), 제천, 진천

국 어국 문학회 _

6. 중부방언 방언구획과 하위방언 군집 분석 결과

- (5) 중부방언의 5차 하위방언(®)
 - 11개의 군집 형성
 - ① 강원도 삼척
 - ② 강원도 영동 지역 (명주(강릉), 정선, 영월, 평창)
 - ③ 충청남도 금산 + 충청북도 옥천, 보은, 영동
 - ④ 충청북도 청원(청주)
 - ⑤ 충청남도(금산 제외)
 - ⑥ 강원도 고성, 양양, 인제
 - ⑦ 강원도 춘성(춘천), 양구, 화천, 홍성, 철원
 - ⑧ 경기도 옹진
 - ⑨ 경기도(웅진 제외)
 - ⑩ 충청북도 단양
 - ⑪ 강원도 원성(원주), 횡성 + 충청북도 괴산, 음성, 중원(충주), 제천, 진천

국 어 국 문 학 회

국어국문학회

6. 중부방언 방언구획과 하위방언 군집 분석 결과

- (6) 중부방언하위방언(©)
 - 총 4개의 군집 형성
 - ① 강원도 영동 지역
 - : 삼척, 명주(강릉), 정선, 영월, 평창
 - ②충청남도

충청북도 남부

- : 청원(청주), 옥천, 보은, 영동
- ③ 강원도 영서 지역
- : 원성(원주), 횡성 제외
- ④ 경기도

충청북도북부

- : 청원(청주), 옥천, 보은, 영동 제외
- : 원성(원주), 횡성 포함

EB VEI A VB EVI EEF VVI Number of components

※ 매개 변수화 된 가우스 혼합 모델

: 초기화된 기대 최대값을 위해 베이지안 정보 기준에 따라 최적의 모델 및 군집 수를 결정하는 모델

국어국문학회

구어구무하히

7. 요약 및 결론

- (1) 중부방언의 하위방언 네트워크
 - 어휘를 활용한 방언 구획
 - 네트워크 중심성 분석
 - : 중심부 방언(충주방언), 주변부 방언(삼척방언)
 - 네트워크 군집 분석
 - : 크게 4개 하위방언권으로 분류

(2) 결론

- 중부방언은 '경기도, 강원도, 충청남북도'라는 행정구역으로 구획하기 어려움
 - : 경기도와 충청북도 북부, 강원도 원성(원주), 횡성을 포함하는 지역
 - : 강원도 영동 지역
 - : 강원도 영서 지역(원성(원주), 횡성 제외)
 - : 충청남도와 충청북도 남부를 포함하는 지역
 - 🤛 충청북도 지역, 한국어 중남부 방언의 대(大)전이지역으로 보임
- 어휘 자료뿐만 아니라 음운, 형태, 문법 등의 언어자료를 활용한 하위방언 구획 필요성
 - ☞ 1980년대 방언 자료 → 현재의 변화를 반영한 방언 자료 필요
 - : 통시적으로 방언의 변화가 어떻게 일어나는지도 확인 가능

국 어국 문학회 -



"중부방언 방언구획과 네트워크 분석"에 대한 토론문

김한샘(연세대)

2000년대에 국내에서 디지털 인문학의 개념이 논의되기 시작한 이후로 디지털화된 정보를 활용한 인문학적 관점의 네트워크 분석이 점차 늘어나고 있으나, 역사학, 사회학 등의 학문 분야에 집중되어 왔습니다. 국어국문학 분야로 범위를 좁혀 보아도 주로 문학 분야에서 네트워크 분석 방법론을 활발하게 활용해 왔으며, 국어학 분야에서는 국어학의 세부 영역에 대한 연구동향 분석, 빅데이터나 사전 엔트리에 기반한 어휘 관계 분석 등이 주를 이루었습니다. 이에 최근 들어 발표자 선생님이 국어학 분야에서의 네트워크 분석의 적용 범위를 언어의 지역 레지스터 연구로 확장하여 방언 연구에 적극적으로 활용하고 계신 것이 반갑습니다. 토론을 맡았으나방언 전문가가 아니기에 발표 자료를 바탕으로 몇 가지 질문을 드리는 것으로 토론을 대신하도록 하겠습니다.

