

K^A_LS 한국언어과학회 2023년 겨울학술대회

언어과학의 새로운 미래

일시 2023년 2월 16일(목)

장소 온라인(Zoom)

주관 부산대학교 언어정보학과

주최 K^A_LS 한국언어과학회

[한국언어과학회 2023 겨울학술대회] 온라인 ZOOM 접속 정보

회의 주소

ZOOM주소: <https://pusan.zoom.us/s/92639892323>

회의 ID: 926 3989 2323

암호: 123456

한국언어과학회 2023 겨울학술대회

“언어과학의 새로운 미래”



일시: 2023년 2월 16일(목)

장소: 온라인(Zoom)

주관: 부산대학교 언어정보학과

주최: 한국언어과학회

12:40 ~ 13:00		개회사 및故안동환교수님 추도 묵념		회장	김재운(신라대)
시간	분과	1분과(형태/통사/의미)	2분과(코퍼스)	3분과(코퍼스/응용)	4분과(응용/언어교육)
		좌장(사회): 이원빈(대전대)	좌장(사회): 황규홍(동아대)	좌장(사회): 권미분(신라대)	좌장(사회): 전지현(부산대)
13:00 ~ 13:30	제목	Form Sequence & Coordinate Structure Constraint	Keyword analysis of the IMO Standard Marine Communication Phrases [ⓔ]	코퍼스 기반 한국수어의 기초 어휘 비교 연구	Voices of Constructing Dialogues in English Cross-Cultural Communication
	발표	권기양(영산대)	Wang, Meihan(한국해양대)	장세은 & 박호민(한국해양대)	송경숙(동의대)
	토론	전혜원(부경대)	Zhang, Guandong(한국해양대)	권순복(부산대)	김규미(안동대)
13:30 ~ 14:00	제목	Linguistic synaesthesia in language contact: A Sino-Korean compound perspective	An Exploration of Metaphorical Expressions in Maritime English from the Perspective of Keywords	수어코퍼스 기반 청각장애 초등학생 교과서 개발의 제시 모델	Instructing how to use Google Translate effectively for research writing [ⓔ]
	발표	Jo, Charmhun(College of International Studies, Southwest University, China)	Xu, Mengyu (한국해양대)	권순복(부산대), 장세은(한국해양대), 최영주(조선대)	허영현(University of Aizu, Japan)
	토론	박기성(부산대)	Liu, Cuilin(한국해양대)	안병길(경상대)	우길주(부산교대)
14:00 ~ 14:30	제목	A Study on the Levels of Grammaticalization of the New Degree Adverbs in Chinese	Lexical bundles in the academic writing of Vietnamese EFL learners [ⓔ]	Gesture and/to Sign in Language	A Study of Common Errors in Chinese EFL Students' Writing [ⓔ]
	발표	Ren, Yang(부산대)	Bui, Thuy Quynh(한국해양대)	장세은(한국해양대)	Du, Jiaqi(동아대)
	토론	송효원(부산외대)	Zhang, Guandong(한국해양대)	최영주(조선대)	신유선(순천대)
14:30 ~ 15:00	제목	Gapping in the Conjoined Structure	From Gender-Biased to Gender-Specific and Gender-Inclusive Language: A Corpus-Based Study on Maritime Conventions and Contemporary English [ⓔ]	'쉬운-한국어(Plain-Korean)'를 통한 사회적 소통의 인문사회과학적 연구	L2 언어 인지·발화 과정에서의 말하기 조정에 대한 성별의 영향
	발표	김희태(계명대)	Liu, Cuilin(한국해양대)	이상수(부경대)	서미교(부산대)
	토론	허세문(한국외대)	Xu, Mengyu(한국해양대)	서민정(부산대)	장하연(부산외대)
15:00 ~ 15:30	제목	중세국어 '곧'과 '다'류 동등비교구문의 정도성 표현 연구 - '이'형 구문과 '와'형 구문을 중심으로	A Hybrid Approach to Keywords Analysis of Aviation Accident Investigation Reports Corpus [ⓔ]	한국어 교육을 위한 한국어 피동 표현 연구 -일본어 피동 표현과의 대조 분석을 중심으로-	CAF에 나타난 한국대학생의 말하기 특징
	발표	윤보영(신라대)	Yang, Qi, 박호민 & 장세은(한국해양대)	김미루(부산대)	신순란(부산대)
	토론	김태우(부산대)	Liu, Cuilin(한국해양대)	김중수(부산대)	김정윤(부산대)
15:30 ~ 16:00	제목	Chinese '(shi)' plus focuses and Search constraint at Determinacy	Lexical bundles Analyses across Maritime English Legal Genres [ⓔ]	중학교 영어 고능력 학습자 맞춤형 프로그램의 효과-창의성과 코메트릭스 언어분석 중심으로	Lexical semantic analysis by means of artificial intelligence
	발표	임익희(한국외대)	Zhang, Guandong(한국해양대)	이주영(한국교원대)	김원빈(University of Georgia)
	토론	정현경(신라대)	Bui, Thuy Quynh(한국해양대)	신운섭(고려대)	장세은(한국해양대)
16:00 ~ 17:00	초청 특강	김정렬(한국교원대): 외국어 언어교육사에서 내용학과 교육학의 상호작용 사회: 서민정(부산대), 토론: 전지현(부산대)			
17:00 ~ 17:30	연구윤리교육: 편집위원장(전지현, 부산대)				
17:30 ~ 18:00	학술상 시상식				
	임시총회 진행: 총무이사(서민정, 부산대)				

ⓔ = Presentation in English

Contents

메인세션 (개회식, 초청특강, 연구윤리교육 및 총회)

- 외국어 언어교육사에서 내용학과 교육학의 상호작용 3
/ 김정렬

1분과 (형태/통사/의미)

- Form Sequence & Coordinate Structure Constraint 19
/ 권기양
- Linguistic synaesthesia in language contact: A Sino-Korean compound perspective 25
/ Jo, Charmhun
- A Study on the Level of Grammaticalization of the New Degree Adverbs in Chinese 39
/ Ren, Yang
- Gapping in the Conjoined Structure 51
/ 김희태
- 중세국어 ‘곧 허다’류 동등비교구문의 정도성 표현 연구 - ‘이’형 구문과 ‘와’형 구문을 중심으로 55
/ 윤보영
- Chinese ‘(shi)’ plus focuses and Search constraint at Determinacy 69
/ 임익희

2분과 (코퍼스)

- Keyword analysis of the IMO Standard Marine Communication Phrases 81
/ Wang, Meihan
- An Exploration of Metaphorical Expressions in Maritime English from the Perspective of Keywords 91
/ Xu, Mengyu
- Lexical bundles in the academic writing of Vietnamese EFL learners 105
/ Bui, Thuy Quynh
- From Gender-Biased to Gender-Specific and Gender-Inclusive Language: A Corpus-Based Study on Maritime Conventions and Contemporary English 121
/ Liu, Cuilin
- A Hybrid Approach to Keywords Analysis of Aviation Accident Investigation Reports Corpus 133
/ Yang, Qi & 박호민 & 장세은
- Lexical bundles Analyses across Maritime English Legal Genres 153
/ Zhang, Guandong

3분과 (코퍼스/응용)

- 코퍼스 기반 한국수어의 기초 어휘 비교 연구 169
/ 장세은 & 박호민
- 수어코퍼스 기반 청각장애 초등학생 교과서 개발의 제시 모델 195
/ 권순복 & 장세은 & 최영주
- Gesture and/to Sign in Language 209
/ 장세은
- ‘쉬운-한국어(Plain-Korean)’를 통한 사회적 소통의 인문사회과학적 연구 251
/ 이상수
- 한국어 교육을 위한 한국어 피동 표현 연구 -일본어 피동 표현과의 대조 분석을 중심으로- 261
/ 김미루
- 중학교 영어 고능력 학습자 맞춤형 프로그램의 효과 275
/ 이주영

4분과 (응용/언어교육)

- Voices of Constructing Dialogues in English Cross-Cultural Communication
/ 송경숙
- Instructing how to use Google Translate effectively for research writing 289
/ 허영현
- A Study of Common Errors in Chinese EFL Students'Writing 307
/ Du, Jiaqi
- L2 언어 인지·발화 과정에서의 말하기 조정에 대한 성별의 영향
/ 서미교
- CAF에 나타난 한국 대학생의 말하기 특징 323
/ 신순란
- Lexical Semantic Analysis by means of Artificial Intelligence 349
/ 김원빈

메인세션

개회식, 초청특강, 연구윤리교육 및 총회

외국어 언어교육사에서 내용학과 교육학의 상호작용

김정렬

(한국교원대학교)

I. 서론

언어는 인간이 의미를 이해하고 생성해내는 하나의 형태적 시스템이다. 언어과학은 언어의 형태적 시스템의 정체를 밝혀내고 형태와 의미의 관계를 탐색해온 인문학의 핵심 학문이다. 언어교육은 학생들이 배우고자 하는 언어의 형태적 시스템을 습득하여 의미를 이해하고 생성해내는 도구로 활용할 수 있도록 하는 언어능력을 배양하는 것을 돕는 학문이다. 이를 위해서는 언어과학의 연구 결과로 밝혀진 언어의 형태적 시스템을 먼저 제시하고 관련 언어자료를 경험하게 하는 연역적 방법과 언어자료를 먼저 경험하고 목시적으로 언어의 시스템을 습득하게 하는 귀납적 방법이 있다.

언어교육에서 목표어로 사용된 언어에는 고전 그리스어나 고전 라틴어와 같은 사어도 있고 언어교육 시기와 같은 시대에 상용되고 있는 살아있는 언어도 있다. 사어의 경우는 언어의 어휘나 발음, 문법이 변하지 않는 일종의 폐쇄적 시스템이고 동시대에 쓰이는 구어를 경험할 수 있는 살아있는 언어의 경우는 어휘, 발음, 문법이 변하는 일종의 열린 시스템이다. 그리고 목표어가 고전어들과 같이 형태적 시스템이 충분히 기술된 언어들이 있는가 하면 2차대전 당시 미군이 태평양 각 군도에서 경험했던 언어학적으로 해당 언어의 시스템이 기술되지 않은 언어도 있다. 폐쇄적 시스템의 언어는 주로 문어의 형태로 남아 있어서 이런 경우는 문어적 의미의 생성보다는 문어의 이해를 위해서 번역하는 방법이 주로 사용되었다. 이에 반해 살아있는 언어는 문어와 달리 구어를 중심으로 화자가 의미를 전달하기 위해서 말하기 중심의 언어교육이 이루어졌고 초기 구어교육의 형태는 고전어 문법의 시스템을 활용하여 살아있는 언어를 기술하고 해당 언어의 형태적 시스템을 말하기에 필요한 문형의 형태로 만들어서 연습하도록 하였다.

본 연구에서는 외국어 교육사를 통해서 현재 대학 영어교육과의 교수요목 영역에서 대체로 세 분야로 이루어진 어학과 문학을 내용학이라고 볼 때에 교육학과의 상호작용이 어떻게 이루어졌는지 살펴보고자 하는 것이다. 이를 위해서 외국어 교육방법을 크게 문헌학적 방법, 언어과학적 방법, 상호작용적 방법으로 나누어 볼 수 있다. 문헌학적 방법은 고전어 문법교육이 지배하던 시기로 문법과 번역방식의 연역적 언어교육이 이루어지던 시기이다. 이 시기에는 최초로 William Bullokar에 의해서 1586년에 발간된 Pamphlet for Grammar라는 영문법 책은 William Lily의 Rudimenta Grammatices의 라틴어의 형태적 분류방식을 그대로 답습하여 영문법을 기술하였다. Lily는 1486년에 Oxford대학을 졸업하고 예루살렘 성

지순례에서 돌아오면서 이탈리아에서 르네상스 운동을 직접 경험한 당시 영국에서 가장 유명한 고전어 학자였다. 따라서 그의 라틴문법 책은 라틴어 문법을 정리한 것에 더해 그의 해박한 고전어의 문헌학적 소양에서 나온 주옥같은 예문과 지문으로 더 유명하다. 이러한 Lily의 고전어 문법기술 전통은 그대로 살아있는 현대어에도 이어졌다. 그래서 언어교육에서는 교육해야 할 문법사항의 선정과 분류를 위해서 교수요목을 구성하는 것보다 해당 문법에 나타내는 좋은 글을 찾는 것이 문헌학자의 작업과 흡사하여 문헌학적 방법이라고 하였다. 좋은 예로 정통종합영어를 당시 마산고등학교 재직중이던 송석문선생이 집필할 때 문법 항목마다 좋은 예시문을 찾기 위해서 박근우선생께서 문학자료를 뒤지면서 좋은 글을 찾는데 많은 고생을 하셨고 자제분은 수시로 찾아서 수합한 예문을 송선생께 배달하셨다는 얘기를 들었다(박근우총장 개인면담).

언어학의 시작은 고전 라틴어, 그리스어, 산스크리트어를 중심으로 인구어의 조상어를 밝히려는 역사비교 언어학이고 근대적 공시언어학 연구이전에는 언어학은 문헌학과 연구분야가 많이 겹치는 역사언어학이었다. 그래서 언어학적 연구가 언어의 변화를 다루는 통시언어학에 머물러 있었는데 Saussure에 의해서 공시언어학의 중요성을 언어학계에 인식시켰고 그 이후에 공시언어학을 포함하는 언어학 연구의 지평이 확대되었다. Saussure와 더불어 공시언어학적 연구를 통해서 유럽의 구조주의가 탄생하였다(Bally & Sechehaye, 1916). 특히 Saussure의 langue와 parole로 대비되는 구조와 현상의 관계에 대한 통찰력은 프라하학파에 속하는 Trubetsky와 Jakobson으로 대표되는 음운학의 성립으로 발전되었다(Newmeyer, 2001; Novak, 1968). 경험주의 전통이 강한 영국의 음성학과 Martinet의 기능언어학과 같은 분야에서 많은 진전을 이루게 하는 이론적 틀이 되었다. 한편 미국의 인디언 연구를 위해서 사용된 discovery procedure와 같은 기술되지 않은 언어의 기술방법을 통해서 미국의 구조주의 언어학이 발전하였다. 기초어휘에 기반해서 해당하는 형태를 찾는다는 의미에서 미국의 구조주의는 유럽의 구조주의와 달리 태생적으로 분포주의적(distributionalism) 성격이 강한 구조주의였다(Bloomfield, 1933). 이러한 유럽과 미국의 구조주의 연구성과는 언어의 형태적 시스템에 대한 이해를 높였고 언어교육에 응용되었다. 이 방법을 언어과학적 이론에 근거한 언어교육 방법으로 분류할 수 있다.

언어학에서 Chomsky혁명 이후에 통사적 자율성(syntactic autonomy)이론과 더불어 문장하부구조(subsentential structure)의 규명을 통한 언어적 능력(linguistic competence)의 설명에 몰두하게 되었다. 그러나 외국어 교육에서는 언어적 능력뿐만 아니라 의사소통능력(communicative competence)의 습득이 중요하다는 주장과 함께 비교적 최근에 이르러서는 언어를 의사소통의 도구로 이해하고 의사소통에 필요한 언어표현과 해당 표현이 의사소통 상황에서 나타내는 의사소통기능을 중심으로 언어교육의 교수요목을 구성하여 가르치고자 하는 노력이 있는데 이를 상호작용적 방법이라고 하겠다. 본 연구에서는 이들 방법들을 하나씩 보면서 언어교육상의 특징과 내용학과 교육학의 상호작용을 살펴보고자 한다.

본론으로 들어가기 전에 언어학과 외국어교육에서 사용되는 “기능”이라는 용어를 정리할 필요가 있다. 우선 언어학에서 사용되는 문장하부구조 기능(function)은 주어, 목적어, 술어와 같은 문법적 기능, 행위자, 수혜자, 피행위자와 같은 의미적 기능, 주제와 진술(topic/comment)과 같은 화용론적 기능이 있다. 외국어 교육에서는 기능은 크게 듣기, 말

하기, 읽기와 쓰기 같은 언어기능이 있고 소개하기, 질문고 답하기 등과 같은 의사소통적 기능이 있다. 본고에서는 언어학적 기능을 통칭할 때는 언어내적 기능이라고 하고 외국어 교육에서 사용되는 언어기능은 그대로 사용하고 의사소통적 기능은 언어학과의 관계에서 언어외적 기능이라고 필요한 경우에 구분해서 지칭한다.

II. 내용학과 교육학의 언어교육 방법별 상호작용

1. 문헌학적 방법

Grammar School은 중세의 신과 성직자 그리고 왕과 영주를 중심으로 한 시대에 597년에 설립된 Canterbury Grammar School을 시작으로 과목과 내용은 변했지만 지금까지 세계 각지에 퍼져서 초중등교육기관이 되었다(Hadow, 1926). 이들 학교들이 시작될 때는 성당이나 수도원에서 장래의 성직자를 양성하기 위한 목적으로 당시의 성서 언어였던 라틴어교육이 중심이었다. 그러나 1503년 르네상스의 도래와 함께 Bridnorth Grammar School과 같은 시민사회와 직능단체들에 의한 Grammar School들이 생겨나면서 고전교육을 통한 인문주의와 르네상스 지식인들의 산실이 되었다. 영국에서는 Henry 8세에 의해서 수도원과 성당의 재산이 몰수되고 16세기 후반으로 들면서 급격하게 성당과 수도원과의 관계가 단절되고 Grammar School은 사회의 직능단체나 왕과 영주들에 의해서 교양있는 시민교육으로 발전되었다. 따라서 공교육으로서 Grammar School의 기능이 시작되면서 고전어 교육은 teaching foreign languages for no obvious reasons(tenor)의 시대가 시작되었다. 학교의 주인은 바뀌었지만 기계적 암기와 육체적 징벌로 특징 지워진 라틴어 교육방식은 그대로였다. 1학년때는 어휘와 품사를 배우고, 2학년때는 배운 어휘와 품사를 가지고 문장을 써보고 3학년때는 라틴어를 번역하는 교육과정으로 이루어져서 14세 졸업시에 교육과정의 목표는 유명한 라틴어 저작자들의 작품을 경험할 수 있도록 구성되었다(Hadow, 1926). 외국어의 형태중심 교육이 이루어지는데 전통적으로 이루어지던 라틴어와 고전 그리스어 교육은 이들이 모두 사어이기 때문에 발음을 익히는 것은 중요하지 않았고 쓰여진 고전문헌을 읽을 수 있도록 기술된 언어의 독해에 필요한 문법중심 교육이 이루어졌다. 동사의 변화나 문법성 시제와 상과 같은 각 문법사항을 익히면 관련된 표현이 나타나는 텍스트를 번역하면서 문법의 변화를 교육시키는 소위 말하는 문법번역(grammar and translation)방식이 주를 이루었다. 이 시대의 외국어교육은 고전어에 대한 문어교육으로 선형적 기술에 의해서 학교 문법을 외우고 학교문법의 교수요목에 따라서 교육이 이루어졌다.

2. 언어과학적 방법

문어중심의 외국어 교육에서 구어중심의 외국어 교육으로 외국어 교육의 초점이 이동되는 변화를 이루었다. 지금도 서구의 많은 대학들에는 Dept. of Modern X Language라는 학과들이 많이 있는데 이는 고전어 중심의 문어교육을 사어가 아니라 살아있는 주변의 외국어에 대한 교육이 필요하다는 인식을 담은 구어중심 외국어교육이라는 반동적 의미도 담겨있다고 볼 수 있다. 이러한 외국어 교육에서 관점의 전환은 언어학이 통시언어학으로부터 해방되어 공시언어학에 대한 관심과 연구를 시작하면서 공시언어학의 이론적 바탕위에 교수요목을 설정한 외국어교육 방법들이 나타나게 되었다. 이러한 외국어교육 방법을 언어과학적 방법이라고 용어를 정리하고 기술해보고자 한다.

Saussure 이후 공시언어학에 대한 연구가 불어나고 langue와 parole의 개념적 구분으로 언어공동체의 langue의 하부 시스템에 대한 개별 언어의 연구가 활발해졌다(Bally & Sechehaye, 1916). 당시 이미 과학적 방법으로 인정을 받던 음성학에 더불어서 Trubetskoy와 Jakobson등이 음성에서 추상적 개념 층위인 음소와 음소의 대립쌍 구조에 대한 연구를 활발히 하면서 Prague학파로 통칭되는 유럽대륙의 구조주의가 일어났다. 기호에서 기호의 형태와 의미 사이의 임의성을 기반으로 대립(contrast)과 분포(distribution)로 체계를 이루는 하나의 자율적 시스템으로서 언어학을 제창하였다. ‘별’이라는 의미와 *star*라는 형태의 관계는 임의적이고 *star*는 *stir*와 최소변별쌍(minimal pair)으로 음운론적 대립관계에 있고 /a/라는 음운은 다양한 상보적 분포(complimentary distribution)를 이루는 이음(allophones)의 음소로서 대표음이다. 대립과 분포는 각각 유럽과 미국 구조주의의 중심개념으로 발달하면서 전자는 기능중심적 언어교육 개념으로 후자는 형태중심적 언어교육 개념으로 발전하였다.

우선 영국에서는 런던대학 음성언어학과를 중심으로 한 음성학의 발달로 외국어발음을 제대로 할 수 있도록 훈련시키기 위해서 Henry Sweet, Daniel Jones와 같은 유럽의 외국어 교사와 교수들에 의해서 International Phonetic Alphabet이 만들어지고 발음 연구뿐만 아니라 외국어 교육에서 음성학적 방식이 도입되었다(IPA, 1999; Jones, 1909). 우선 유럽의 외국어 교재에 신출어휘에 대한 발음기호가 들어가고 이에 대한 정확한 발음훈련 강습회 등이 교사단체를 중심으로 열리고 교실에서 학생들에게 충실하게 발음교육을 하게 되면서 듣기 말하기와 같은 구어기능 교육에 관심을 가지게 되었다. 이와 더불어 음성 음운론적 개념에 근거한 근대적 의미의 외국어 교육에 대한 각성이 시작되었다.

음성 음운학과 더불어서 어휘에 대한 연구도 활발하게 이루어졌다. Michael West(1953)는 역사언어학에서 음운적 대응어 발견을 위해서 사용하던 기초어휘의 개념을 발전시켜서 외국어교육을 위한 기초어휘 1000선, 2000선과 같은 어휘목록을 제안하고 기초어휘를 중심으로 고전문학 작품의 다이제스트판 graded readers의 사용을 권장하였다. 그 이후에 Michael Lewis에 의해서 어휘접근법(Lexical Approach)은 구체적 어휘중심의 교수요목으로 발전하였다. 그리고 세계다독학회(International Extensive Reading Association)에서는 지금도 Graded Readers의 발간과 사용을 적극적으로 권장하고 매년 우수한 작품에 대해서 상을 주고 있다. 이에 따라 어휘론적 개념에 근거한 근대적 의미의 외국어 교육에 대한 각성이 시작되었다. 현재도 꾸준히 외국어 어휘교육에 필요한 기초어휘, 일반어휘와 일반과 전문 학술어휘, 직능어휘 등으로 나누어서 신출어휘의 효율적 교육을 통해서 외국어 교육의 효과를 도모하자는 외국어교육 운동은 지속되고 있다. 최근에는 코퍼스기반 언어학의 발달과 더불어서 대용량 코퍼스의 구축과 클라우드 컴퓨팅의 발전으로 자료기반 언어학습(Data-driven Language Learning)으로 발전하고 있다(O’Keefe, McCarthy & Cater, 2007).

구어 영어교육의 초기에는 직접식교수법(Direct Method)이라는 방법이 널리 쓰였다. 이 방법은 원어민이 언어과학 이론에 기반한 체계적인 교수요목없이 자신의 모국어를 가르치는 방법이다. 체계적인 언어연구에 기반하지 않았기 때문에 원어민이면 자신의 모국어를 잘 가르칠 수 있다는 민간의 일반적인 잘못된 믿음이 반영된 비과학적 교수법이라고 볼 수 있다.

직접식교수법에 대한 반작용으로 발음중심 교수법과 더불어 기술된 언어의 기초어휘와 구조를 연습시키는 Oral Method가 영국을 중심으로 발달하게 된다(Palmer, 1921). 이는 기존의 발음 교수요목에 더해 중요한 문장의 통사적 구문을 모아서 연습하고 필요한 어휘도 널리 쓰이는 기초어휘로 어휘제한을 두어서 해당 구문이 나타나는 대화문을 연습하는 형태로 외국어 학습활동을 구사하였다(Palmer & Hornby, 1937). 대표적인 인물이 Harold Palmer이고 그는 직접교수법과 Oral Method를 혼용해서 사용하는 Berlitz Language School에서도 일했다. 유명한 소설가 James Joyce도 젊은 시절 Trieste의 Berlitz에서 일한 것으로 알려져 있다. Palmer는 Hornby와 더불어서 일제가 한반도를 통치하던 시절에 일본의 문부성에 문정관으로 임용되어서 14년간 일본에 머물렀다. 당시 일제치하에서 발행된 조선의 고등보통학교 영어교과서 구성체제에도 많은 영향을 미친 인물이다. 따라서 당시 식민지 조선에 보급되었던 고등보통학교와 고등여자보통학교의 영어책에 사용된 교수요목도 일부 Oral Method의 영향을 받았다고 할 것이다. Oral Method에서 좀 더 발전해서 그림으로 대화문이나 읽기 지문의 상황을 그림이나 실물과 함께 제시하는 방법으로 외국어를 가르치는 것이 Situational Method이다. 처음에는 고전어를 공부하는 Grammar School보다는 일반 사설학원이나 앞으로 해당 외국어로 상호작용을 해야하는 직업을 갖게 될 평민들을 대상으로 하는 교재에 많이 적용되었다.

한국의 영어 교육에 지배적인 영향을 미친 사람들은 Charles Fries와 Robert Lado로 이들 미국의 언어학자들은 외국어 교육에서 Aural-Oral Method (일명 Oral Approach 또는 속칭 Audio-lingual Method)의 중요성을 주장하였다(Ney, 1968; Brooks, 1975). Fries의 구조주의적 방법론은 유용하게 쓸 수 있는 기초어휘와 더불어서 미국의 구조주의 언어학에서 나온 discovery procedure의 직접적 영향으로 인디언 언어의 기술을 통해서 해당 언어의 음운시스템을 발견해내고 문장의 kernel sentence pattern을 발견해 내는 과정을 통해서 해당 언어의 문법을 기술하였다. 이러한 미국 구조주의 전통의 영국의 Oral Method와 함께 영어의 kernel sentence pattern을 중심으로 영어를 배운다는 의미의 Michigan Action English나 English 900와 같은 책들이 출간되었다. 900개의 문장패턴과 이들이 나타나는 대화문을 원어민 녹음으로 구성한 테이프들이 우리나라의 영어 열풍과 더불어서 날개 돋힌 듯 팔려나갔다. 이를 미국에서는 언어학의 통사적(syntagmatic) 구조에 기반한 Audio-lingual Method라고 발음, 어휘제한과 중심문장을 기반으로 한 교수요목에서 출발한 Oral Method와 차별화해서 불렀다.

Audio-Lingual Method의 특징은 교재의 선택, 그 배열 및 지도기술의 구체화에 있다. 어휘 중심이 아닌 구조 중심의 교재의 배열, 음성이나 구조의 분포적 요소를 기본으로 하는 지도법, 음성 및 구조의 완전학습, 언어운용을 목표로 하는 습관형성, 최소한의 시간에 최대한의 연습량의 확보, 문형연습 및 모방 암기법에 의한 효과적인 연습 등의 기본적인 원칙을 그 내용으로 하고 있다. 그는 또한 교재 작성의 지침으로 목표어와 모국어의 대조분석에 의한 교재의 작성, 구조중심적 언어 재료의 단계적 도입, 제공되는 언어 재료와 관련된 장면의 설정을 강조하였다.

정리해 보면 전통적인 문법과 번역 중심의 전통적 외국어교육 방법은 살아있는 언어는 사어와 달리 구어이며 외국어 교육의 본질은 구어의 습득이라는 점을 무시하고 형태적인 문어의

문법과 문법을 나타내는 문어의 구문정독과 해석에 몰입된 경향이 있다. 그러나 오랫동안 문법편중주의를 통하여 전통문법은 체계화되고 거의 완벽에 가까운 정도로 형식화되고 정비되어서 수동적인 학습에 편의를 제공하고 교양 및 이성적 판단력의 양성을 위하여 많은 효과가 있었다. 이들 방법은 모두 20세기 초에 이르러 전통적인 문법번역식 교수법에 대한 비판과 더불어 개혁적 이론이 대두되고 그것은 주로 음성과 형태중심의 구두교수를 주창하는 내용들이었다.

언어학이 제공한 이론적 기반위에서 이루어진 형태중심의 외국어교육은 일부 집중식 외국어 교육기관에서 성공을 거두었다. 미국의 Defense Language School이나 미국 물문교 선교사 양성기관이 BYU의 Mission Training Center같이 학습자들을 모아놓고 합숙을 시키면서 일정기간 집중훈련을 하는 어학원들은 괄목할 만한 성공을 보였다. 현재도 사실 어학원 가운데서 Berlitz와 SDA어학원과 같이 유명한 학원들에서 여전히 이 방법을 채택하여 부분적으로 활용하거나 전체적으로 교수요목을 구성하여 쓰고 있다.

이러한 성공에도 불구하고 공교육 기관에서 이루어지는 산발적 외국어교육에서는 이렇다할 성과를 내지 못하였다. 일주일에 2-3시간 이루어지는 외국어 교육에 더해 매학기마다 끝나는 대로 맞게 되는 방학이라는 긴 외국어교육 휴지기간으로 이루어진 공교육에서 형태중심 외국어교육은 배우는 학습자들에게 증발성이 높은 것으로 판명되면서 문장의 패턴이 안착하지 못하고 배우고 잊어버리는 효과성이 낮은 교육으로 많은 비판을 받게 되었다. 사실 이와같은 비효율성은 교수방법의 본질적 문제라고 보기보다 관점에 따라서는 외국어교육 시간운영이라는 운영의 문제로 볼 수도 있다. 그러나 주어진 여건에서 좀 더 효과적인 방법을 찾아가는 것이 교육하는 사람들의 조건행동이라고 할 때에 형태중심 외국어교육에 대한 비판으로 의미중심의 외국어 교육을 찾게되는 것은 어쩌면 당연한 귀결이라고 할 수 있을 것이다.

언어학적 교수법으로 통칭할 수 있는 발음교수법, Oral Method와 Audio-lingual Method의 대안을 의미중심 외국어교육 방법을 찾아가던 중에 위의 언어학적 교수법을 정의적으로 좀 더 친근한 방법으로 학습자 중심 외국어 교육방법을 만들고자 하는 노력들이 있었다. Suggestopedia는 kernel sentence pattern을 바로크풍의 고전음악과 함께 들려주고 합창독으로 연습하게 하는 방법을 제안하였고 Silent Way는 교사의 제스처나 그림과 도구를 활용해서 학생들이 kernel sentence pattern을 발화할 수 있도록 하였다. 그 외에도 Community Language Learning에서는 학생들이 모둠으로 자신들이 발화하고 싶은 목표어를 kernel sentence pattern에서 찾거나 교사의 도움을 통해서 문장을 완성해서 함께 합창독을 할 수 있도록 하였다.

3. 상호작용적 방법

상호작용적 방법은 다른 말로 언어과학적 발견에 따른 언어의 형태를 습득의 중심에 두는 교수법과 대조적으로 의미중심적이고 언어 사용의 목적인 의사소통 기능을 중심에 두는 외국어 교육 방법이라고 할 수 있다. 의미중심의 의사소통 기능을 익힐 수 있는 상호작용 방법을 찾게 된 사람들은 영국을 중심으로 Notional-functional 교수요목을 주장하였다. 개념

(notion, 개인적으로는 ‘관념’이라는 표현이 더 나오나 관례를 따라서 ‘개념’을 사용)은 시간, 공간, 관계, 양태(modality), 양(quantity), 질(quality), 지칭과 같은 언어 사용자들의 세계관이 언어의 시스템에 장착된 거시적 개념과 수학, 언어, 과학 등과 같은 내용교과에 대한 어휘와 중심표현으로 구성된 미시적 개념으로 나누어 볼 수 있다.

거시적 개념을 중심으로 언어를 기술하고 언어유형을 나누는 방법으로 준거로 사용한 언어과학적 학파로는 이미 유럽의 기능주의 언어학파가 있다. Hjelmslev, Malinowski 그리고 Firth의 기능주의 언어학 전통을 이어받은 Halliday의 Systemic-Functional Grammar에 기능주의 언어학 이론적인 바탕이 비교적 잘 드러나 있다(Langendoen, 1968). Firth(1957)의 “You shall know a word by the company it keeps.” (p. 11) 말이 기능주의 언어학을 요약해주고 있다. Halliday는 그의 이론에서 기본적으로 언어는 사회적 의사소통의 도구이고 이는 크게 경험을 공유하고자 하는 ideational function, 사회적 관계를 유지하고자 하는 interpersonal (social) function, 그리고 일관성있고 응집력 있는 담화를 구축하고자 하는 textual function으로 의사소통 기능을 나누었다(Halliday, 1994).

거시적 개념을 중심으로 의미와 언어외적 기능 중심의 상호작용적 외국어 교육은 위의 Notional-functional syllabus를 core외국어 교과로 다루는 학교와 미시적 개념을 중심으로 일부 캐나다와 같은 곳에서는 타교과내용을 통제로 목표 외국어로 가르치는 Immersion 교육으로 구분해서 볼 수 있다.

1) 교실에서 Core외국어 교과에서 이루어지는 상호작용

우선 Notional-functional교수요목을 살펴보면 교수요목을 실제로 외국어를 배워서 쓸만한 의사소통기능으로 분류한다. 현재 우리나라에서 1992년 이후 영어교과에서 채택하고 있는 것이 Notional-functional교수요목인데 교과서의 단원구성을 살펴보면 자신을 소개하기, 장소 묻고 답하기, 가격묻고 답하기, 안부묻고 답하기 등과 같은 일상적인 생활에서 필요한 대화에서 사용할 수 있는 의사소통기능을 교수요목에 담아서 단원을 구성하고 있다. 중학교 고등학교 가면서 주제중심의 읽기 자료들이 등장하지만 큰 틀에서는 이들 읽기 자료들도 언어의 형식을 익히는 정독자료라기보다는 의미를 빨리 익힐 수 있도록 하는 다독자료이다. 물론 현장에서 교사들이 지문에 나오는 문장의 형식이나 구문의 특징을 중심으로 정독을 시킬 수 있지만 수능능력시험에서 측정하고자 하는 읽기능력을 보면 읽기에서 정독보다는 다독을 중요시한다는 것을 알 수 있다. 다시말해서 수능의 읽기를 주어진 시간에 다 풀고자 하면 분당 120단어 이상을 읽고 의미를 파악할 수 있는 능력을 요구하고 있는데 이는 고등학교 때까지 영어교과서를 정독중심으로 읽고 형태를 익히는 방식으로는 다다를 수 없는 읽기의 유창성이고 읽기 속도이다.

Notional-functional 교수요목에서는 외국어 학습의 목표에 맞추어서 의미적 교수요목을 먼저 구성한다. 이 때에 의미라는 것은 언어학적 의미라기 보다는 의사소통에서 특정 의미를 전달하려고 하자면 이라는 의사소통의 기능적 의미이다. 따라서 언어학에서 어떤 표현의 텍스트적 의미라기보다는 표현이 의사소통 상황에서 일어나는 기능적 의미로서 담화적 의미라고 볼 수 있다. 의사소통기능은 소통이 일어나는 시간, 장소, 화자와 청자 그리고 이들 사

이의 사회적 관계와 심리적 관계라는 복잡한 상황 속에서 어떤 표현이 담아내는 의미이다. 이는 교과서에 나오는 표현에 언어의 향기를 불어넣어서 실제로 사용되는 담화를 배우고 익히는 교수요목이다. 따라서 이를 제대로 수행하기 위해서는 의사소통기능을 제대로 익힐 수 있는 교수방법이 필요한데 언어적 상호작용을 중요시하는 의사소통중심 교수법과 의미적 상호작용을 중요시하는 과제중심 교수법이 의사소통기능을 가르치고 배우는데 최적화된 교수법이라고 볼 수 있다. 이들 교수법은 외국어 교육에 종사한 교사들이나 외국어 교육의 전문가들이라고 할 수 있는 학자들이 공동으로 개발한 교수법이다. 이런 의미에서 언어교육사에서는 처음 있는 일이라고 할 수 있다. 이전에는 외국어 교육이라는 영역이 독자적인 학문적 영역으로서 전문가를 양성하고 외국어 전문 교사를 양성하는 연계성을 갖지 못했다. 외국어를 잘하는 사람들이 외국어를 잘 가르칠 수 있다고 사람들은 전통적으로 생각하거나 외국어에 대해서 잘 아는 언어학자들이 내놓은 방법이 최적화된 외국어 교육방법이라고 생각하고 외국어 교육을 전문적인 영역이라고 보지 않았던 것이다.

Notional-functional 교수요목은 기존의 문법적 표현인 언어의 notion을 의사소통상황 속에서 요구하는 표현의 기능적 의미와 연결한 것으로 이를 뼈대로 해서 교수요목을 만든 것이다. 따라서 본질적으로 담화에서 일어나는 의사소통기능을 중심으로 내용을 선정하고 유용한 표현에서 덜 유용한 표현 또는 널리 쓰이는 표현에서 다소 국부적인 표현의 순서로 배열하는 교수요목이다. 이러한 내용을 뒷받침하는 이론적 틀은 Andre Martinet의 역사언어학에서 push-chain과 pull-chain등의 개념적 기능주의 언어학에서 그 기원을 찾을 수 있으며 언어학의 언어현상에 대한 기능적 기술은 근대에 Mark Halliday의 Functional Grammar에서 그 맥을 잇고 있다. 이는 그동안 sub-sentential grammar에서 supra-sentential grammar로의 관점이동이며 각 표현의 의사소통상황에서 기능적 의미를 중심으로 교수요목을 구성하고자 하는 사람들에게 많은 영향을 미쳤다. 크게 의사소통기능은 잡담나누기를 중심으로 하는 social function과 중요한 정보의 교환을 중심으로 하는 transactional function으로 개념적으로 나눌 수 있다. 그리고 각 개념적 의사소통 기능 범주에 수없이 많은 하위 의사소통기능과 그 대표적 언어표현을 나누어서 해당 언어표현이 의사소통 목적에 맞게 나타나는 실제적 대화를 배우고 활동, 게임, 역할극을 통해서 해당 의사소통 표현을 의사소통목적에 맞게 유창하게 쓸 수 있도록 교육하고자 하는 것이다. 이를 위해서 대표적으로 사용되는 교수법이 바로 의사소통중심 교수법과 과제기반 교수법이다.

2) 몰입적 상호작용

몰입적 상호작용은 거시적 환경이 목표어가 사용되는 L2환경이거나 학교의 방침에 따라서 학생들이 몰입식으로 해당 외국어를 사용하는 환경에서 교육을 받고 기숙사 같은 곳에서 머물면서 목표어만을 사용하는 것으로 생활규칙으로 삼아서 이루어지는 언어교육 환경을 말한다.

언어교육사에서 모국어 교육은 문자교육, 발음교육, 고전어 교육으로 나누어서 교수요목이 개발되고 근대적 공교육이 이루어지기 전에는 주로 계급에 따라서 이루어졌다(권오량과 김정렬, 2010; 김정렬, 2017). 외국어 교육의 경우에는 기능적인 교육으로 일반인들의 필요에 의해서 이루어지거나 역관들의 교육을 위해서 전문적인 분야의 의사소통을 목적으로 전문통역사나 번역사를 양성하기 위한 목적으로 시작되었다. 1572년 8월 24일에 있었던 프

랑스 파리에서 시작되어서 루앙, 리옹, 오를레앙의 프랑스 각지로 번져나간 성바르톨로메오 축일의 대학살로 구교도들의 당시 위그노라 불리던 신교신자들에 대한 박해가 극에 이르자 많은 신교도들이 신변의 위협과 종교적 박해를 피해서 비교적 종교적 자유가 보장되었던 영국으로 갔다. 위그노들 중에 일부는 노르만 정복이후에 영국사회에 만연했던 불어 사교육에 종사했고 위그노들 역시 이국땅에서의 삶을 영위하기 위해서 영어교육이 필요했다(Howatt & Widdowson, 2004). 영국에서는 이 때부터 위그노들의 불어교육과 영국인들이 위그노들 에게 실시한 영어교육이 실용적 목적을 가지고 상호간에 이루어졌다.

한편 한반도에서는 고려시대의 통문관에서 이루어지던 역관양성 업무를 조선에 와서는 사역원이라는 기관을 두어서 한어, 몽어, 여진어, 왜어 역관을 양성하였다. 이들 중 일부는 어렸을 때부터 역관집안에서 자라거나 변방지역에서 해당 언어를 익힐 수 있는 기회가 있거나 일부는 인질로 해당언어를 사용하는 나라에 잡혔다가 탈출한 경우가 있었지만 대부분 잡과에 해당하는 과거시험인 역관선발 시험을 통해서 선발되어 사역원에 입소하고 2-3년간 해당외국어만 사용하면서 기숙생활을 하는 고강도 훈련을 통해서 엄격한 외국어능력의 검증후에 역관으로 임무를 부여받았다. 이 때에 교본으로 사용된 책이 노걸대인데 이 책은 한어 역관을 훈련시키는 중심교본으로 주로 조선의 사신을 명칭의 수도에 수행하고 갔다가 돌아오는데 필요한 한어표현을 익히도록 구성되어 있다. 예를 들어서 여관을 찾아서 방이 있는지, 말을 먹일 여물은 제공이 되는지 가격은 얼마이고 필요한 물건들을 살 수 있는 시장은 어디인지 등을 묻고 답하는 것을 알아들을 수 있도록 훈련하였다(김정렬, 2017). 더불어서 역관들에게는 인삼을 일정분량 무역하여 출장비로 충당하도록 하였으므로 인삼을 살 곳이 있는지 묻고 가격을 협상하는 과정에 필요한 표현을 익힐 수 있도록 하고 있다(이상각, 2011). 따라서 한반도의 외국어 교육의 시작도 일부 외국어가 필요한 사람들에게 필요한 의사소통기능을 담은 외국어 표현을 익히도록 하는 실용적 외국어 교육으로부터 시작되었음을 알 수 있다.