1. 네트워크 분석의 대상 확장 가능성

연구의 분석 대상인 방언 정보 원자료는 이익섭 외(2008) <한국언어지도>입니다. 이는 한국 정신문화연구원에서 1987년에서 1995년에 걸쳐 출간한 <한국방언자료집>을 지도화한 것이고 발표 자료에 있는 것처럼 현지 조사는 1985년에 마무리되었으니 21쪽의 하위 방언 네트워크 자료는 최소한 35년 전의 정보를 담고 있는 것입니다. 현지 조사에 기반한 방언 연구가 매우 전문적인 영역이고 현재 활성화되어 있지는 않지만 한국정신문화연구원의 대규모 조사 이후에도 국립국어원 등의 기관이나 개인 및 연구 그룹에서 방언 연구를 진행한 바 있습니다.

1985년 이후의 방언 조사 결과를 향후 연구에서 네트워크 분석의 대상으로 포함하는 것이 가능할지, 만약 가능하다면 구체적으로 어떻게 서로 다른 조건에서 조사된 결과들을 통합할 수 있을지 궁금합니다. 이러한 연구 대상의 확대가 불가능하다면 최신의 방법론을 활용한다고 하더라도 국내 방언 연구가 1980년대 방언 자료에서 벗어나기 힘들다는 근본적인 한계를 가지게됩니다. 더불어 상대적으로 최근 시기의 방언 자료를 연구 대상으로 포함할 수 있다면 한국어방언에 대해 방법론적으로 공시적 분석, 혹은 개별 공시적 분석에 기반한 방언 변화 연구에서한걸음 나아가 자료 기반 시계열 분석을 시도할 가능성은 없을지에 대해서도 고견을 듣고 싶습니다.

2. 방언 조사 단위와 행정구역

발표자께서는 자료에 기반한 네트워크 군집 분석을 통해 방언, 특히 중부방언은 경기도, 강원도, 충청남북도라는 행정구역으로 구획하기는 어렵다는 결론을 제시하셨습니다. 전통적으로 방언 구획은 행정구역의 경계를 전제로 하여 연역적으로 접근했던 것이 사실입니다. 그런데 방언을 행정구역으로 구획하기 어렵다는 결론을 이끈 이 연구에서 기준으로 삼은 142개 조사 지점도 군 규모의 행정구역을 단위로 하고 있습니다. 섬진강 유역처럼 교류가 활발하거나 반대로산맥 관통 등으로 물리적인 교류 단절이 일어나는 곳 등에서는 군 규모의 행정구역이 방언 구

획에 있어 유효하지 않을 수도 있을 듯합니다. 주어진 자료의 계량적 분석을 수행하시는 과정에서 방언 조사 지점 및 단위의 설정에 대한 시사점이 혹시 있지 않았는지 여쭈어 봅니다.

3. 네트워크 분석의 시각화

네트워크 분석 방법론의 큰 장점 중 하나는 규모가 큰 데이터를 분석한 결과를 중요한 정보가 잘 드러나도록 제시할 수 있다는 것이고, 이를 통해 데이터 전문가가 아니더라도 연구 결과를 이해하고 관련 연구를 진행하는 데에 인사이트를 얻을 수 있습니다. 물론 최소 단계 거리를 이용한 나뭇가지그림(dendrogram)을 통해 계층 관계가 있는 하위방언권의 관계 정보를 볼 수 있지만 기본 네트워크를 나타낸 26쪽의 가중치가 있는 1원 네트워크(one-mode network)도 담고 있는 정보가 잘 드러나게 시각화할 것을 제안합니다. 중요하지 않은 관계들을 제거해 전 반적이고 거시적인 연결 구조를 시각화한다면 네트워크 분석 결과를 더 효과적으로 표현하고 이해할 수 있지 않을까 합니다.

발표를 듣기 전에 자료만으로 토론문을 작성하다 보니 필자의 의도를 완벽히 파악하지 못하여 내용을 잘못 이해하였을 수 있습니다. 너그럽게 이해해 주시면 감사하겠습니다.