다른 한편의 의미중심 외국어교육의 방향은 외국어를 교실언어로 사용해서 교과교육을 해당 외국어로 실시하는 몰입교육이다. 몰입교육에서는 외국어 학습이 하나의 도구교과의 역할을 하고 학습내용인 해당교과 내용이 도구인 외국어를 통해서 상호교환되는 의미의 역할을 한다. 이를 시작한 나라는 캐나다인데 Immersion이라고 하고 캐나다 온타리오주의 경우는 외국어를 core curriculum으로 가르칠 수도 있고 immersion으로 가르칠 수도 있는 선택을 학교 담당 교사, 관리자, 학부모가 할 수 있도록 하고 있다. core curriculum으로 가르친다면 우리와 같이 외국어 교과를 따로 설정해서 가르치는 것이고 아니면 immersion으로 부분적으로 또는 전면적으로 내용교과를 목표 외국어로 가르치는 것이다. 실제로 온타리오주에서는 불어를 퀘벡주에서는 영어를 immersoin으로 가르치는 경우가 많으며 immersion학교가 core curriculum학교에 비해서 외국어 능숙도가 뛰어난 것으로 보고하고 있다. 이에 따라서 미국과 호주 일본 한국 등에서도 몰입학교들이 생겨났으나 한국에서는 초등사립학교를 중심으로 몰입학교들이 있었으나 교육부에서 못하도록 교육과정 운영에 간섭하고 중학교 진학시에 연계된 몰입학교들이 주변에 없어서 그 효과를 제대로 보지 못하고 있는 실정이다.

유럽에서는 몰입과는 달리 CLIL (content and language integrated learning)이라는 방법으로 일부 교과를 목표 외국어를 가르치는 경우가 핀란드, 스웨덴과 같은 북유럽에서 시작

해서 EU산하에 중점 연구기관을 두고 EU회원국들이 외국어 습득을 위해서 학교 단위로 많이 채택하도록 관련 자료들과 교수요목을 제공하고 있다. 북아메리카와 유럽의 몰입교육과 CLIL의 차이는 목표 외국어를 쉽게 접할 수 있는 제2언어 환경의 여부와 외국어 교육 방식에서 sheltered class를 통해서 교수언어를 따라가지 못하는 학생들을 격리시켜서 쉬운 목표어로 내용을 소개하는 방식을 채택하느냐 여부에 따라서 달라진다. 몰입교육에서는 목표외국어로 강의가 이루어지고 따라가지 못하는 학생들은 따로 sheltered class로 뽑아내서(pull-out) 쉬운 목표 외국어로 내용을 가르치는 것이다. 그리고 학생들의 거시적 언어환경이 쉽게 목표외국어를 소통할 수 있는 제2언어환경에 노출되어 있다. 그러나 유럽의 CLIL은 거시적 언어환경은 모국어 환경이고 외국어로서 해당 목표어를 배우고 해당 목표어로 구성된 교수요목을 학교에서 배운다는 점에서 북아메리카의 몰입교육과는 다르다. 비교적 모국어 환경에서 외국어를 배우는 언어적 경험이 유사한 학생들이 모여서 교실을 구성한다는 점에서도 차이가 있다. 북아메리카의 경우에는 학생들의 목표 외국어에 대한 언어적 경험이 다양하고 구어와 문어에서 수준차를 다양하게 보이고 있어서 sheltered class가 필요하지만 유럽의 CLIL 환경에서는 이와같은 sheltered class의 편성이 필요없다는 것이다 (김정렬, 2011).

우리나라의 경우도 사립초등학교와 중학교를 중심으로 번져가던 영어몰입교육이 교육부의 제재로 막히면서 현재는 교육부 인가 제주와 송도의 국제학교, 외국인 학교, 미인가 대안학교를 중심으로 영어몰입교육이 실시되고 있다. 또한 대구광역시 교육청과 제주특별자치도 교육청 산하에서 실시되던 IB (International Bacalorea) 교육과정을 경기도를 비롯한 광역 교육청 단위에서 자율적으로 국가교육과정 대신에 채택하는 길이 열리게 됨으로서 일부 교과를 중심으로 부분몰입을 실시할 수 있는 여지를 남겨두고 있다(임태희, 개인소통).

III. 대학 입학고사의 변천과 사범대학 영어교육과의 변화

우리나라 대학입시에서 영어시험은 일제 식민지 시대 경성제대에서 시행된 1924년 3월에 시행된 입학시험이 그 시작이다. 그리고 해방후에는 대학입시제도의 변천과 더불어서 약 10회 정도의 영어시험에 대한 변화가 있었다(권오량과 김정렬, 2010). 그 구체적인 변천과정을 보면 다음과 같다.

- (1) 대학별 단독 시험제 (1945-1953)
- (2) 대학입학 국가연합고사 & 본고사 병행제 (1954)
- (3) 대학별 단독 시험제 (1955-1961)
- (4) 대학입학자격 국가고사제 (1962-1963)
- (5) 대학별 단독시험제 (1964-1968)
- (6) 대학입학 예비고사, 본고사 병행제 (1969-1980)
- (7) 대학입학 예비고사, 내신 병행제 (1981)
- (8) 대학입학 학력고사, 내신 병행제 (1982-1985)
- (9) 대학입학 학력고사, 내신, 논술고사 병행제 (1986-1992)
- (10) 대학수학능력시험, 본고사, 내신 병행제

권오량과 김정렬(2010, pp. 508-9)

대체로 (1)-(6)시기는 문법과 번역, 해석문제가 주를 이루는 문헌학적 시기로 볼 수 있고, (6)-(9)시기는 발음, 어휘, 문법의 형태적 특징을 주로 묻는 언어과학적 시기로 볼 수 있고 (10)시기는 개념-기능교수요목에 근거한 언어의 의사소통적 기능에 중점을 둔 개념과 기능 중심의 언어기능(듣기, 간접말하기, 읽기, 간접쓰기)의 문항으로 이루어진 시기로 볼 수 있다.

이러한 구분은 우리나라 사범대학 영어교육과를 구분하는 기준으로 볼 수 있다. 우리나라 사범대학 영어교육과는 크게 3개의 시기로 나누어 설립되었다. 해방직후인 1946년부터 서울대학교 사범대학과 경북대학교 사범대학에 영어교육과가 설립되었고 이어서 1960년에 충청 전라 지역에 영어교사 수급을 담당하기 위해서 공주대학교에 영어교육과가 설립되었다. 1960년대 후반부터 1970년대까지 베이비붐 세대가 중학교에 입학하면서 영어교사의 수급난을 해소하기 위해서 사립대학에도 영어교육과 설립을 허용하고 각 시도의 국립대학에 의 무발령을 전제로 사범대학 영어교육과를 추가로 설립하였다.

우선 가장 오랜 전통을 가진 서울대학교 영어교육과와 경북대학교 영어교육과를 살펴보면 각 대학의 외국어교육과의 교원전공현황의 변화에서도 문학, 어학, 영어교육이 차지한 교사 양성 교수요목의 비중을 짐작해 볼 수 있다. 서울대학교와 경북대학교에 교수들의 전공을 보면 설립 당시에는 압도적으로 문학 전공자의 비율이 높았고 60년대에서 70년대로 오면서는 어학전공 교수의 비율이 높고 80년대와 90년대에 들어서야 영어교육 전공자들을 찾아볼 수 있다. 각 대학마다 최근에 은퇴한 비교적 젊은 최근 5년내 은퇴교수들 가운데 소수의 영어교육 전공 박사들을 볼 수 있다는 것은 앞에서 언급한 외국어 교육 분야의 학문적 성립시기가 짧다는 것을 방증한다고 볼 수 있다.

이는 기본적으로 외국어 교육이 외국어를 잘 하는 문학전공자들이 중심이 되어 문헌학적 전통의 외국어교육 중심이었음을 방증하는 것이고 이어서 형태중심 외국어 교육을 이끈 어학전공 교수들이 언어과학적 탐구를 중심으로 교육과정을 주도하였다. 최근의 상호작용적 방법론과 더불어서 영어교육으로 박사를 받은 학자들이 영어과를 중심으로 외국어 교육이라는 독립된 학문분야를 만들어 가는 과정에 있으나 아직은 내용학 전공자들과 협업을 통해서 상호의존적 협력관계를 설정하고 있는 단계에 머물러 있다.

사범대학 영어교육과의 학과 명칭도 대학의 공식적인 명칭과 달리 학교 구성원들 간에 통칭되는 학과명은 따로 있다. 처음 한동안 문헌학적 방법이 주를 이루던 문학전공 학자들이 대부분이던 시절에는 사범대 영문과로 불렸다. 그리고 언어과학적 시기에 와서는 어학전공 학자들이 영어교육과의 주요 구성원을 이루면서 사범대 영어과로 불렸다. 최근에 들어서 영어교육 전공자들이 1/3이상 구성하고 있는 학과들이 많은데 지금도 사범대 영어과는 영어교육과와 함께 통칭되고 있는 이름이다.

이는 학과에 관계된 사람들의 관점을 보여주는 단편적인 정보의 하나라고 볼 수 있다. 영어를 가르치는 사람들을 길러내는 곳은 영문학 작품을 잘 읽어내는 독해력과 독해력의 기본이 되는 문법을 잘 아는 사람이라는 문헌학적 시기가 있었고 문장의 구조와 소리를 잘 알아는 사람을 길러내야 한다는 언어학적 소양을 중시한 언어과학적 시기가 있었고 현재는 영어

교과 교육학적인 입장에서 영어의 교수방법과 교수요목 등에 밝은 사람을 길러내야 한다는 상호작용적 관점과 언어과학적 관점이 조화를 이루고 있는 시기라고 볼 수 있다.

IV. 결론

외국어 교육은 흔히 응용언어학의 한 분야로 간주된다. 응용언어학이란 언어과학의 탐구 결과를 응용하여 외국어 교수에 활용하는 학문이라는 말이다. 그런 의미에서 보면 문어중심의 외국어 교육이 주를 이루던 문헌학적 시기를 제외하면 대체로 외국어 교육 방법은 구조주의 언어학과 기능주의 언어학의 언어과학적 탐구 결과를 수용하여 그 내용을 발전시켰다고 볼 수 있다.

그런데 우리나라에서 최근의 상호작용적 외국어 교육방법이 언어과학과 결별한 것처럼 보이는 것은 그동안 우리나라에 Systemic-functional Grammar와 같은 기능적 전통의 언어과학적 탐구 방법이 제대로 소개되지 못하고 언어과학 연구의 주류를 이루는 문장구조의 문법성을 따지는 문법적 능력의 규명에 대부분의 연구역량이 집중되었기 때문으로 보인다. 물론 중요한 연구 아제다임에 틀림없으나 외국어 교육적 입장에서는 의미를 구성하는 표현의 발화와 이해를 교육시키는 의사소통능력의 함양이 중요하다. 따라서, 기능주의 언어과학 탐구에서 밝혀낸 개념과 의사소통기능을 중심으로 하는 개념-기능적 교수요목이 1992년 제6차 영어과 교육과정에 도입되었다. 개념-기능적 교수요목으로 가르쳐야 할 어휘, 의사소통기능의 선정과 의사소통 표현을 자연스럽게 생성해 낼 수 있는 상황의 설정이 중요했다.

이런 맥락에서 보면 우리나라의 외국어 교육과 언어과학 간에 현재에 나타난 현상적 분리상태는 기능적 언어과학 탐구방법이 주류 언어학에서 밀려나서 제대로 우리나라에 자리잡지 못한 데서 일어난 일시적 현상이라고 볼 수 있다. 좀 더 다양한 언어과학의 연구방법이 소개되고 그러한 연구방법이 적용될 필요가 있다. 최근에 언어과학과 외국어 교육 간에 많은 연구결과를 공유하고 발전하고 있는 분야 중의 하나가 코퍼스 언어학과 코퍼스 기반 외국어 교육이다. 결국 언어과학의 다양한 탐구결과가 외국어 교육의 교수요목과 교수방법 구성에 길을 안내하는 내용학으로서 언어과학과 그 결과를 외국어 교육에 담아내는 응용언어학으로서 외국어 교육은 다시 상호의존적 탐구과정과 결과의 긍정적 순환과정을 함께 가져갈 동반자 관계이다.

참고문헌

- 권오량, 김정렬. (2010). *한국영어교육사*. 서울: 한국문화사.
- 김정렬. (2011). 영어 몰입교육의 이론과 현황. 서울: 경문사.
- 김정렬. (2017). 노걸대 분석을 통해서 바라본 우리반도의 외국어 교육. *한국콘텐츠학회 논문지*, 17(6), 408-414.
- 이상각. (2011). *조선역관 열전*. 서울: 서해문집.
- Bally, C., & Sechehaye, A. (eds.) (1916) *Cours de linguistique générale*. Paris: Payot.
- Bloomfield, L. (1933). *Language*. New York: Holt, Rinehart & Winston.
- Brooks, N. (1975). The meaning of audiolingual. *The Modern Language Journal*, 59(5-6), 234-240.
- Bullokar, W. (1977). *Booke at large for the amendment of English orthographie (1580) with a bref grammar for English (1586)*. Ann Harbor, MI: University of Michigan Scholars' Facsimiles & Reprints. ISBN 978-0-8201-1287-9.
- Firth, J. (1957). *Studies in linguistic analysis*. London: Wiley-Blackwell.
- Hadow, W. (ed.) (1926). *The Education of the Adolescent*. London: HM Stationery Office.
- Halliday, M. (1994). *Introduction to Functional Grammar* (2nd ed.). London: Edward Arnold.
- Howatt, A., & Widdowson, H. (2004). *A history of English language teaching*. London: Oxford University Press.
- IPA. (1999). *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press.
- Jones, D. (1909). *The Pronunciation of English*. Cambridge: Cambridge University Press.
- Langendoen, T. (1968). *The London School of Linguistics*. Cambridge: Cambridge University Press.
- Lewis, M. (1997). *Implementing the Lexical Approach*. Hove, England: Language Teaching Publications.
- Newmeyer, F. (2001). The Prague School and North American functionalist approaches to syntax. *Journal of Linguistics*, 37, 101 - 126
- Ney, J. (1968). The Oral Approach: A re-appraisal. *Language Learning*, 18(1-2): 3-13.
- Novak, P., & Sgall, P. (1968). *On the Prague functional approach*. Tuscaloosa, AL: Univ. Alabama Press.
- O' Keefe, A., McCarthy, M., & Carter, R. (2007). *From corpus into classroom: Language use and language teaching*. London: Cambridge University Press.
- Palmer, H. (1921). *The Oral Method of teaching languages: A monograph on conversational methods together with a full description and abundant examples of fifty appropriate forms of work*. Cambridge, England: Heffer and Sons.
- Palmer, H., & A. Hornby. (1937). *Thousand-word English: What it is and what can be done with it*. London: Harrap.
- Ward, A. and Waller, A. (1908.) *The Cambridge history of English and American literature*. Volume III: Renaissance and reformation. Cambridge: Cambridge University Press.
- West, M. (1953). *A general service list of English words*. London: Longman, Green and Co.

형태/통사/의미

Form Sequence & Coordinate Structure Constraint

Kwon Kiyang

(Youngsan University)

1. Introduction

- ◆ **Coordinate Structure Constraint (CSC, Ross 1967):** In a coordinate structure, no conjunct may be moved, nor may any element contained in a conjunct be moved out of that conjunct

- ◆ **CSC–regular coordinate structure**

- (1) a. *John_i seems to be [t_i and Mary] in the room.
 b. *[All the heaviness]_i had [t_i and the height] gone.
 c. *Who_i did John kiss [t_i and a girl]?
 d. *What table_i will he put the chair between [t_i and some sofa]?
 e. *The speaker who_i I watched [t_i and Bill] was vain.
 f. *[Go to the club]_i John might [t_i and [have some fun]].
 g. *Can_i you [t_i and will] stay at home? (Zhang 2007:2)

- (2) a. *Mary_i seems to be [John and t_i] in the room.
 b. *[The height]_i had [[all the heaviness] and t_i] gone.
 c. *Who_i did John kiss [a girl and t_i]?
 d. *What sofa_i will he put the chair between [some table and t_i]?
 e. *The speaker who_i I watched [Bill and t_i] was vain.
 f. *[Have some fun]_i John might [[go to the club] and t_i].
 g. *Will_i you [can and t_i] stay at home? (Zhang 2007:2)

- (3) *Which surgeon_i did Kim date [[friends of t_i] and [a lawyer]]? (Ross 1967: 89)

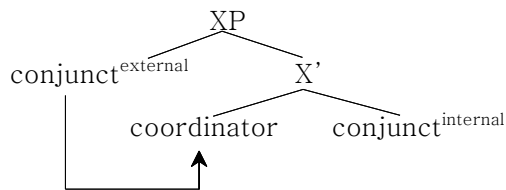
- ◆ **CSC–exceptional coordinate structure**

- (4) a. What_i did Harry [[go to the store] and [buy t_i]]? (*temporal coordination*, Lakoff 1986:152–153)
 b. That' s the news_i that the child [[heard t_i] and [broke down in tears]]. (*causal coordination*, Goldsmith 1985:135)

- c. How much_i can you [[drink t_i] and [still stay sober]]? (adversative coordination, Lakoff 1986:152)

2. Previous Analysis of CSC

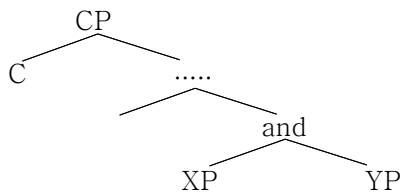
- (5) Immobility of Conjuncts (Munn 1987, Zoerner 1995, Johannessen 1998, Zhang 2007)



[C-feature]-inheritance

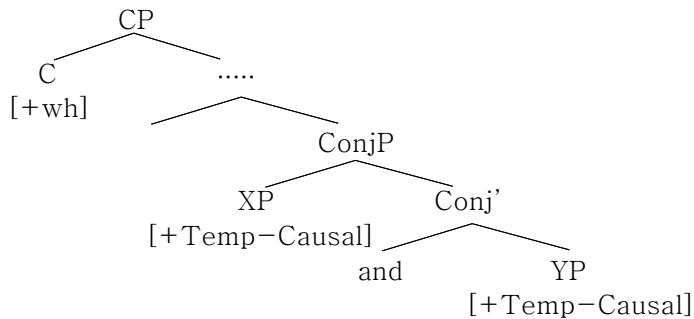
- (6) a. You can depend on [DP [DP my assistance] and [CP that he will be on time]].
 b. *You can depend on [CP that he will be on time].
 c. *You can depend on [CP [CP that he will be on time] and [DP my assistance]].
 → It is the external conjunct that satisfies the c-selection of the preposition.
- (7) *Which surgeon_i did Kim date [DP [DP t_i and a lawyer]]?
 → External conjunct cannot be extracted since it lacks a categorial feature.
- (8) *Who_i did John kiss [a girl and t_i]?
 → Internal conjunct cannot be extracted because of Relativized Minimality.

- (9) CSC-regular coordinator structure (Sato & Chua 2019)



→ XP and YP do not precede or asymmetrically c-command each other but rather are floating in separate dimensions equidistant to the interrogative C head. This means that syntactic operations must treat both XP and YP equally since they are hierarchically equal. Consequently, extraction operations cannot take place on one conjunct without necessarily affecting the other.

(10) CSC–exceptional coordinator structure (Sato & Chua 2019)



→ XP and YP share a matching feature, [+temporal-causal], which denotes this relationship and triggers event restructuring, thus collapsing two events into one macro-event. The two conjuncts XP and YP are merged into the tree on a single dimensional plane, Consequently, no special condition exists that prohibits extraction of either the specifier or complement of a regular phrase.

3. Form Sequence and Utilization of CSC

(11) Form Sequence (FSQ, Chomsky 2019b/UCLA, 2020/LSJ, 2021b/WCCFL)

- a. FSQ = $\langle (\&), X_1, \dots, X_n \rangle$: an order-restricted flat-formation operation, especially in order to generate unbounded unstructured sequences in (12).
- b. FSQ produces a pure sequence, yielding a flat structure where there is no formal c-command relation, but there is a strong tendency for a left-to-right counterpart.

- (12) a. I met someone young, happy, eager to go to college, tired of wasting time, ... (Chomsky 2019b:49),
- b. John, Bill, my friends ... ran, danced, took a vacation.(Chomsky 2020:)
- c. John, Mary, the man who lives on the first floor, ... (Chomsky 2021a:8)

- (13) In Pair-Merge, elements in $\langle a, b \rangle$ cannot be accessed or extracted, while in Form Sequence, elements in $\langle a, b \rangle$ can be. Chomsky (2021b/WCCFL) suggests that the reason why the elements in $\langle a, b \rangle$ formed by Form Sequence can be accessed and extracted is that Form Sequence is applied “after” set-Merge is applied to the elements a and b, forming $\{a, b\}$.

◆ FSQ & CSC—regular coordinator structure

- (14) a. John lived on a farm with his family.
 b. John lived on a farm and with his family. (Chomsky 2021: 32)
- (15) a. Which farm did John live on—~~which farm~~ with his family? → **set / wh-movement**
 b. *Which farm did John live on—~~which farm~~ and with his family?
 → **sequence / *CSC**
- (16) a. I wonder what [{John bought t_i } and {Bill handed t_i to Tom}]
 b. *I wonder what [{John bought t_i } and {Bill handed **a sandwich** to Tom}]
 (Chomsky 2021: 34)
- (17) a. I wonder [what₃, C, <&, {John bought ~~what~~₁}, {Bill handed ~~what~~₂ to Tom} >]
 → **sequence / ATB**
 b. I wonder [what₁, C, <&, {John bought ~~what~~₂}, {Bill handed **a sandwich** to Tom} >] → **sequence / *CSC**

◆ FSQ & CSC—exceptional coordinator structure

- (18) a. *Which surgeon_i did Kim date [friends of t_i and a lawyer]? (Ross 1967: 89)
 b. What_i did Harry go to the store and buy t_i ? (Lakoff 1986:152–153)
- (19) a. *Which surgeon C did Kim date <&, {friends of ~~which surgeon~~} {a lawyer} >]?
 → **sequence / *CSC**
 b. What did Harry {go to the store and buy what}? → **set / wh-movement**

◆ FSQ & Adjunct Condition Construction

- (20) a. *Which girl_i did John arrive [after Bill kissed t_i]? (Boeckx 2012: 16)
 b. *[which girl C <&, {did John arrive} {after Bill kissed which girl} >]
 → **sequence / *CSC**

◆ FSQ & Parasitic Gap Construction

- (29) a. What did you file without reading? (Boskovich 2020: 6)
 b. *What did you file the book without reading?
- (30) a. [what₁ C <&, {did you file ~~what~~} {without reading ~~what~~} >]
 → **sequence / ATB**
 b. *[what₁ C <&, {did you file the book} {without reading ~~what~~} >]
 → **sequence / *CSC**

5. Conclusion

◆ Proposal

We will suggest that CSC based FSQ system can explain CSC–regular coordinate structure, CSC–exceptional coordinate structure, Adjunction Constraint structure, and Parasitic gap structure.

6. References

- Bošković, Željko. 2020. On Unifying the Coordinate Structure Constraint and the Adjunct Condition, In A. Błrłny, T. Biberauer, J. Douglas, and S. Vikner (eds.) *Syntactic architecture and its consequences II : Between syntax and morphology (Open Generative Syntax 10)*, 227–258, Language Sciences Press
- Chomsky, Noam. 2019a. Some Puzzling Foundational Issues: The Reading Program. *Catalan Journal of Linguistics Special Issue*, 263–285.
- Chomsky, Noam. 2019b. Fundamental Issues in Linguistics, MIT Lectures, April 10th and 12th.
- Chomsky, Noam. 2019c. UCLA Lectures, April 29th, 30th, May 1st, 2nd.
- Chomsky, Noam. 2021. “Minimalism: Where are we now, and where can we hope to go. *Gengo Kenkyu 160*, 1–41.
- Chomsky, Noam. Ingal Gallego, and Dannis Ott. 2019. Generative Grammar and the Faculty of Language: Insights, Questions, and Challenges. *Catalan Journal of Linguistics*, 229–261.
- Goldsmith, John. 1985. A principled exception to the coordinate structure constraint. In *CLS 21. Part 1: Papers from the General Session*, ed. William H. Eilfort, Paul D. Kroeber and Karen L. Peterson, 133–143. Chicago: Chicago Linguistic Society.
- Grosu, Alexander. 1973. On the nonunitary nature of the coordinate structure constraint. *Linguistic Inquiry 4*. 88–92.
- Johannessen, Janne Bondi. 1998. *Coordination*. Oxford: Oxford University Press.
- Lakoff, George. 1986. Frame Semantic Control of the Coordinate Structure Constraint. In *Papers from the Chicago Linguistic Society 22, Part 2, The Parasession on Pragmatics and Grammatical Theory*, ed. Anne M. Farley, Peter T. Farley and Karl–Erik McCullough, 152–167. Chicago: Chicago Linguistic Society.
- Kiyang Kwon & Lee. Wonbin 2021. Notes on Adjunct Condition and MERGE. *Journal of Language Sciences 28*(1), 39–58.
- Munn, Alan. 1987. Coordinate structure and X–bar theory, *McGill Working Papers in Linguistics 4*, McGill University.
- Napoli, Donna Jo. 1993. *Syntax: theory and problems*. New York : Oxford University Press.
- Oda, Hiromune. 2017. Two types of the coordinate structure constraint and rescue by PF deletion. In Andrew Lamont & Katerina Tetzloff (eds.), *Proceedings of the North East Linguistic Society*, vol. 47, 343–356. Amherst, MA: Graduate Linguistic Student

- Association, University of Massachusetts.
- Pesetsky, David. 1982. *Paths and categories*. Doctoral dissertation, MIT, Cambridge, MA.
- Ross, John Robert. 1967. *Constraints on variables in syntax*. Cambridge, MA: Massachusetts Institute of Technology Dissertation.
- Sag, Ivan A., Gerald Gazdar, Thomas Wasow & Steven Weisler. 1985. Coordination and how to distinguish categories. *Natural Language and Linguistic Theory* 3. 117-171.
- Yosuke Sato and Jacky Chua. 2019. A Dual-Structure Analysis of Coordination, ms. lingbuzz/004704
- Zoerner, Cyril Edward, III. 1995. *Coordination: The Syntax of andP*. Doctoral Dissertation, University of California, Irvine.
- Zhang, Niina Ning. 2007. The Syntactic Derivations of Two Paired Dependency Constructions. *Lingua* 117:2134-2158.
- Zhang, Niina Ning. 2009. *Coordination in syntax*. Cambridge: Cambridge University Press.

Linguistic synaesthesia in language contact: A Sino-Korean compound perspective

Jo, Charmhun

Southwest University, China

Introduction

Synaesthesia signifies involuntary brain phenomena where “sensory events in one modality take on qualities usually considered appropriate to another” (Marks 1982: 15). On the other hand, linguistic synaesthesia (for instance, *warm color*) is usually accepted as metaphor (Geeraerts 2010). That is, *warm color* is often treated as metaphor and in this expression a perception of sight (*color*) is described using a word related to touch (*warm*), where *color* becomes the target domain of the transfer and *warm* is the source domain.

Since the idea of synaesthetic metaphors was introduced by Stephen Ullmann (1963[1957]), his theoretical framework of hierarchical directionality has been generally accepted as a probable universal principle in the process of synaesthetic mapping. Subsequent examinations have gradually expanded from English to other languages such as Italian, Hebrew, Chinese and from poetic texts to conventional language data (e.g., Williams 1976; Shen 1997; Yu 2003; Strik Lievers 2015).

On the other hand, several studies have recently reported somewhat different findings with regard to linguistic synaesthesia (e.g., Zhao and Huang 2015; Zhao et al. 2019). According to Zhao et al.’s (2019) report, for instance, the corpus-based analysis of synaesthetic metaphors in Mandarin Chinese presents that Mandarin synaesthesia has three types of directionality models (namely, the rule-based unidirectionality, the frequency-based biased-directionality, and the bidirectionality without any directional preference), which strongly challenges the universality hypotheses claimed by Ullmann (1963[1957]) and Williams (1976). The recent discoveries that challenge the so-called cross-linguistic universality of linguistic synaesthesia mostly based on the data analysis from Indo-European languages suggest language/culture-specific variations in linguistic synaesthesia.

In this context, I then note synaesthetic phenomena in language contact, which no study has looked at in different languages so far. That is to say, if the directionality of linguistic synaesthesia is language-universal, then language contact or borrowing will not significantly affect it. However, if there are significant differences in synaesthetic directionalities between two languages in language contact, that is, loan words from the source language show any differences with indigenous words in the synaesthetic directionality, it would be a serious challenge to the cross-linguistic universality claim of linguistic synaesthesia.

In this paper, accordingly, I propose to investigate the linguistic synaesthesia of Sino-Korean compound words, comparing it with the directionality of synaesthetic metaphors in pure Korean compounds. Furthermore, the current study conducts the examination of morpho-semantic aspects of Sino-Korean words, which can make important contributions to the study of Sino-Korean, given that most traditional studies on Sino-Korean have focused on phonology (e.g., Eom 1999, 2015, 2019a, 2019b; Han 2001).

Directionalities of linguistic synaesthesia in different languages

As a seminal study of linguistic synaesthesia, Ullmann (1963[1957]) presented a theoretical framework of hierarchical directionality by analyzing the 19th century poetic texts invented in English, French, and Hungarian. In the framework, he proposed three overall tendencies in synaesthetic mappings: first, the majority of synaesthetic transfers have the standardized direction of “Touch □ Heat □ Taste □ Smell □ Sound □ Sight”, where the transfers tend to move from the lower sense to the higher sense; second, derived from the first tendency, the most frequent source mode of mappings is tactility, the lowest stage of the senses; third, the most frequent target mode for synaesthetic mappings is audition, not vision.

Following Ullmann's (1963[1957]) study of synaesthetic metaphors for poetic language, many researchers have investigated linguistic synaesthesia in poetic textual data (e.g., Shen 1997; Shen and Cohen 1998; Jo and Jhang 2019) or in daily-language corpora (e.g., Strik Lievers 2015; Zhao et al. 2018; Zhao et al. 2019). They mostly focused upon the phrasal or clausal synaesthesia analysis within a text. In this context, an exceptional study of synaesthetic metaphor is that conducted by Williams (1976), who explored synaesthetic transfer patterns in everyday language lexicon, that is, the historical change of the meanings of sensory adjectives in everyday English.

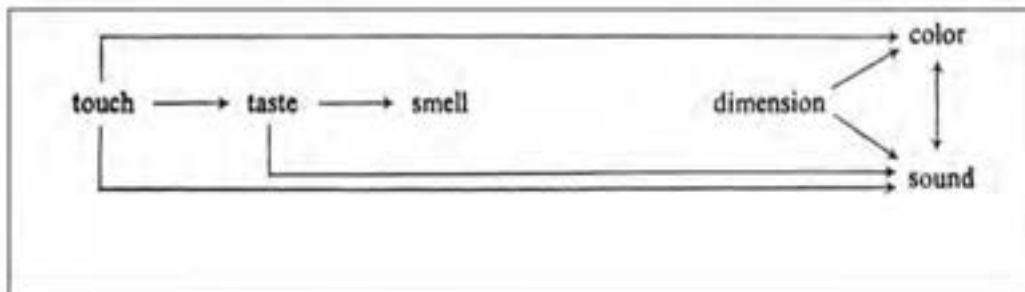


Figure 1. Williams's (1976: 463) synaesthetic transfer route

Research Questions

1. What type of mapping directionality does Sino-Korean and pure Korean compound synaesthesia show? And how are both of them different or similar?
2. How are the two synaesthetic directionalities revealed, compared with previous studies?
3. What are the features of Sino-Korean synaesthesia based on the morpho-semantic analysis?

Data collection

For Sino-Korean compound-word synaesthesia, data collection has started with several compound synaesthesia examples discovered in daily language use. They have been then extended through Korean WordNet (<http://www.wordnet.co.kr/>) and the *Standard Korean Grand Dictionary* (<http://stdweb2.korean.go.kr/main.jsp>). For instance, the related synaesthesia compounds of an existent compound-word synaesthetic example of 경음 硬音 *kyengum* ‘fortis’ were detected based on Korean WordNet and the *Standard Korean Grand Dictionary*, as shown in (1):

- (1) 경음 硬音 *kyengum* ‘fortis’ > 강음 強音 *kangum* ‘strong sound’, 약음 弱音 *y*
a

kum ‘weak sound’, 된소리 *toynsoli* ‘fortis’

Table 1. Korean compound-word synaesthesia gathered for the study (type)

Korean compound synaesthesia	SK compound synaesthesia	IK compound synaesthesia	ISK compound synaesthesia
Sub-total	19	18	4
Total	41		

Synaesthetic data from Sino-Korean and pure Korean compounds

As aforementioned, I have totally gathered 41 types of Korean compound synaesthesia via the *Standard Korean Grand Dictionary*, Korean WordNet, and the native Korean author’s intuition. Table 2 shows their mapping distribution between sensory domains based on Sino-Korean (SK), indigenous Korean (IK), and indigenous and Sino-Korean (ISK) compounds. In order to further understand Sino-Korean compound synaesthetic phenomena in Korean, I need to analyze the data word by word in detail, comparing mutually corresponding pairs of SK, IK, and ISK compounds.

Table 2. Distribution of synaesthetic mappings between sensory domains in Korean compounds (type)

Synaesthesia Mapping	SK compound synaesthesia	IK compound synaesthesia	ISK compound synaesthesia	Total
Touch □ Taste	1	1	1	3
Touch □ Sight	2	1	0	3
Touch □ Hearing	3	3	0	6
Taste □ Smell	0	4	0	4
Taste □ Sight	1	1	0	2
Taste □ Hearing	2	1	0	3
Sight □ Smell	2	0	0	2
Sight □ Hearing	8	7	3	18

Total	19	18	4	41
-------	----	----	---	----

Mapping directionalities of linguistic synaesthesia of Sino-Korean and pure Korean compounds

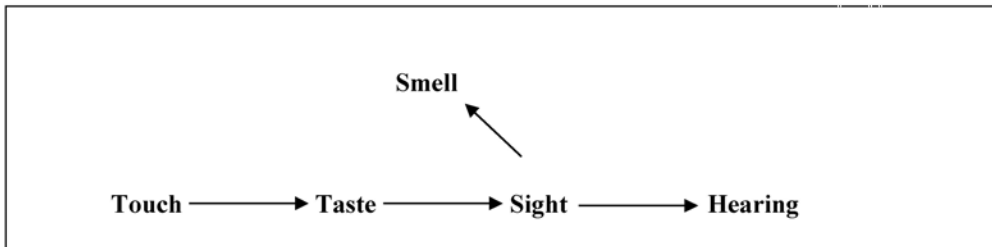


Figure 2. Mapping directionality of Sino-Korean compound synaesthesia

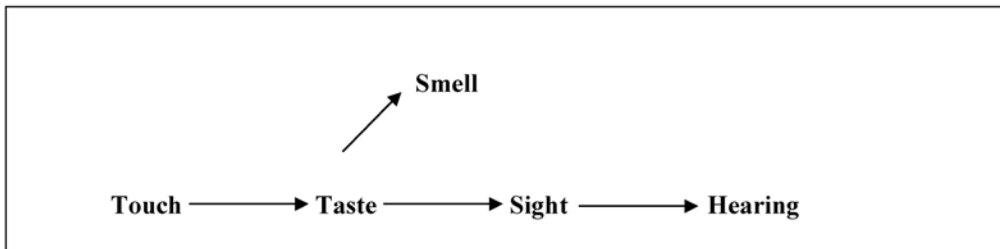


Figure 3. Mapping directionality of indigenous Korean compound synaesthesia

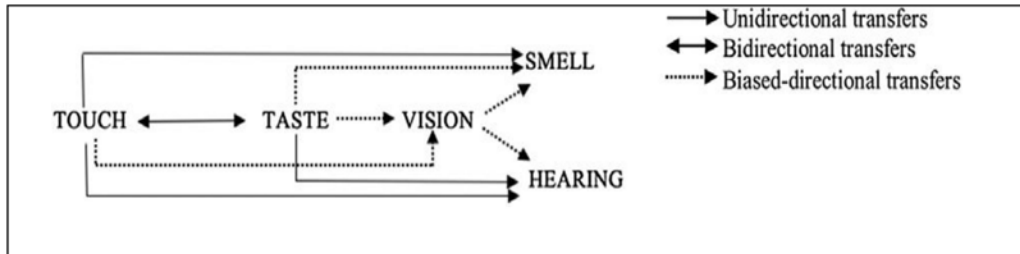


Figure 4. Zhao et al.'s (2019: 9) directionality model of Mandarin synaesthesia

To sum up, both findings of mapping directionality of linguistic synaesthesia between Sino-Korean and indigenous Korean compounds do not show a similar linear hierarchy, as observed in Figures 2 and 3. Furthermore, Sino-Korean synaesthesia is analogous to Mandarin Chinese synaesthesia analyzed by Zhao et al. (2019) in the transfer directionality, whereas indigenous Korean synaesthesia parallels Indo-European language synaesthesia studied by Williams (1976) and Korean synaesthesia discussed by Jo (2019). It means that the directionality of linguistic synaesthesia in language contact is not language-universal, because loan words (i.e., Sino-Korean) still tend to follow the source language regulation in regards to transfer directionality, showing considerable dissimilarity with native words (i.e., Korean). Thus, this study presents that linguistic synaesthesia has language/culture-bound variations for directional tendency, which strongly challenges universality theories of linguistic synaesthesia that has been claimed based on Indo-European languages (e.g., Ullmann 1963[1957]; Williams 1976; Shen 1997).

Additional morpho-semantic analyses of Sino-Korean compound synaesthesia

Number	Compound word	Frequency	
		Token	Type
1	큰소리 <i>khunsoli</i> 'vain boast or scolding'	943	1
2	냉소 冷笑 <i>nayngso</i> 'cold smile'	513	1
3	잔소리 <i>cansoli</i> 'useless talk'	488	1
4	쓴웃음 <i>ssunwusum</i> 'bitter smile'	161	1
5	고함 高喊 <i>koham</i> 'big voice'	155	1
6	감칠맛 <i>kamchilmas</i> 'savory taste'	109	1
7	고음 高音 <i>koum</i> 'high sound'	84	1
8	된소리 <i>toynsoli</i> 'fortis'	76	1
9	저음 低音 <i>ceum</i> 'low sound'	74	1
10	잔말 <i>canmal</i> 'useless talk'	68	1

11	쓴소리 <i>ssunsoli</i> ‘bitter remarks’	60	1
12	고언 곱말 <i>koen</i> ‘bitter word’	50	1
13	단내 <i>tannay</i> ‘sweet smell’	48	1
14	감언 곱말 <i>kamen</i> ‘sweet word’	44	1
15	경음 硬音 <i>kyengum</i> ‘fortis’	29	1

Table 3. Frequency list of Korean compound synaesthesia in Sejong Modern Written Corpus

Group	Compound word	Frequency	
		Token	Type
1	냉소 冷笑 <i>nayngso</i> ‘cold smile’	513	1
	찬웃음 <i>chanwusum</i> ‘cold smile’	0	1
2	고언 곱말 <i>koen</i> ‘bitter word’	50	1

	쓴소리 <i>ssunsoli</i> ‘bitter remarks’	60	1
3	고소 苦笑 <i>koso</i> ‘bitter smile’	7	1
	습소 濕笑 <i>supso</i> ‘bitter smile’	0	1
	쓴웃음 <i>ssunwusum</i> ‘bitter smile’	161	1
4	고함 高喊 <i>koham</i> ‘big voice’	155	1
	고언 高言 <i>koen</i> ‘vain boast’	0	1
	큰소리 <i>khunsoli</i> ‘vain boast or scolding’	943	1
5	세설 細說 <i>seysel</i> ‘useless talk’	1	1
	잔사설(辭說) <i>cansasel</i> ‘useless talk’	1	1
	잔소리 <i>cansoli</i> ‘useless talk’	488	1
	잔말 <i>canmal</i> ‘useless talk’	68	1

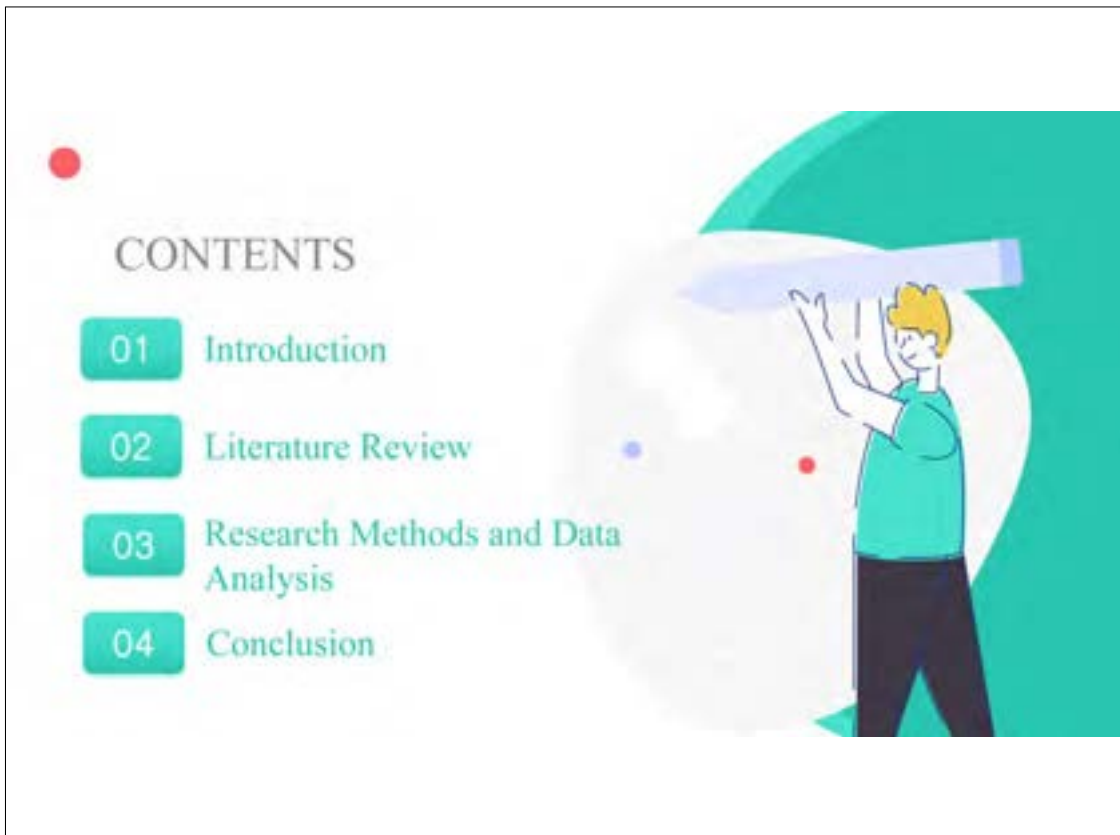
Table 4. Five groups of Korean compound synaesthesia


Group	Compound word	Perceptual intensity	Formality	Sentiment polarity
1	냉소 冷笑 <i>nayngso</i> ‘cold smile’	+	+	-
	찬웃음 <i>chanwusum</i> ‘cold smile’	+	0	-
2	고언 苦言 <i>koen</i> ‘bitter word’	+	+	+
	쓴소리 <i>ssunsoli</i> ‘bitter remarks’	+	-	+/-
3	고소 苦笑 <i>koso</i> ‘bitter smile’	-	+	-
	습소 濕笑 <i>supso</i> ‘bitter smile’	-	+	-
	쓴웃음 <i>ssunwusum</i> ‘bitter smile’	+	0	-
4	고함 高喊 <i>koham</i> ‘big voice’	+	-	-
	고언 高言 <i>koen</i> ‘vain boast’	-	+	-
	큰소리 <i>khunsoli</i> ‘vain boast or scolding’	+	-	-

5	세설 細說 <i>seysel</i> ‘useless talk’	0	+	-
	잔사설(辭說) <i>cansasel</i> ‘useless talk’	+	-	-
	잔소리 <i>cansoli</i> ‘useless talk’	+	-	-
	잔말 <i>cammal</i> ‘useless talk’	+	-	-

Table 5. Semantic contents in groups of Korean compound synaesthesia

Comprehensively speaking, Sino-Korean compounds are in general more formal than indigenous Korean ones, whereas the former is usually lower than the latter in perceptual intensity and use frequency. It signifies that Sino-Korean language based on Chinese characters should be still studied proactively by Korean people as a loaned/foreign language (not as a mother language) in the process of language contact of Korean and Chinese. In addition, it is noticeable that most Sino-Korean synaesthesia examples, along with native Korean synaesthetic words, act in the negative sentiment (see Table 5). However, based on the whole data of this study, it cannot be suggested that Sino-Korean synaesthesia as well as native Korean synaesthesia has any specific bias in sentiment polarity. The Sino-Korean compound’s sentimental negativity (or positivity) for semantic distribution requires a further, in-depth study in future.





1. The parts of speech in language:

- Content words: nouns, verbs, adjectives
- Function words: preposition, connectives, pronouns and demonstratives

Adverbs: in the intermediate status, some of which are classified into function words.

2. Role of Grammaticalization:

Crucial to the formation of the function words, many of which were derived from the content words, and has contributed greatly to the expansion of the lexicons.

3. Grammaticalization in Chinese:

Some Chinese content words are changing to be function words in the process of grammaticalization, constantly producing new vocabulary, some of which are degree adverbs.

4. Previous Studies:

5. Argumentation in the present paper:


6. Research Methods:

- Their classifications (Wang, 1943; Ma, 1988; Zhang, 2004, etc.
- Their functions (Pei, 2009; Xie and Yuan, 2009; and Yao, 2010);
- Cross-linguistic comparisons between Chinese and other languages (Guo, 2013; Zhao and Fang, 2017; and Li, 2018).
- Some on degree adverbs in Chinese dialects (Yu, 2018; Gong and Huang, 2018; Yang, 2019).
- Special grammatical phenomena of degree adverbs (Li, 2012; Zhang, 2017; Zhang, 2019).
- Few studies on the grammaticalization of the new emerging degree adverbs.
- - Prototypical cases of grammaticalization: "hǎo(very)" and "nǎi(too)" are originally verbs, and changed into complete function words.
- The new emerging adverbs "chāo(super)", "zì(third)", "wài(bare)", "jǐ(giant)" still hold parts of their lexical stages. Some scholars argued that they already hold the basic feature of the degree advs, and have been defined as the degree advs in Chinese dictionary.

1. "chāo", "zì", "wài", "jǐ" belong to different levels or stages of grammaticalization, given that they have their rather different frequencies of use and/or their more or less different structural distributions in the real texts.

2. The frequency of use as advs in the data is crucial to determine the level of grammaticalization, and the observation of their structural distributions is also a quite important process of defining in which level of grammaticalization a degree adverb is.

- Three corpus for data analysis:
 - Beijing Language Chinese University Corpus (BBC): 13 billion modern words & 2 billion ancient words.
 - Chinese Linguistics Corpus (CCL): 200 million modern words & 500 million ancient words.
 - Chinese Corpus (CNCO): 129 million modern words & 100 million ancient words.





PART 02

Literature Review

- The classification and function of some prototypical form of the Chinese conventional degree adverbs
- The adverbialization process of the four new degree adverbs.
- The theory of grammaticalization.

2.1 Conventional Degree Adverbs: Classification and Types

1. Chinese degree adverbs:

A closed word class, less morphological changes. (English: adj+-ly) (Guo, 2013; 7,14)

2. Classifications of the Chinese degree adverbs:

- (1) Two main types, absolute degree adverbs & relative degree adverbs (Wang, 1943: 131-132).
- (2) Controversial in subdivision:
 - Ma (1988: 83-86): according to their grammatical meanings.
 - Zhang (1997: 92-96): according to their magnitude in the five types of comparative sentences.
 - Chu (1999: 36-44): according to the compared objects with the relationship between the individual and entirety.
 - Zhang (2000: 51-63): according to the comparable relationship and the gradable semantic differences of the absolute and relative.
 - Lin and Guo (2003: 74) classify them according to the hierarchy of magnitude.
- (3) The classifications are made with reference to the three parameters: comparative objects, comparative relation, and degree of magnitude.
- (4) The classification of the Conventional Degree Adverbs.

Table 1. The classification of the conventional degree adverbs

Word Class	Mean value	Subclasses	Word	Mean (Frequency in BEC)
Absolute Degree Adverbs	Degree of Total Appl.	Higher	全 (all), 全 (entire), 全 (whole)	全 (all), 全 (entire), 全 (whole)
		Lower	半 (half), 半 (part), 半 (portion)	半 (half), 半 (part), 半 (portion)
		Equal	等 (equal), 等 (same), 等 (equivalent)	等 (equal), 等 (same), 等 (equivalent)
		Contrastive	反 (opposite), 反 (reverse), 反 (contrary)	反 (opposite), 反 (reverse), 反 (contrary)
Relative Degree Adverbs	Degree of Part Appl.	Higher	更 (more), 更 (further), 更 (deeper)	更 (more), 更 (further), 更 (deeper)
		Lower	少 (less), 少 (fewer), 少 (smaller)	少 (less), 少 (fewer), 少 (smaller)
		Equal	同 (same), 同 (similar), 同 (alike)	同 (same), 同 (similar), 同 (alike)
		Contrastive	反 (opposite), 反 (reverse), 反 (contrary)	反 (opposite), 反 (reverse), 反 (contrary)

3. Structural distribution of the Chinese degree adverbs: (two main types)

Simple Structure:

(1) *adverb + modified element*. (*adv* modifies the head elements and operates as an adverbial)

Main function: modify other elements. (most are adjs, few verbs, and occasional nouns)

- Modified Adjectival Elements:

a. + **conditional adj.**

E.g., Ta men dou yu zhe xie xie yong xi ji de shuang shou zao chu lai de zao pin, gan diao *fen cheng jiao shi*. (CNCO)
 'They felt so proud of these masterpieces made by their own hands.'

b. + **trait adj.**

E.g., Bei jiu xia du, jiu te xin geng ri. (CNCO)
 'With one alcohol drunk, the heart is even better than the alcohol.'

- Modified Verbal Elements:

a. + **mental verb**

E.g., Suo xue ke mu zhong ta zai xi huan ying yu. (CCCL)
 'Among all the subjects, he likes English most.'

b. + **verbal phrase**

E.g., Wo tu ran yu Shi Xiang jiao Chao. (BBC)
 'I suddenly extremely want to meet Cristiano Ronaldo.'

c. + **general verb**: + **verbal phrase**

E.g., Ta dai ta hen hao de le, zhi dao ta hen xiu xiu ji. (BBC)
 'He knows her quite well, and know how charming she is.'

d. + **verbal phrase**:

E.g., Zhou li bu: "wo ting hen hao de xi ji xin tai de..." (CCCL)
 'Zhou libo said: "I can adjust my mentality very well..."'

e. + **verbal phrase**:

E.g., Huai sui zhe me shuo, dian kan de chu ta dai jiu guo yu dian er bu man yi. (CCCL)
 'Even saying like that, it is obvious to find that he is not quite satisfied with the result.'

(2) *modified elements + de (structural auxiliary) + adv* (restricted only to the conventional degree advs "bian")

E.g., a. Yue Bu Qun que zong shi yu xiu di di hen. (CCCL)

b. *Yue Bu Qun que zong shi yu xiu di di hen cheng.
 'Yue Bu Qun is always with a very big ambition.'

(3) *modified elements + adverb (+ Modal Auxiliary "le")*

a. (only disyllabic degree adv)

E.g., dao le Dao Guang nian jian, wo men de fa zhi yu ming wu shi, guan li fa bu, min sheng *ting jiu yin jin*. (CCCL)

'In Dao Guang years, our legal system was nominally and the officials were corrupted, people lived in deep agony.'

b. (or monosyllabic degree adv + "le")

E.g., ta jiao de hen xiao ji de. (CCCL)
 'He thinks it's extremely ridiculous.'

Compound Structure:

(1) **Attributive Compound Structure:** modify Subject or Object.
 • **adverb + modified element + de (structural auxiliary)**
 E.g., Rén hé rén zài gōng hǎn de guān xi jù shì jù jī guān xi. (CCLJ)
 "The most basic relationship among people is the class relation."

(2) **Adverbial Compound Structure:** modify Predicate
 • **adverb + modified element + di**
 E.g., dǎng nián lái mǎn zhū zhuō shàng shì, rén men fēn zī yì zhī de xīng qǐng dǎo le Yán lín shì dà. (CNCO)
 "When talking about the spirit of democracy, people are thinking of the Yan an era very naturally."

(3) **Complemental Compound Structure:** modify Predicate
 • **de + adverb + modified elements** (a deviation of the normal structure, adv has been postposed)
 E.g., a. wǒ zhǎng jù de zhè lǐ bǐ jiào shì hěn de gōng yòng xīn xi, gōng háo wán. (CNCO)
 "I think that having class here is more interesting than in the classroom."
 b. tā men de dòu zhēng hěn de shì shì jìn xi. (CNCO)
 "The combat among them shows to be extremely sharp."

(4) **Negative structure:**
 • **Express negation:** běn (or others) + bù (negative adv) + modified elements
 E.g., a. guò jì xīng shì dòng dǎng duǎo biàn, bù cǐ yīn sù zōng duǎn, tiān xià hái fēn fēn bù bù zēng. (CNCO)
 b. guò jì xīng shì dòng dǎng duǎo biàn, bù cǐ yīn sù zōng duǎn, tiān xià hái shì shì bù bù zēng.
 "The world is still very unpredictable in the global situation is in chaos and changeable, and the unpredictable factors are also increasing."
 • **Reduce the degree:** bù (negative adv) + nǐ / gāo / dà (only this three) + modified elements
 E.g., a. wǒ gān dào wǒ men lì cǐ hái bù shì rú guān. (CNCO)
 "I felt like that we have not that familiar with each other yet."
 b. *wǒ gān dào wǒ men lì cǐ hái bù rú guān.

2.2 The Adverbialization of the New Degree Adverbs

1. The adverbialization of "chāo", "zēi", "mán", "jū":

- the One important stream of Adverbialization: adverbs mainly derive from verbs, adjectives, or even nouns (Zhang, 2000:3).

"chāo"




(1) Original meaning: as a verb, "jump" (from the Origin of Chinese Characters (1963: 36)).
 (2) In modern Chinese: multiple identities, like verb, adjective, a single word or a morpheme in phrases.
 (3) Adverbialization of "chāo":
 a. Some scholars: completed during the 'Northern and Southern Dynasties of China'.
 b. Zhu (2018: 74-75): the usage of "chāo" as degree adverb is a result of the combined effects with the introduction of foreign languages, the pragmatic interpretations of Chinese expression, and the inheritance of the developing rules.
 c. Zhou (2006:47): The semantic change of "chāo" from a morpheme (adj) to a single word (degree adverb) plays a decisive role, but also restricted by pragmatics in a great extent.
 d. The Modern Chinese Dictionary (7th Edition, 2016: 32): adv, a high degree of expression.

"zēi"




(1) Original meaning: a verb, "to massacre, to harm".
 (2) Transformations:
 a. Evolved to be a noun, "one who carries out destruction, assassin", further expanded into "traitor, the robber of property, thief".
 b. Continue to be an adj, "cunning, wicked".
 c. During the "Ming and Qing dynasties", started to stand in front of the verbs and adjs, as an adv (Wang, 2020: 79).
 d. At first, had a strong regional color, and was mostly categorized as a degree adverb in the dialects in northeastern China.
 e. The Modern Chinese Dictionary (7th Edition, 2016: 1638): adv, "very, very much", represents a high degree.

"mán"



(1)Original meaning: as a noun, "a self-named term given by minority groups in the south of ancient China", then refers to a representation of a minority groups in the south, but also refer to the surrounding minority groups (Zhao and Zhou, 2010: 117-118).
(2)Transformation:
a. At first, people of the Central Plain use "mán" to speak out their derision and contempt to the southern region.
b. The pejorative color gradually faded, can also to describe the positive meaning.
c. Due to the preservation of the self-esteem of the southern people, it turns to be a degree adverb to praise in the southern dialects.
d. The function as a degree adv has been authorized and the emotional color has been diluted.
e. The Modern Chinese Dictionary (7th Edition, 2016: 874): adv, "very, quite".

"jū"




(1)Original meaning: the expansion of the meaning of rules and regulations, then replaced by "jū"(rectangle, rule).
(2)Transformation:
a. From the 'Spring and Autumn period and the Warring States period': as an adj. "big", and together with many other meanings (Wang, 2019: 38-41).
b. Gradually placed in front of adjs and verbs, as an adv, "most, extremely".
c. In modern chinese: the multiple meanings gradually disappeared, and began to evolve into the simple meaning "big", and occurs as a morpheme in a phrase more frequently than as a single word.
d. Becomes a kind of popular network language, can be searched as an adv on The Chinese Dictionary Online. The degree adv of "ju" is gradually transferring from spoken language to written language (Gao and Yang, 2006: 34).

2. The features of the adverbialization of "chāo", "zéi", "mán", "jū":

- (1) They are distinctly territorial and have a strong emotional dimension.
- (2) The content words themselves are intended to have a great influence on the modifying functions of the adverbialization words.
- (3) With regard to their classification:
 - They can express an extremely high degree and do not require a comparison with other objects, which express one's own statement of mood or a description of fact.
 - a. wǒ jīn tiān chāo hài xìng. (BBC)
'I'm extremely happy today.' (One's own statement of a highly expressed mood)
 - b. tāng dào zhè gè xiāo xī tā zéi jī dòng. (BBC)
'He was so thrilled to hear the news.' (One's own statement of a highly expressed mood)
 - c. suī rán qián bù duō, dàn wǒ mán mǎn zú de. (BBC)
'Even I don't have too much money, I'm quite satisfied.' (One's own statement of a highly expressed mood)
 - d. zhè gè rén jū yóu qián. (BBC)
'This man is extremely rich.' (An exaggerated description of a fact.)
 - They all express a high degree of meaning as "very/extremely", thus are all intensifying modifiers, absorbed most of the grammatical functions and features from the conventional degree adverbs.

➔ To check whether they are fully grammaticalized.



2.3 Grammaticalization

Meillet (1912: 131) and several linguistics:

First raised this term.

Saussure (1971: 58):

Grammaticalization "consists of intake from lexis", it takes place when a word becomes "sufficiently empty of lexical meaning".

Hopper and Traugott (2003: 1):

Concerned with such questions as how lexical items and constructions come in certain linguistic contexts to serve grammatical functions or how grammatical items develop new grammatical functions.



Kurylowicz (1965:52):

The term is a process which turns lexemes into grammatical formatives and makes grammatical formatives still more grammatical (widely recognized as the most typical definition).

Heine, et al. (1991: 2):

We are dealing with grammaticalization where a lexical unit or structure assumes a grammatical function, or where a grammatical unit assumes a more grammatical function, which is on both synchronic and diachronic sides. And grammaticalization is unidirectional, from less grammatical to more grammatical.



- One important way to define the level of grammaticalization is to count the frequency of use (Heine, et al., 1991:38)



- Research method:
 - (1) Selected a total of 1000 items for each of the four new degree adverbs (according to the proportional ratios of total numbers from each corpus):
 - 100 items from CNCO
 - 400 items from BBC
 - 500 items from CCL
 - (2) From the selected items, we investigate their word classes in the sentences and frequency of use as adverbs to evaluate their levels of grammaticalization.

Results:



Table 2. The Frequency of Word Classes for the Four Words in Three Corpus

Data searched in corpus (1000 sentences in total) with word 'so'				
	CNCO(1000sentences)	BCC(4000sentences)	CCL(5000sentences)	Total
Word Class	Volume (N)	Volume (N)	Volume (N)	Volume (N)
adverb	11 (1%)	5 (0.125%)	2 (0.04%)	18 (0.18%)
adjective phrase	2 (0.2%)	28 (0.7%)	63 (1.26%)	93 (0.93%)
verb	0 (0%)	0 (0%)	7 (0.14%)	7 (0.07%)
nominal phrase	17 (1.7%)	60 (1.5%)	66 (1.32%)	143 (1.43%)
verbal phrase	0 (0%)	1 (0.025%)	0 (0%)	1 (0.01%)
others	1 (0.1%)	9 (0.225%)	0 (0%)	10 (0.1%)

Data searched in corpus (1000 sentences in total) with word 'soo'				
	CNCO(1000sentences)	BCC(4000sentences)	CCL(5000sentences)	Total
Word Class	Volume (N)	Volume (N)	Volume (N)	Volume (N)
adverb	1 (0.1%)	0 (0%)	1 (0.02%)	2 (0.2%)
adjective phrase	1 (0.1%)	0 (0%)	23 (0.46%)	24 (0.24%)
verb	17 (1.7%)	15 (0.375%)	41 (0.82%)	73 (0.73%)
nominal phrase	20 (2%)	22 (0.55%)	11 (0.22%)	53 (0.53%)
verbal phrase	1 (0.1%)	0 (0%)	11 (0.22%)	12 (0.12%)
others	12 (1.2%)	11 (0.275%)	11 (0.22%)	34 (0.34%)
total word	0 (0%)	100 (2.5%)	0 (0%)	100 (1%)

Data searched in corpus (1000 sentences in total) with word 'soo'				
	CNCO(1000sentences)	BCC(4000sentences)	CCL(5000sentences)	Total
Word Class	Volume (N)	Volume (N)	Volume (N)	Volume (N)
adverb	1 (0.1%)	0 (0%)	10 (0.2%)	11 (0.11%)
adjective phrase	0 (0%)	1 (0.025%)	1 (0.02%)	2 (0.02%)
verb	40 (4%)	10 (0.25%)	30 (0.6%)	80 (0.8%)
nominal phrase	14 (1.4%)	130 (3.25%)	200 (4%)	344 (3.44%)
verb	1 (0.1%)	0 (0%)	2 (0.04%)	3 (0.03%)
verbal phrase	2 (0.2%)	0 (0%)	1 (0.02%)	3 (0.03%)
others	2 (0.2%)	10 (0.25%)	0 (0%)	12 (0.12%)


Data searched in corpus (1000 sentences in total) with word 'soo'				
	CNCO(1000sentences)	BCC(4000sentences)	CCL(5000sentences)	Total
Word Class	Volume (N)	Volume (N)	Volume (N)	Volume (N)
adverb	1 (0.1%)	0 (0%)	2 (0.04%)	3 (0.03%)
adjective phrase	10 (1%)	10 (0.25%)	20 (0.4%)	40 (0.4%)
verb	0 (0%)	20 (0.5%)	20 (0.4%)	40 (0.4%)
nominal phrase	10 (1%)	100 (2.5%)	10 (0.2%)	120 (1.2%)
verb	0 (0%)	100 (2.5%)	1 (0.02%)	101 (1.01%)
verbal phrase	0 (0%)	10 (0.25%)	200 (4%)	210 (2.1%)
others	0 (0%)	50 (1.25%)	0 (0%)	50 (0.5%)

Data Analysis:

- The tokens of use as an adverbs of the four words:
 - "chǎo": 110 (11%), the second high
 - "zǎi": 63 (6.3%), initial level
 - "mǎi": 411(41.1%), highest level
 - "yǐ": 10 (1%), initial level
 - "mǎi" > "chǎo" > "zǎi" > "yǐ"
- The highest frequency of the word class for each adv:
 - "chǎo": verbal phrases (39.6%)
 - "zǎi": nominal phrases (66.5%)
 - "mǎi": adverbs (41.1%)
 - "yǐ": adjectival phrases (72.6%)
- The frequency of the phrasal structure of each of the four are much higher than when they are as the single word.
 - "chǎo": verbal phrases (39.6%)
 - "zǎi": nominal phrases (66.5%)
 - "mǎi": adjectival phrases (22.6%) (just a little less than when as adverbs (41.1%))
 - "yǐ": adjectival phrases (72.6%), nominal phrases (22.6%)
- The frequency of use as adverbs when in the three corpus:
 - "chǎo": BCC (24%) > CNCO (8%) > CCL (1.2%)
 - "zǎi": BCC (14.25%) > CNCO (3%) > CCL (0.6%)
 - "mǎi": CNCO (57%) > BCC (55.25%) > CCL (26.6%)
 - "yǐ": BCC (2.25%) > CNCO (1%) > CCL (0%)

BCC includes a wide range of linguistic materials from social networks and media platforms, such as Weibo and People's Daily, while the other two corpus mainly focus on formal texts.

→ All four words are deeply influenced by online culture and structure more frequently in informal language and daily conversations.



Results:


(3) Investigate whether the four new degree adverbs have simple or compound structures to get more evidence of the level of the grammaticalization.

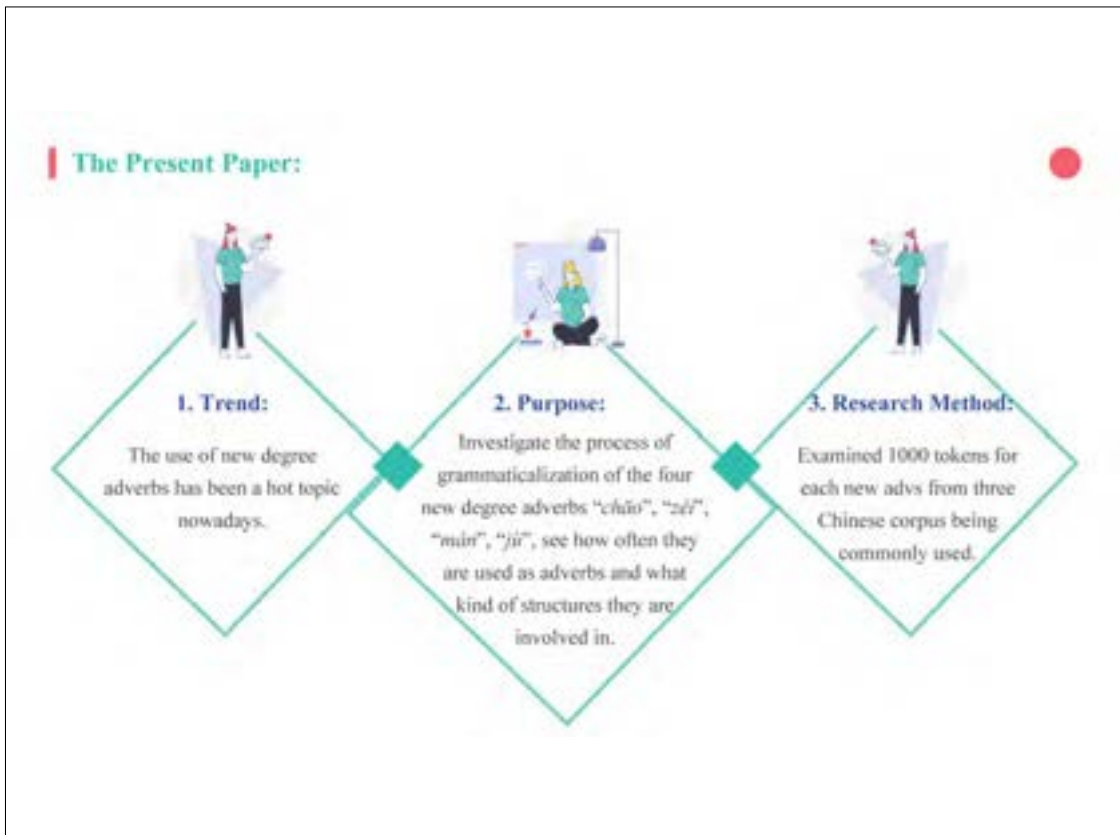
Table 3. Single or Compound Structures for the Use As Adverbs

word items	Total Tokens (adv)	Simple Structures	Compound Structures		
		adv + modified element 1	Attributive structure adv+modified element-adv	Adverbial structure adv+modified element-adv	Complemental structure adv+adv+modified element
yi	10	8 (80%)	2 (20%)	0 (0%)	0 (0%)
chao	110	75 (68.18%)	32 (29%)	3 (2.73%)	0 (0%)
mai	411	347 (84.43%)	5 (1.24%)	0 (0%)	0 (0%)
yi	10	7 (70%)	3 (30%)	0 (0%)	0 (0%)

- Four words all get more simple structures than compound structures.
- Nearly all the words "mǎi" (91.48%) and "zǎi" (92.06%) get the simple structures.
- "chǎo" takes the most compound structures, attributive structures are most often.

→ The four new degree adverbs are in a relatively low level of grammaticalization in that the structures which they get are mostly simple ones.





Results:

(1) Each of the four has a low percentage in the frequency of uses as adverb and their preferences for simple structures instead of compound ones.

(2) The grammaticalization level of the four new degree advs is still in the low stage, except for "mán" in the relatively intermediate level.

(3) "mán" and "chǎo" covers three kinds of ways to form structures, can tell some extent of higher level of grammaticalization.

(4) When compare the three corpus, people prefer to use the new degree adverbs in the oral or media language expressions, not to much in the formal genres, except for "mán".

Thanks for your
listening

E-mail: shiguangmin1018@163.com

The harder you work, the more luck you have.

Gapping in the Conjoined Structure

Huitae Kim

(Keimyung University)

1. Introduction

The goal of this paper is to present a principled explanation on the Gapping in the conjoined structure. Let us consider the sentences such as (1) and (2).

- (1) a. John works at a school and Mary ~~works~~ at a bank.
 b. John will work at a school and Mary ~~will work~~ at a bank.
- (2) a. John ate fish, Mary ~~ate~~ rice, and Tom ~~ate~~ potato.
 b. Bill wants to try to begin to write a novel, and

Jane ~~wants~~ to try to begin to write an essay.

Jane ~~wants to try~~ to begin to write an essay.

Jane ~~wants to try to begin~~ to write an essay.

Jane ~~wants to try to begin to write~~ an essay.

(cf. Ross 1970)

As seen in (1) and (2), the repeated occurrences of the verbs in the coordinate sentences are deleted: *works* in (1a), *will work* in (1b), *ate* in (2a), and *wants*, *wants to try*, *wants to try to begin*, or *wants to try to begin to write* in (2b).

2. Previous Accounts

2.1. Syntactic Deletion Transformation (Ross 1970)

Gapping is considered as the syntactic transformation rule which deletes indefinite number of many occurrences of identically repeated verbs in a conjoined structure. The order of Gapping depends on the order of elements. That is, when identical elements are on left branches, forward Gapping occurs, as seen in (1) and (2), whereas, if they are on right branches, backward Gapping happens, as seen in (3). In addition, Russian can be gapped not only forward to produce (4a) like (1) and (2) but also backward to produce (4b) like (3).

- (3) Watasi-wa sakana-o ~~tabe~~ Bill-wa gohan-o tabeta. (Japanese)
 I-Top fish-ACC eat Bill-Top rice-ACC ate
 ‘I ate fish and Bill rice.’ (cf. Johnson 2006)
- (4) a. Ja pila vodu i Anna ~~pila~~ vodca. (Russian)
 I drank water and Anna ~~drank~~ vodca
 b. Ja vodu ~~pila~~ i Anna vodca pila.
 I water ~~drank~~ and Anna vodca drank

2.2. PF Deletion (Chomsky and Lasnik 1993)

Phonetic form (PF) deletion of Gapping is argued to be a phenomena which deletes the identical verb elements in the PF based on the PF Crash Theory.

- (5) a. John wants to try to begin to write a novel, and
 Mary ~~wants~~ to try to begin to write a play.
 b. Bill wants to try to begin to write a novel, and
 Jane ~~wants to try~~ to begin to write a play.
 c. Bill wants to try to begin to write a novel, and
 Jane ~~wants to try to begin~~ to write a play.
 d. Bill wants to try to begin to write a novel, and
 Jane ~~wants to try to begin to write~~ a play.

In order to determine whether or not to delete the identical verbs in PF, the procrastinate principle called look-ahead should be accepted.

2.3. ATB Movement (López and Winkler 2003)

VP coordination and Across The Board (ATB) raising of the verbs to T is assumed under the framework of D-model where Narrow Syntax, LF and PF are interactive at the same time.

- (6) a. John ate pizza and Jane __ rice.
 b. [_{CP}[_{TP} John [T ate [_{VP} John ate pizza] and [_{VP} Jane ate rice]]].



3. Alternative

So as to give a principled explanation for the Gapping in the conjoined structure such as (1), (2), (3) and (4), López and Winkler’s (2003) D-model is adopted and changed for the derivation in the vein of Syntactic Deletion, rather than PF Deletion or ATB

Movement. And the interpretation is applied in the Interface of C–I and S–M for the satisfaction of the Strong Minimalist Thesis (SMT). It is supported by Chomsky’s Phase (Chomsky 1999, 2001), Phase higher evaluation (Chomsky 1999), and the workspace of the MERGE-based system (Chomsky 2019a, 2019b, 2021).

- (7) a. [John works at a school and Mary ~~works~~ at a bank.]
 b. [[_{CP}[_{TP} John [_{VP} John works at a school]]] [and [_{CP}[_{TP} Mary [_{VP} ~~works~~ at a bank]]]].]
- (8) a. [Bill wants to try to begin to write a novel,
 and Jane ~~wants to try to begin to write~~ a play.]
 b. [[_{CP}[_{TP} Bill wants [_{CP}[_{TP} to try [_{CP}[_{TP} to begin [_{CP}[_{TP} to write a novel]]]]]]]]] [and [_{CP}[_{TP} Jane ~~wants~~ [_{CP}[_{TP} ~~to try~~ [_{CP}[_{TP} ~~to begin~~ [_{CP}[_{TP} ~~to write~~ a play]]]]]]]]]].]
- (9) a. [Watasi-wa sakana-o ~~tabe~~ Bill-wa gohan-o tabeta.] (Japanese)
 I-Top fish-ACC eat Bill-Top rice-ACC ate
 ‘I ate fish and Bill rice.’
 b. [[_{CP}[_{TP} Watasi-wa [_{VP} sakana-o ~~tabe~~]]] [[_{CP}[_{TP} Bill-wa [_{VP} gohan-o tabeta]]]].]
- (10) a. [Ja pila vodu i Anna ~~pila~~ vodca.] (Russian)
 I drank water and Anna ~~drank~~ vodca
 a’ [[_{CP}[_{TP} Ja [_{VP} pila vodu]]] [i [_{CP}[_{TP} Anna [_{VP} ~~pila~~ vodca]]]].]
 b. [Ja vodu ~~pila~~ i Anna vodca pila.]
 I water ~~drank~~ and Anna vodca drank
 b’ [[_{CP}[_{TP} Ja [_{VP} vodu ~~pila~~]]] [i [_{CP}[_{TP} Anna [_{VP} vodca pila]]]].]

4. Closing Remarks

References

- Boeckx, C. 2008. *Bare Syntax*. New York: Oxford University Press.
- Chomsky, N. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- Chomsky, N. 1998. Minimalist Inquiries: the Framework. Ms, MIT.
- Chomsky, N. 1999. Derivation by Phase. *MIT Occasional Papers in Linguistics* 18, 1–44.
- Chomsky, N. 2001. Beyond Explanatory Adequacy. Ms, MIT.
- Chomsky, N. 2005. On Phases. Ms, MIT.
- Chomsky, N. 2013. Problems of Projection. *Lingua* 130, 33–49.
- Chomsky, N. 2014. Problems of Projection: Extensions. Ms, MIT.
- Chomsky, N. 2019a. Some Puzzling Foundational Issues: The Reading Program. *Catalan Journal of Linguistics Special Issue*, 263–285.
- Chomsky, N. 2019b. UCLA Lectures. Ms
- Chomsky, N. 2021a. WCCFL Plenary 1: Genuine Explanations.
- Chomsky, N. 2021b. LSJ Lecture: Minimalism: where we are now, and where we are going. Ms.
- Chomsky, N. and H. Lasnik. 1993. The Theory of Principles and Parameters. In Jacobs, J., von Stechow, A., Sternefeld, W. and T. Vennemann (eds). *In Syntax: An International Handbook of Contemporary Research*, 506–569. Berlin: Walter de Gruyter.
- Grohmann, K. 2003. *Prolific Domains. On the Anti-Locality of Movement Dependencies*.

Amsterdam: Benjamins.

Johnson, K. 2006. Gapping. In Everaert M. and H. van Riemsdijk (eds). *The Blackwell Companion to Syntax*. Volume 2, 506–569. MA: Blackwell Publishing.

López L. and S. Winkler. 2003. Variation at the Syntax–semantics Interface. In Schwabe, K. and S. Winkler (eds). *The Interfaces: Deriving and Interpreting Omitted Structures*, 227–248. Amsterdam: John Benjamins Publishing Co.

Ross, R. 1967. *Constraints on Variables in Syntax*. Ph.D. Dissertation. MIT.

Ross, R. 1970. Gapping and the Order of Constituents. In M. Boerwosjcj & K.E. Heidolph (eds.), *Progress in Linguistics*, 249–259. The Hague: Mouton.

중세국어 '곧ㅎ다'류 동등비교구문의 정도성 표현 연구

- '이'형 구문과 '와'형 구문을 중심으로

윤보영

(신라대학교)

1. 서론
2. '곧ㅎ다'류 구문 유형 검토
 - 2.0. 선행 연구 검토 및 제안
 - 2.1. 'NP1-이 NP2-이 곧ㅎ다' 구문
 - 2.2. 'NP1-이 NP2-와 곧ㅎ다' 구문
3. '곧ㅎ다'류 동등비교구문의 정도성 표현
 - 3.1. 부분 일치 구문
 - 3.2. 완전 일치 구문
4. 결론

1. 서론

중세국어 동등비교구문에서 고빈도 출현 양상을 보이는 '곧ㅎ다'류 동등비교구문은 현대국어와는 대비되는 조사 결합상의 다양한 구조적 특징을 보인다. 이에 그 개별 문법 기능 및 의미 기능에 대한 면밀한 검토가 요구되면서 지속적인 연구의 대상이 되어 왔다. 김정아(1993)은 중세국어의 동등비교구문의 구조를 조사를 중심으로 7가지로 유형화하고, 'NP2'에 연결되는 조사의 형태에 따라 '와'형, '이'형, '와로'형, '으로'형, '에'형으로 구분하여 논의를 진행하였다. 이를 통해 다양한 조사와 결합하여 형성된 중세국어 '곧ㅎ다'류 동등비교구문은 각각 독립적 의미 기능을 수행하고 있음을 밝히고 있다. 이는 비교적 다양한 문헌 자료 검토를 통해 중세국어 동등비교구문의 독립된 형태·통사적 특성을 자세히 다루고 있어 중세국어 동등비교구문을 이해하는 데 상당한 기여를 하였다고 할 수 있다.

이러한 지금까지의 연구 성과를 통해 개별 형태들의 통사적, 의미적 기능에 대한 궁금증이 어느 정도 해소되었다면, 이제 원점에서 동등비교라는 공통된 기능에 집중하여 각 개별 유형들의 상호 연관성에 대해 좀 더 깊이 들여다볼 필요가 있다. 이 가운데 특히 주목받아 온 'NP1-이 NP2-이 곧ㅎ다' 구문과 'NP1-이 NP2-와 곧ㅎ다' 구문은 '비유'와 '비교'라는 현대국어의 결과론적 해석 방식에서 벗어나 15~16세기 언중들이 두 가지 구문을 어떠한 기준으로 구분하여 사용하였는지에 대한 심도 있는 검토가 요구된다. 이에 본고에서는 이 두 구문이 나타나는 다양한 문헌 자료의 용례를 해석함에 있어 '동등함의 정도성'에 대한 언중의 인식이 비교구문의 형식적 차이로 표현되고 있음을 포착하고 이에 대한 연구를 진행해 보고자 한다.

송홍규(2016)은 비교구문의 정도차 표현은 비교대상의 특성이나 성질이 얼마나 다른지를 표현하는 기능을 갖기에 차등비교구문에서만 나타날 수 있다고 하였다. 곧 차등비교구문에서는 서술어가 가지는 정도성에 정도부사나 수량화구 등이 결합하여 비교대상 사이의 정도 차이를 나타내며, 동등비교구문은 기본적으로 두 비교 대상이 갖는 정도가 차이가 나지 않음을 표현하는 것이므로 정도차 표현은 쓰일 수 없다고 보았다. 그러나 본고는 비교대상의 ‘다름’뿐만 아니라 ‘같음’에 대한 ‘정도성’ 역시 인간의 인식 가능 영역이며, 이러한 ‘같음의 정도’를 표현하는 언어 형식 역시 존재하였음을 밝히고자 한다.

이에 우선 중세국어 ‘ㄴㅎ다’류 동등비교구문의 통사·의미적 특성을 면밀히 검토하여 각 유형들의 독립적 의미 기능을 확인하고, 이들 사이의 상호 연관성을 밝히려 한다. 또한 ‘NP1-이 NP2-이 ㄴㅎ다’ 구문과 ‘NP1-이 NP2-와 ㄴㅎ다’ 구문의 비교기준(NP2)과 결합하는 조사 선택의 차이는 언중이 ‘같음’에 대한 판정을 언어적으로 표현함에 있어 ‘같음’의 판정에 ‘정도성’을 더하여 드러내고자 문법적 표지를 달리 적용한 결과임을 입증해 보고자 한다.¹⁾ 이러한 논의는 앞선 연구 성과를 이해함에 있어 예외적 현상으로 분류되는 용례들에 대한 구문 해석의 실마리를 찾고자 한 것에서 출발한다. 즉 기존 연구에서 밝힌 ‘이’형 구문과 ‘와’형 구문이 갖는 ‘비유’와 ‘비교’라는 의미 기능적 특성은 두 구문이 갖는 단편적 속성에 불과함을 인식하고 이를 보다 폭넓게 해석할 수 있는 새로운 대안을 모색해 보고자 하는 것이다.

2. ‘ㄴㅎ다’류 구문 유형 검토

2.0. 선행 연구 검토 및 제안

먼저 중세국어 ‘ㄴㅎ다’류 구문이 어떻게 나타나는지 그 구문 유형을 살펴보기로 한다. 비교구문에 나타나는 통사 구조의 다양성은 비교와 관련된 제반 요소들을 비교 행위자가 어떤 측면에서 어떻게 인지하느냐에 따라 나타나는 현상이다. 따라서 이러한 인지 작용에 관여하는 요소들이 비교구문 속에서 어떻게 실현되는가를 검토함으로써 중세국어 비교구문의 성격을 이해할 수 있을 것이다. 특히 조사 선택, 각 명사구의 성격 및 후행 조사와의 관계, 각 유형의 비교구문에 관련되는 주요 형태들의 특성 등에 대한 세밀한 검토가 각 비교구문의 통사·의미적 특성을 파악함에 있어 중요한 관건이 될 것이다. 이 가운데 15세기 ‘ㄴㅎ다’ 구문에서 가장 고빈도로 나타나는 ‘와’형 구문과 현대국어에서는 관찰되지 않는 유형인 ‘이’형 구문에 대한 부분은 더욱 철저한 검토가 필요하다.²⁾

김정아(1998)은 ‘와’형 구문과 ‘이’형 구문은 모두 화자가 비교대상에 대해 유표적 태도를 취하는 구문으로서, 두 대상 중 한쪽에 대해 특별한 태도를 취하는 공통점을 가지고 있으나 ‘와’형 구문에서 ‘와’ 명사구의 개체는 단순히 비교되는 대상물로서 개체의 외연적 의미로서 사용되는 반면, ‘이’형 구문에서 ‘이’ 명사구(NP2)는 비교주체의 성질을 설명하기 위한 보조적 존재로서 그 자신의 존재가 아니라 그것이 갖고 있는 성질 즉 개체의 내포

1) 기술상의 편의를 위해 구문 구조에 대한 상세한 기술이 요구되지 않는 경우는 비교기준에 결합된 조사를 그 특징으로 삼아 ‘이’형 구문, ‘와’형 구문 등과 같은 형태로 구문 유형을 제시한다.

2) 본고에서는 비교구문을 구성하는 주된 요소 가운데, 비교 행위자(평가자)가 비교의 중심적 자격을 부여하는 NP1을 ‘비교대상’으로, 평가 행위의 잣대가 되는 NP2를 ‘비교기준’으로 규정하고 분석 내용을 기술해 나간다.

적 의미가 사용된 것으로 파악하였다.³⁾ 또한 이러한 입장에서 '와'형 구문이 단순 비교를, '이'형 구문이 비유를 위한 형식으로 사용되었음을 이해할 수 있다고 하였다. 즉 서로 다른 개체를 비교대상으로 하여 그들의 어떤 한 성질이 같음을 보이는 '와'형 구문을 일반적인 비교구문으로 보고, 이와는 달리 서로 구체적인 연관성이 없는 명제를 연결하여 그들의 관련성을 찾아내는 '이'형 구문을 비교의 하위 개념인 비유구문으로 구분한 것이다. 그러나 이러한 논의 방식에 따른 유형적 특성에 근거하여 중세국어 용례를 재검토함에 있어 쉽게 설명되지 않는 다음과 같은 용례가 확인되어 이에 대한 정밀한 검토가 요구된다.

(1) ㄱ. 各別히 法愛를 내야 브토틐 플와 옷과 같하야 <楞嚴9:100a>

ㄴ. 뉘음과 글음과 因호미 같하야 聖과 凡과 두 길히 업스니 <楞嚴5:11a>

(1 ㄱ)은 '各別히 法愛를 내야 브툼'을 NP1으로 '플와 옷'을 NP2로 가지는 '와'형 비교구문이다. 그러나 이 문장의 앞뒤 문맥을 고려할 때 이는 단순 비교구문으로 보기 어렵다. 즉 '사람으로 변하여 나타난 마귀를 스승인 본래의 선지식이라 알아 (중생이) 각별히 법애를 내어 붙음'이 '플와 옷' 같음을 비유적으로 표현한 것으로 이해할 때 자연스러운 해석이 가능하다.⁴⁾

한편, (1 ㄴ)은 '뉘음과 글음'을 NP1으로 '인호미'를 NP2로 가지는 '이'형 비교구문이다. 이 역시 김정아(1998)의 유형 분류에 따라 비유구문으로 해석하는 것은 적절하지 않으며, '매는 것'과 '푸는 것'은 '인'이라는 측면에서 동일함을 나타내고자 한 단순 비교구문의 형태로 파악하는 것이 옳을 듯하다.

이상에서 우리는 '와'형 비교구문과 '이'형 비교구문을 비교 행위의 측면에서 '비교'와 '비유'라는 이분법적 사고로 바라볼 수 없음을 확인하였다. 그렇다면 위와 같은 언어적 형식을 취하게 된 근원적 이유는 무엇일까? 우리는 이에 대한 답을 구하기 위해 '같다'류 동등 비교구문 가운데 '이'형 구문과 '와'형 구문이 독립적 의미 기능을 획득하게 되는 과정을 다시 따져 보아야 할 것이다.

'같다'가 나타나는 구문에 대한 검토에 앞서 '같다'의 개별 의미 특성을 먼저 파악해 보기로 하자. 중세국어 평가형용사 '같다'는 현대국어에서처럼 '서로 다르지 않고 하나이다'라는 의미뿐만 아니라 '똑같지는 아니하지만 전체적 또는 부분적으로 일치하는 점이 많다'라는 의미를 함께 가졌음을 앞선 연구를 통해 확인한 바 있다.⁵⁾ 이는 [동일함]의 의미와 [유사함]의 의미로 달리 표현할 수 있으며, 이를 통해 우리는 중세국어 '같다'가 현대국어의 '비슷하다'의 의미까지도 포괄하고 있었음을 알 수 있다. 이러한 '같다'가 갖는 의미적 특성은 구문에 따라 그 의미가 더욱 명확하게 구별되어 나타난다.

3) 김정아(1998)는 비교구문의 NP1과 NP2 전체를 비교대상으로 규정하고, NP1을 비교주체, NP2를 비교기준으로 구분하여 논의를 진행하였다.

4) 해석에 대한 타당성 검토의 편의를 위해 해당 구문이 나타나는 전체 문맥을 함께 제시한다.

(i) 그 ㅁ 天魔ㅣ 엇와 그 ㅁ을 得하야 精을 놀여 사르믈 브터 이베 經法을 니르게 하야든 그 사르미 本來 魔ㅣ 著호 ㄷ 아디 ㅁ하야 ㅁ 닐오디 내 無上涅槃을 得호라 호고 더 應求호는 善男子處에 와 座 ㅁ고 說法호디 能히 듣는 衆으로 ㅁ간 그 모미 百千歲 ㅁ호믈 보고 ㅁ스매 愛染을 내야 能히 브려 여희디 ㅁ하야 모미 ㅁ이 ㅁ외야 四事로 供養호디 잇부믈 아디 ㅁ호며 各各 座下엿 사르미 ㅁ스ㅁ로 이 ㅁ것 스승인 本來人 善知識이라 아라 各別히 法愛를 내야 브토틐 플와 옷과 같하야 아리 잇디 아니호믈 得호라 호리니 이 사르미 어리 迷惑하야 惑하야 菩薩이라 하야 그 ㅁ스매 親 近히 하야 부텃 律儀를 ㅁ오 <楞嚴9:100a>

5) 김정아(1998)은 '같다'가 '同, 等, 如, 均, 猶, 若, 似, 像' 등의 다양한 한자와 대응함을 검토하고, 이를 통해 중세국어 '같다'의 의미적 특징을 제시한 바 있다.

중세국어 ‘같다’류 구문 가운데 ‘이’형 구문과 ‘와’형 구문의 의미 기능적 특성은 먼저 비교대상 (NP1)과 비교기준(NP2)의 관계적 특성 파악을 통해 미루어 짐작해 볼 수 있으며, 비교대상과 비교기준의 관계성은 후행문에 나타나는 비교 내용을 통해 보다 명확하게 설명될 수 있다. 이에 중세국어 ‘같다’가 보이는 구문 유형을 살펴, 비교 내용이 구체적으로 드러나는 용례를 중심으로 살펴 나가기로 한다.

2.1. ‘NP1-이 NP2-이 같다’ 구문

먼저 현대국어에 나타나지 않는 ‘NP1-이 NP2-이 같다’ 구문을 살펴보자.

- (2) 一와 異왜 쁘디 ㄹ죽흐며6) 解와 惑과 根源이 같하야 一은 眞 비취요물 니르시고 異는 俗 비취요물 니르시니 眞俗이 비록 다르나 妙旨는 다르디 아니흐니 解惑도 例 그러흐니라 <永嘉下48b>

(2)는 ‘一와 異 (NP1)’는 ‘쁘 (NP2)’이 같으며 ‘解와 惑 (NP1)’은 ‘根源 (NP2)’이 같아서… (一와 異)는 眞俗이 비록 다르나 妙旨 (묘한 뜻)는 같고 解惑도 (역시) 그러함을 나타낸 구문으로 해석된다. 이는 1차적으로 복수의 비교대상이 비교기준에 비추어 서로 일치함을 나타내고자 한 것으로 해석될 수 있다. 이때 복수의 비교대상을 평가하기 위해 제시한 비교기준은 각각의 비교대상의 의미 영역에 공통적으로 포함된 부분적인 존재이며, 이러한 내재된 부분적 비교기준에 한하여 두 비교대상이 서로 동등함을 드러내고자 한 단순 비교구문으로 볼 때 그 의미 해석이 자연스럽다 할 것이다.⁷⁾ 이에 (2)는 ‘비교대상1의 부분적 속성(a)과 비교대상2의 부분적 속성(a’)이 서로 일치함’을 나타내려는 구문이며, 부분적 속성이 비교기준 ‘NP-이’의 형태로 나타난 구문 구조로 파악해 볼 수 있다. 다시 말해 (2)는 비교대상과 비교기준이 전체와 부분의 관계를 지니는 이중주어문으로 파악할 수 있다. 다음의 (3) 역시 단순 비교로서의 쓰임을 보이는 ‘이’형 구문의 용례에 해당한다.

- (3) ㄱ. 十方 國에 各各 부터 드외야 다 訖 號 | 같하야 일후미 寶相 如來 應供正遍 知明 行足善逝 世間海無上土 調御丈夫 天人師佛世尊이리니 <法華4:64a>
 ㄴ. 根과 塵과 根原이 같흐며 悤음과 버숨과 둘 업스며 <圓覺下1-2:45a>

6) (2)’ 一와 異왜 쁘디 ㄹ죽흐며 <永嘉下48a>

중세국어 ‘ㄹ죽흐다’는 한자 ‘齊’와 의미적 대응 관계를 보인다. 이에 (2)’의 ‘ㄹ죽흐며’는 ‘가 지런하다, 같다, 바르다’ 등의 다의적 의미로 해석될 수 있으나, 문맥을 고려할 때 위 용례에서는 ‘같다’의 의미로 해석함이 가장 적절할 듯하다. 이에 ‘같다’비교구문과 함께 검토해 보기로 한다.

7) 이현희 (1999)는 중세국어 [양호함]의 의미를 가지는 ‘동-’ 구문이 ‘NP1-이 NP2-이 동-’ 유형으로 나타남을 검토한 바 있다. 이때 NP1은 NP2를 포함하는 의미영역을 가지며, 이는 바꾸어 NP2가 NP1의 부분적인 존재나 속성 등에 해당하는 것으로도 설명할 수 있다 하였다. 또한 이 구문은 [NP1-이 [NP2-이 동-]]의 구조를 가지기 때문에 관계 구성으로 나타날 때 NP1은 핵심명사가 될 수 있으나 NP2는 핵심명사가 되지 못한다고 하였다. 이러한 중세국어 평가형용사 ‘동-’ 구문에 대한 해석은 ‘NP1-이 NP2-이 같다’구문의 의미 기능적 특성을 이해하는 것에 도움이 될 것이다.

(3ㄱ)은 비교대상(NP1)이 밑줄 친 비교구문의 문면에 드러나지 않으나 선행문의 내용에 비추어 '各各 부터'로 상정해 볼 수 있다. 즉 부처 각각의 동등함은 '號'라는 한정된 기준에 대한 것임을 나타낸 것으로 이해할 수 있다. (3ㄴ) 역시 '根과 塵'의 시초(根原)에 대한 부분적 일치성을 표현한 단순 비교구문으로 파악할 수 있다.⁸⁾ 이렇듯 중세국어의 '곧ㅎ다'류 '이'형 비교구문에서도 비유적 진술이 아닌 단순 비교구문과 같은 쓰임을 일부 확인할 수 있다.⁹⁾ 이러한 통사 구조를 지닌 중세국어의 비교구문은 비교대상에 내포된 일부분을 비교기준으로 취하고 있으므로 비교대상과 비교기준의 의미적 상관성을 유추하기 위한 비교기준이 가진 속성에 대한 구체적 진술이 필수적으로 요구되지 않는다. 따라서 후행문에 비교기준이 가진 속성에 대한 진술이 뒤따르지 않음은 지극히 당연한 현상으로 받아들일 수 있다.

한편, 다음의 용례는 조금 다른 양상을 보인다.

(4) 四大와 五蘊과 거우루엿 像이 곧ㅎ야 뷔며 뷔여나 업스며 쏘 사르미 업도다 <金三 3:31a>

(4)는 복수의 비교대상을 비교기준에 근거하여 평가하는 형식을 취함은 (3)과 동일하나, 비교기준이 비교대상의 의미 영역에 포함된 일부가 아니라는 점에서 (3)과의 차이점을 확인할 수 있다. (4)의 비교기준인 '거우루엿 像'은 복수의 비교대상의 속성을 설명하기 위한 보조적 장치로 활용된 것이며, 이 속성은 비교기준이 가진 여러 가지 속성 가운데 일부분에 해당한다. 따라서 (4)와 같은 구문은 비교대상에 부합하는 비교기준의 한정적 속성에 대한 구체적 내용이 후행문으로 뒤따라 비교 내용을 반드시 보강해 주어야 의미가 명확하게 전달될 수 있다. 즉 (4)는 서로 구체적인 연관성이 없는 명제를 연결하여 관련성을 찾아내는 비유적 진술 방식으로 이해할 수 있다. 이러한 '이'형 비교구문이 보이는 비유적 진술 방식은 다음의 용례를 통해서도 파악해 볼 수 있다.

8) i. 法身은 量이 虛空과 곧ㅎ실씨 이럴씨 <金剛30a>

위의 용례는 (2~3)과 의미적으로 유사한 진술 방식을 취하고 있으나 구조적으로 다른 구문에 해당한다. 통사 구조를 분석하는 방식에 따라 (i)은 [[虛空과 法身]은 量이 곧ㅎ실씨]와 같이 '이'형 비교구문의 구조가 어순재배치에 의해 변형된 것으로 볼 여지가 있다. 그러나 (i)의 '法身은 量이'를 '法身이 量이'로 표현하여도 전체 문맥 해석에 문제가 발생하지 않는다. 즉 주 제어 '法身'은 '量'과 전체와 부분의 관계(法身の 量)를 가진다는 점에서 이를 하나의 문장 내에 존재하는 성분주제어로 보아 '法身은 量이'를 이중주어로 처리할 수 있으며, '虛空과'는 부사어로서의 자격을 갖는 것으로 보는 것이 좋을 듯하다. 이에 (i)은 어순재배치에 의한 '이'형 비교구문이 아닌 '[[법신의 양]이 [허공(의 양)]과 같아]'라는 '와'형 완전 일치 비교구문으로 보고자 한다.

9) 두 비교대상의 이질성을 드러내는 다음의 용례에서 'NP1-이 NP2-이 다르다' 구문 유형을 확인할 수 있다.

i. 心和 形과 用이 달오되 <楞嚴10:81b>

ii. 溫陵環師 | 一生을 門 담고 世間과 즐규미 달아 호오사 黃卷 聖賢을 피스와 <楞嚴1:3a>

(i)은 복수주어 '心'과 '形'이 '用'의 측면에서 같지 않음을, (ii)는 '溫陵環師'와 '世間(세상 사람)'의 '즐김(좋아함)'이 같지 않음을 표현한 것이다. 이때 이질성을 판단하는 비교기준 '用'과 '즐김'은 복수주어에 속한 특성의 일부에 해당한다. 이렇듯 비교대상과 비교기준이 전체와 부분의 관계를 지닐 때 'NP1-이 NP2-이 곧ㅎ다/다르다'의 구문 형태를 취하는 것은 중세국어에서 보편적인 현상이었음을 짐작해 볼 수 있다.

- (5) 가. 道 | 큰 바를 | 곧 하야 더욱 드디웃 더욱 기프리라 <蒙山高운사37b>
 나. 眞機 | 銀山鐵壁 | 곧 하야 구더 들디 어려우며 눕고 머리 더위잡디 못호고 <金三序二5a>
 다. 罪性 | 서리와 눈패 | 곧 하야 智慧人 히 곳 올오매 혼 點도 업는들 모로매 아롤디니라
 <南明上8b>
 라. 奧 | 듣는 들헛 | 블와 번겿 | 비치 | 곧 하야 블가 어루 보리로디 <金三序二5a>
 마. 모든 罪 | 는 서리와 이슬왜 | 곧 하야 智慧人 히 能히 스러 더는니라 험시니라 <南明上9a>

(5가)은 ‘道’와 ‘큰 바를’을, (5나)은 ‘眞機’와 ‘銀山鐵壁’을 각각의 비교대상과 비교기준으로 취하고 있다. 이들 각각의 개체는 구체적 상관성이 직접적으로 드러나지 않으므로, 비교 행위자(평가자)는 이러한 비교구문의 의미 파악 과정에서 발생할 수 있는 다의적 해석의 가능성을 차단하기 위해 비교기준의 부분적 속성에 대한 구체적 설명을 후행문에 제시하게 된다. (5다~마)은 하나의 비교대상과 복수의 비교기준이 공통적으로 지닌 부분적 속성에 대한 동등함을 표현한 구문 구조로 이 역시 비교 행위자가 동등하다고 판단하게 된 비교기준의 부분적 속성에 대한 내용이 뒤따르고 있음을 확인할 수 있다.¹⁰⁾

이상에서 중세국어 ‘NP1-이 NP2-이 곧 하다’ 구문이 단순 비교구문과 비유구문으로 모두 문헌에 나타남을 확인하였다. 이는 ‘NP1-이 NP2-이 곧 하다’ 구문을 모두 비유구문으로 일반화하여 설명할 수 없음을 증명하는 근거가 된다. 그러나 단순 비교구문(2~3)과 비유구문(4~5)의 용례 모두 비교대상을 비교기준에 근거하여 평가하되, 전체가 아닌 부분적 속성에 기반한 상호 일치성을 드러내고자 했다는 점에서 공통점을 찾을 수 있으며, 단순 비교구문과 비유구문이 동일한 ‘이’형 구문으로 나타날 수 있었던 본질적 이유 역시 이러한 공통점에 기반하여 다시 유추해 낼 수 있다 생각된다. 나아가 ‘NP1-이 NP2-이 곧 하다’ 구문을 ‘부분적(조건적) 동등비교구문’으로 유형화하여 이들을 아울러 살펴보는 것이 중세국어 비교구문의 문법적 체계를 마련함에 있어 더욱 경제적인 방식이 될 것이라 여겨진다.¹¹⁾

2.2. ‘NP1-이 NP2-와 곧 하다’ 구문

앞서 언급하였듯이 기존 연구에서 중세국어 ‘와’형 비교구문은 단순 비교라는 용법을 가짐을 검토한 바 있으며, 다음의 용례를 통해 이를 확인할 수 있다.

- (6) 가. 이 踏七寶華佛國土 莊嚴과 목숨과 劫數와 教化 | 흔 弟子와 正法 像法이 또 山
海慧自在通王如來와 곧 하야 달오미 업스리며 또 이 부텃기 몬아드리 드외리니
 <月釋15:36b>
 나. 여러 모땃 微塵이 世界 | 딩ᄃᆞᆫ | 는 微塵과 곧 하야 | 곧 하야 | 곧 하야 | 곧 하야 | 곧 하야
 고티요미 업서 微塵과 自性 이 서르 다티 아니하며 刀兵에 니르러도 또 다호미 업슨들 보아 <楞嚴5:68b>

10) (5라~마)의 용례는 구문이 나타나는 자료의 전체적 문맥을 고려하였을 때 비교대상(NP1)이 주제화되어 ‘NP1-은 NP2-이 곧 하다’의 형식을 취한 것으로, 이 역시 ‘NP1-이 NP2-이 곧 하다’ 구문과 동일한 맥락에서 이해될 수 있다.

11) 동등비교구문의 정도성 표현에 대한 내용은 제3장에서 자세히 다루기로 한다.

(6ㄱ)은 표면적으로는 비교대상 '蹈七寶華佛國土 莊嚴과 목숨과 劫數와 教化혼 弟子와 正法 像法'이 비교기준 '山海慧自在通王如來'와 일치함을 나타내는 비교구문으로 보이나, 앞 뒤 문맥을 고려하면 비교기준 NP2는 '國土 莊嚴과 목숨과 劫數와 教化혼 弟子와 正法 像法'이 생략된 것으로 보아야 한다.¹²⁾ 이에 (6ㄱ)은 기본적으로 '蹈七寶華佛의 國土 莊嚴과 목숨과 劫數와 教化혼 弟子와 正法 像法'과 '山海慧自在通王如來의 國土 莊嚴과 목숨과 劫數와 教化혼 弟子와 正法 像法'의 동일함을 드러내고자 한 구문이며, 이때의 동일함은 문맥적으로 1 : 1 대응 관계를 지닌 각 개체 상호간의 전면적 일치를 의미하는 것이라 추측된다. (6ㄴ)은 '몸의 微塵이 세계 만든 微塵과 차별이 없이 같아 微塵과 自性이 서로 부딪히지 않으며 刀兵에 이르러도 (刀兵까지도) 또한 부딪힘이 없음을 보았음'을 나타낸 것으로 직역된다. 이때 비교대상(여러 모땀 微塵)과 비교기준(世界 밍그랏는 微塵)은 동일한 의미영역(개체를 이루는 요소)을 지녀 구체적 연관성을 쉽게 파악할 수 있으며, 이러한 의미 구조를 가지는 '와'형 비교구문을 단순 비교구문으로 해석함에는 어려움이 따르지 않는다.

그러나 '와'형 비교구문 가운데 비유적 진술 방식의 특성에 보다 근접한 용례를 확인할 수 있어 이에 대한 검토가 요구된다.

- (7) ㄱ. 부텃 짜히 雜말 업시 淸淨호고 겨지비 업스며 惡趣ㅣ며 受苦르힐 소리 업고
瑠璃 짜히 드외오 金노호로 길힐 느리고 城이며 지비며 羅網이 다 七寶로 이
러 이쇼미 또 西方 極樂 世界와 곧하야 글히요미 업고 <釋詳9:11a>
ㄴ. 너브신 복이 하늘와 ㅁ트샤 브름도 고르며 비도 순하야 <翻朴1a>

비교의 하위 개념으로서 비유는 어떤 사물이나 현상을 비슷한 현상이나 사물에 빗대어 우회적으로 표현하는 것으로, 비교대상과 비교기준은 표면적으로 직접적인 연결 고리를 형성하지 않는 것이 특징이다. 이러한 비교대상과 비교기준과의 관련성만을 고려했을 때 (7)의 용례는 비유구문으로 해석될 여지가 있다. 그러나 '이'형 구문의 비유적 진술 방식과는 달리 '와'형 구문은 비교대상의 범위를 제한하는 구체적 진술을 통해 한정된 비교대상이 비교기준과 전면적으로 일치됨을 표현하고 있다. 즉 (7ㄱ)은 구체적 특징으로 한정된 '부텃 짜히'와 '西方 極樂 世界'의 특징이 전면적으로 일치함을 드러내고 있으며, (7ㄴ)은 '넓음'으로 제한된 '복'의 특성과 '하늘(의 넓음)'이 완전히 일치함을 드러낸 것으로 해석할 수 있다. 이에 우리는 비유적 진술 방식에 있어 '이'형 구문이 비교대상과 비교기준의 부분적 속성에 대한 동등성을 드러낸다면, '와'형 구문은 한정된 비교대상과 비교기준 사이의 전면적 동등성을 표현하는 것으로 그 의미 기능을 유추해 볼 수 있다. 이러한 '와'형 동등비교구문의 의미 기능적 특징은 선행문과 후행문의 의미적 관련성을 앞서 살핀 '이'형 구문과 비교하여 검토해 봄으로써 보다 명확하게 파악해 볼 수 있다. 다음 두 용례를 살펴보자.¹³⁾

12) 문맥 해석의 타당성 검토의 편의를 위해 관련 내용을 부분 발췌하여 제시한다.

(i) 그뻘 부테 羅睺羅드려 니르샤티 네 오는 누에 부테 드외야 號를 蹈七寶華如來 應供 正遍知 明行 足 善逝世間解 無上士 調御丈夫 天人師佛 世尊이리니 十世界 微塵等數 諸佛如來를 供養하야 상네 諸佛入 거미 몰아드리 드외야 오닐날 곧하리라 이 蹈七寶華佛國土 莊嚴과 목숨과 劫數와 教化혼 弟子와 正法 像法이 또 山海慧自在通王如來와 곧하야 달오미 업스리며 또 이 부텃기 몰아드리 드외리니 <月釋15:36b>

- (8) ㄱ. 道 | 큰 바리리 곧 하야 더욱 드디웃 더욱 기프리라 <蒙山49a>
 ㄴ. 이 踏七寶華佛國土 莊嚴과 목숨과 劫數와 教化혼 弟子와 正法 像法이 쏘 山
海慧自在通王如來와 곧 하야 달오미 업스리며 쏘 이 부텃기 몰아드리 드외리니
 <月釋15:36b>

(8ㄱ)과 (8ㄴ)의 선행문과 후행문의 1차적 의미는 ‘곧 하야’에 의한 인과적 관계로 해석될 수 있다. 그러나 이들 구문의 선행문과 후행문의 비교 행위에 기반한 원인과 결과의 구체적 관계는 달리 포착된다. 먼저 (8ㄱ)은 구체적 연관성이 없는 개체를 서로 비교대상(道)과 비교기준(큰 바리)으로 삼아 그들이 가진 동일성에 대해 설명하고 있다. 이에 (8ㄴ)은 선행문의 동일성에서 비롯된 후행문의 (비교대상이 지닌) 결과적 [상태성]의 표현으로 그 전체 의미를 파악할 수 있다. 이에 반해 (8ㄴ)은 비교대상(踏七寶華佛國土 莊嚴과 목숨과 劫數와 教化혼 弟子와 正法 像法)과 동일한 의미 영역을 지니며 구체적 연관성이 있는 개체를 비교기준(在通王如來의 國土 莊嚴과 목숨과 劫數와 教化혼 弟子와 正法 像法)으로 삼고 있어 둘의 일치성을 설명하기 위한 비교기준이 가진 특성이 뒤따르지 않아도 그 의미를 이해하는 것에 문제가 발생하지 않는다. 따라서 (8ㄴ)의 후행문은 선행문의 동일성에서 비롯된 비교대상의 결과적 [행위성]에 대한 내용으로 보아 구문 전체가 지닌 의미를 해석함이 보다 적절하다 여겨진다. 이러한 비교대상과 비교기준의 의미적 연관성과 선행문과 후행문 사이의 인과적 관련성을 중심으로 (8ㄴ)과 유사한 형식을 보이는 다음의 용례들을 검토해 보자.

- (9) ㄱ. 모든 幢王刹에 흘러 브서 거우룻 소매와 드러 내 모매 드러 모미 虛空과 흐
야 서르 마가 ㄹ료미 업스며 모미 能히 微塵國土에 이대 드러 佛事를 너비 行
 하야 큰 隨順을 得호니 <楞嚴5:79a>
 ㄱ'. 諸幢王刹이 來入鏡內涉入我身하야는 身同虛空하야 不相放礙하며 身能善入
 微塵國土하야 廣行佛事호대 得大隨順하니
 ㄴ. 두 사르미 서르 스랑하야 돌히 스랑하야 念호미 기프면 이 ㄹ티 生을 브터 生
 에 니르리 얼굴와 그림자와 곧 하야 서르 어괴여 다르디 아니호리니 <楞嚴:85b>
 ㄴ'. 二憶念深하면 如是乃至從生至生히 同於形影 不相乘異하리라

(9ㄱ~ㄴ)은 모두 ‘NP1-이 NP2-와 곧 하야’ 구문을 포함한 선행문과 선행문의 비교 행위에서 비롯된 결과적 내용이 후행하여 뒤따르는 구조를 지니고 있으나, (9ㄱ)은 ‘몸’과 ‘허공’을, (9ㄴ)은 ‘두 사람’과 ‘얼굴(형체)과 그림자’를 각각 비교대상과 비교기준으로 각각 취하고 있어, 표면적으로 드러나는 비교대상 비교기준과의 관계가 (8ㄴ)과는 다르다. 그러나 (9)는 비교대상의 범위를 한정하는 조건적 내용이 앞에 구체적으로 제시되어 있으며, 이러한 조건은 직접적 의미 연관성이 없는 비교대상과 비교기준과의 동일성을 유추해 내는 사고 과정에 개입하여 의미 파악을 돕는 보조적 역할을 하게 된다. 이에 이들 구문 역시 제한된 비교대상과 비교기준의 전면적 일치를 드러낸 것으로 볼 수 있으며, 후행하는 비교 행위의 결과 역시 이러한 동일함에서 비롯된 비교대상의 행위적 의미로 해석해 볼 수 있을 것이다.¹⁴⁾

13) 설명의 편의를 위해 앞의 용례 일부를 다시 인용한다.

3. '곧하다'류 동등비교구문의 정도성 표현

앞서 우리는 '곧하다'류 구문의 문맥 의미를 세밀하게 검토하여 중세국어에 특수하게 나타나는 'NP1-이 NP2-이 곧하다' 구문이 비교대상 (NP1)과 비교기준(NP2)의 부분적 일치성을 나타내는 언어적 형식임을 확인하였으며, 'NP1-이 NP2-와 곧하다' 구문은 비교대상과 비교기준과의 전면적 일치성을 드러내는 언어적 표현임을 부분적으로 파악해 보았다. 이는 당대 언중들이 동등성에 대한 정도를 언어적 형태로 표현하고자 하였음을 입증할 수 있는 하나의 근거가 된다. 이제 우리는 이러한 의미 기능적 특성이 '곧하다' 비교구문 전반에 두루 적용 가능한지를 검토해 보아야 한다.

3.1. 부분 일치 구문

먼저 기존 연구의 관점에 따른 구문 해석 방식에 문제가 있음을 보이기 위해 앞에서 인용한 용례를 다시 검토해 보기로 하자.

- (10) ㄱ. 먹음과 글음과 因호미 곧하야 聖과 凡과 두 길히 업스니 <楞嚴5:11a>
 ㄱ'. 結解同所因이오 聖凡無二路라

(10 ㄱ)은 복수의 비교대상과 비교기준 사이에 표면적으로 드러나는 구체적 연관성이 포착되지 않으므로 비유구문으로 해석될 여지가 있으나, 구문에 나타나는 비교기준 '因'은 비교대상인 '먹음'과 '글음'이라는 행위가 지니는 근원적 원인에 해당하므로 이는 두 비교대상이 공통적으로 함의한 부분적 속성의 일치를 표현한 것이라 풀어내는 것이 더욱 적합하리라 여겨진다. 또한 비교대상을 평가하기 위한 비교기준은 두 비교대상 모두에 내포되어 있는 공통된 부분적 특성에 해당하므로 비교기준에 대한 구체적 진술이 필수적으로 요구되지는 않는다는 점에서 단순 비교구문의 특성에 보다 부합됨을 보인다. 이러한 '이'형 구문의 단순 비교적 용법은 복수의 비교대상에 공통적으로 포함된 부분적 속성을 비교기준으로 취하는 용례에서 주로 확인할 수 있다.

- (11) ㄱ. 鷗는 산 類를 먹고 土梟는 서르 먹느니 殺食이 又고 鷲는 툇기 잘 잡고 鷲는
 주거를 먹느니 盜食이 又고 <法華2:108b>
 ㄴ. 菩薩이 누니 넙고 큰 靑蓮華葉이 곧하야 <月釋18:77b>
 ㄷ. 一切相을 노기샤미 바리리 흐르는 물 드료미 곧하실씨 넙고 크고 깊고 머
 다 하시니라 <月釋11:99a>

(11 ㄱ)은 '鷗와 土梟'는 '殺食'이라는 측면에서, '鷲와 鷲'는 '盜食'이라는 측면에서 서로 일치함을 나타내는 단순 비교구문이다. 또한 (11 ㄴ)과 (11 ㄷ)은 각각 '(보살의) 눈'과 '一切相을 노기삼'을 '靑蓮華葉'과 '바리리 흐르는 물 드림'에 빗대어 상호간의

14) (9 ㄱ)과 (9 ㄴ) 구문의 비교기준과의 동일함에서 비롯된 비교대상이 보이는 결과적 행위는 각각 '서로 막아 보이지 않게 됨이 없으며 몸이 능히 微塵國土에 들어가 佛事를 넓게 행하여 큰 隨順을 얻음'과 '서로 어긋나지 아니함'으로 유추해 볼 수 있다.

일치함을 비유적 방식으로 설명하나, 이 역시 비교기준이 지닌 부분적 속성과의 동일성을 드러내고 있다는 점에서 (1 1 ㄱ) 과 동일한 유형적 속성을 지닌 것으로 분류될 수 있다. 따라서 중세국어 ‘이’형 비교구문은 그것이 단순 비교적 성격을 보이든 비유적 성격을 보이든, 모두 비교기준에 기반한 부분적 동등함을 나타내고자 한 것으로 의미 기능적 성격을 일반화할 수 있다.

3.2. 완전 일치 구문

다음으로 ‘와’형 비교구문에 나타나는 동등성의 정도 표현을 살펴보자.

아래 (1 2 ㄱ) 은 구체적 연관성이 없는 두 개체를 비교대상 (各別히 法愛를 내야 브툼) 과 비교기준 (플와 옷) 으로 삼아 그들이 가진 동일성에 대해 설명한 구문이다.

- (12) ㄱ. 各別히 法愛를 내야 브툼 플와 옷과 곤흐야 아리 잇디 아니호물 得호라
 흐리니 이 사르미 어리 迷惑흐야 惑흐야 菩薩이라 흐야 그 므스매 親近히
 흐야 부텃 律儀를 헐오 <楞嚴9:100a>
 ㄱ'. 別生法愛하고 粘如膠漆하야 得未曾有라하니 是人愚迷하야 惑爲菩薩하며
 綿愛其心하야 破佛律儀하고 潛行貪慾하리라

이때의 비교기준 ‘플와 옷’은 (1 2 ㄱ') 의 ‘膠漆 (아교와 옷) ’에 대응하는 것으로, 중세국어 시기 ‘膠漆’은 ‘아교와 옷’이라는 기본 의미와 이들이 가진 속성에서 비롯된 ‘서로 떨어질 수 없을 정도로 사이가 아주 친밀함’이라는 내포적 의미를 함께 지닌다.¹⁵⁾ 이에 (1 2 ㄱ) 이 비유적 의미를 드러내고 있음에도 불구하고 ‘와’형 구문 구조를 취하게 된 근본적 이유는 ‘膠漆’ 자체의 부분적 속성과의 일치성이 아닌 내포적 의미와의 전면적 일치성을 강조하기 위한 비교 행위자의 (평가자) 인식이 통사 구조에 투영된 결과일 것이라 짐작해 볼 수 있다. 즉 (1 2 ㄱ) 은 언중이 비교대상과 비교기준의 완전한 일치성을 인식한 후 이를 언어적 형식으로 나타내려고 한 결과적 양상으로 파악해 볼 수 있는 것이다. 이러한 접근 방식은 (1 2 ㄱ) 과 같은 유형의 비교구문을 일반적인 ‘와’형 비교구문이 보이는 주된 특징에 어긋나는 예외적 현상으로 두어야 하는 문제를 해결해 주고, 중세국어 비교구문이 지니는 언어적 현상을 보다 효율적이고 경제적으로 설명하는 것에 도움을 줄 수 있다. ‘와’형 비교구문의 [전면적 동등성] 이라는 의미 기능은 다음의 용례를 통해서도 설명될 수 있다.

- (13) 일후미 等流果 | 니 果 | 因과 곤흐 전치라 <月釋11:49a>

(1 3) 은 이름이 ‘等流果’인 연유가 ‘果’와 ‘因’이 같음에서 비롯된 것임을 설명하고 있다. 불교에서 ‘等流果’는 원인에서 결과가 생길 때, 그 원인과 똑같은 성질이 있는 결과를 이르는 것으로, 이러한 의미에 근거할 때 (1 3) 의 비교 행위자는 ‘果’와 ‘因’의 완전한 일치성을 드러내기 위해 ‘와’형 구문을 선택한 것이라 추측된다. 이 외에도 ‘와’형 비교구문

15) ‘膠漆’의 기본 의미와 내포 의미로서의 쓰임은 다음을 통해 확인할 수 있다.

ㄱ. 시혹 뵈에 膠漆호모로 [膠는 프리오 漆은 오시라] <法華1:219a>

ㄴ. 天下엿 벌들흐 다 膠漆마타 사괴더니라 <杜詩3:62a>

을 '비교대상과 비교기준 사이의 완전한 일치성에 대한 언중의 인식을 반영한 것'으로 특징 지을 수 있는 근거는 다음의 용례에서도 일부 확인할 수 있다.

- (14) ㄱ. 여러 모뎨 微塵이 世界 甞ᄃᆞᆫ 微塵과 곧ᄃᆞᆫ 곧ᄃᆞᆫ미 업서 <楞嚴5:68b>
 ㄱ'. 身微塵이 與造世界한 所有微塵으로 等無差別하야
 ㄴ. 내 몸과 界와 두 드트리 곧ᄃᆞᆫ 곧ᄃᆞᆫ미 업서 <楞嚴5:69a>
 ㄴ'. 我以諦觀身界二塵이 等無差別이라

(14 ㄱ~ㄴ)은 '와'형 비교구문 구조를 통해 비교대상과 비교기준의 일치성을 드러내는 한편, '곧ᄃᆞᆫ미 업서'라는 표현을 덧붙임으로써 두 개체의 완전한 동등에 대한 내용을 한 번 더 강조한 것으로 유추할 수 있다. 이는 (14 ㄱ'~ㄴ')의 대응 구문을 통해서도 확인할 수 있다.¹⁶⁾

한편, (15 ㄱ~ㄴ)은 밑줄 친 비교구문에 비교대상(NP1)이 문면에 드러나 있지 않으나, 문맥적 의미를 고려할 때 생략된 NP1의 내용은 선행문의 내용 전체로 상정해 볼 수 있다.

- (15) ㄱ. 이 몸 버리는 저근 因을 ㅎ디 아니ᄃᆞᆫ면 비록 無爲를 일위도 반드기 도로 사
 르미게나 아랫 비들 가포디 내 馬麥과 正히 곧ᄃᆞᆫ 달음 업스리라 <楞嚴6:105a>
 ㄱ'. 若不爲此捨身微因하면 縱成無爲라도 必還生人하야 酬其宿債호대 如我馬
 麥하야 正等無異하리라
 ㄴ. 이 呪를 誦持ᄃᆞᆫ면 壇애 드러 行道ᄃᆞᆫ 功德과 도로 곧ᄃᆞᆫ 달오미 업스리라
 <楞嚴7:53b>
 ㄴ'. 縱不作壇하고 不入道場하고 亦不行道라도 誦持此呪하면 還同入壇 行道功德
 하야 無有異也하며

이때 (15 ㄱ)은 선행문에 제시한 행위의 인과적 속성과 '내 馬麥함'의 인과적 속성이 동일함을 나타내며, (15 ㄴ)은 '이 呪를 誦持함'이 '壇애 드러 行道ᄃᆞᆫ 功德'과 동일함을

16) (14 ㄴ)의 비교대상과 비교기준의 의미적 관련성이 아래 구문 유형과 유사하여 이에 대한 비교 검토가 요구된다.

- (i) ㅅᄃᆞᆫ과 ㅅᄃᆞᆫ과 因호미 곧ᄃᆞᆫ <楞嚴5:11a>

앞서 우리는 (i)가 복수 비교대상 사이의 비교기준으로 한정된 부분적 일치성을 드러낸 유형임을 검토하였다. 이는 (i)는 [[ㅅᄃᆞᆫ과 ㅅᄃᆞᆫ]이 [因호미]이 곧ᄃᆞᆫ] 구조를 가지며, 두 비교의 대상이 공통적으로 내포한 '인호미'는 비교기준의 자격으로 두 번째 명사구 'NP-이'의 형태로 나타난 것으로 파악한 것이다. 한편, (14 ㄴ)은 비교의 대상과 비교기준의 의미적 상관성, 즉 티끌이 몸과 세계의 부분적 요소인 점으로 미루어 볼 때 (i)와 유사한 관점에서 해석될 여지가 있다. 그러나 두 구문의 언어적 형식은 이들이 동일하지 않음을 직접적으로 드러낸다. (i)는 중세국어 공동격 조사 '와 / 과' 결합 방식의 전형적 특성을 그대로 따르고 있음에 반해, (14 ㄴ)에는 마지막 명사구 '界'에 실질적 의미 기능을 수행하는 조사가 확인되지 않는다. 이는 실제 의미 기능을 하는 조사가 생략된 구문 구조로 보는 것이 옳을 듯하며, 이에 문맥적 의미를 함께 고려하여 (14 ㄴ)은 (i)와는 다른 [[내 몸과 界]의 [두 드틀]]이 곧ᄃᆞᆫ]의 구조로 그 의미를 파악하는 것이 보다 타당하리라 여겨진다. 이러한 구조상의 차이는 의미상 주된 비교 대상을 특정하려는 언중의 인식에서 비롯된 것으로 유추해 볼 수 있다. 즉 (i)가 평가자가 비교 행위의 중심적 자격을 'ㅅᄃᆞᆫ과 ㅅᄃᆞᆫ'에 부여하여 이들의 부분적(因) 동일함을 보이기 위한 구조라면, (14 ㄴ)은 '두 드틀'이 비교 행위의 중심이 되어 '(내 몸의) 드틀과 (界의) 드틀'의 전면적 동일함을 드러내기 위한 것으로 이해해 볼 수 있다.

나타냄을 전체 문맥을 통해 어렵지 않게 유추할 수 있다. 또한 ‘正히 같흐야 달음 업스리라 (正等無異)’, ‘달오미 업스리라 (無有異也)’와 같은 부가적 진술을 통해 [완전한 일치] 를 보다 강조하고 있음도 다시 한번 확인할 수 있다.

4. 결론

본고는 중세국어의 구체적 용례 분석을 통해 ‘NP1-이 NP2-이 같흐다’ 구문과 ‘NP1-이 NP2-와 같흐다’ 구문의 통사-의미적 기능을 밝히고, 이를 통해 두 비교구문의 언어의 형식적 차이는 동등함의 정도를 언어적으로 구별하고자 한 언중의 인식에서 비롯된 것임을 논의해 보았다.

본고는 앞선 연구에서와 마찬가지로 ‘NP1-이 NP2-이 같흐다’ 구문과 ‘NP1-이 NP2-와 같흐다’ 구문이 각각의 독립된 의미 기능을 수행하고 있다는 점에는 동의하나, 이 두 유형을 ‘비유’와 ‘비교’의 관점에서 일반화하는 것이 타당하지 않음을 용례 검토를 통해 입증하였다. 하나의 언어 형식에 할당된 일반적 의미 기능에 대한 예외적인 개별 용례가 발견됨은 문법 체계를 마련함에 있어 효율적이지 않으므로, 기존 연구의 결과에서 벗어난 예외적 용례를 모두 아우를 수 있는 방안을 모색하여 ‘NP1-이 NP2-이 같흐다’ 구문과 ‘NP1-이 NP2-와 같흐다’ 구문의 의미 기능을 다음과 같이 정리하였다.

먼저 중세국어에 특수하게 나타나는 ‘NP1-이 NP2-이 같흐다’ 구문이 비교대상 (NP1)에 대한 비교기준 (NP2)의 부분적 (조건적) 일치성을 나타내는 언어적 형식임을 확인하였으며, 앞선 연구에서 밝힌 바 있는 ‘비유’라는 의미 기능 역시 ‘비교’의 하위 개념으로서 비교기준의 부분적 속성에 빗대어 비교대상을 표현하는 방식이기에 ‘NP1-이 NP2-이 같흐다’ 구문이 비유를 위한 가장 적절한 형식으로 자리잡게 되었음을 예측해 보았다.

한편, ‘NP1-이 NP2-와 같흐다’ 구문은 비교대상과 비교기준과의 전면적 일치성을 드러내는 언어적 표현임을 다양한 용례 분석을 통해 확인하였다. 비교대상과 비교기준의 의미 영역상의 구체적 연관성을 검토하고, ‘같흐야’를 포함한 선행문과 이에 뒤따르는 후행문 사이의 의미적 관련성을 추적하여 ‘와’형 구문이 언중의 비교대상과 비교기준에 대한 완전한 일치성을 나타내는 유형적 특징을 보임을 파악해 보았다. 또한 완전한 일치성을 드러내는 구체적 표현들이 덧붙여 이를 강조하고자 하였음을 용례 검토를 통해 확인하였다.

이상의 검토 내용을 바탕으로 우리는 당대 언중들이 동등성에 대한 정도를 언어적 형태로 표현하고자 하였음을 입증해 보았다. 앞서 지적하였듯이 본 연구는 앞선 연구들의 성과에 대해 일정 부분 동의하면서, 기존 논의에서 도출된 결과가 아우르지 못하는 예외적 용례들에 대한 구문 해석의 실마리를 풀어보고자 한 것에서 출발하였다. 따라서 이번 논의의 결과는 예외적 현상으로 분류되는 용례들에 대한 해석상의 난점을 해결할 수 있는 대안적 방편 마련을 위한 첫걸음이 될 수 있으리라 생각한다. 이에 덧붙여 언어 변화 과정 속에 나타나는 ‘같흐다’에서 파생된 다양한 형태들의 통사-의미적 특성과 이들이 가지는 범주의 유동성을 이해함에 있어 본 연구의 결과물이 하나의 기초 자료로 활용될 수 있을 것이라 기대한다.

<<참고문헌>>

- 고영근·구본관(2008) 『우리말 문법론』, 집문당.
- 김정대(1991), 『한국어 비교구문의 통사론』, 계명대 박사학위논문.
- 김정아(1986), 「'갈하다'구문의 통사·의미적 특성」, 진단학보 62, 127-142면, 진단학회.
- 김정아(1993), 『15세기 국어의 비교구문 연구』, 서울대 박사학위 논문.
- 김정아(1998), 『중세국어의 비교구문 연구』, 태학사.
- 송홍규(2016), 「차등비교구문과 정도차 표현」, 동악어문학 69, 145-169면, 동악어문학회.
- 오경숙(2003), 「국어 후치사 비교구문 연구」, 서강대 박사학위논문.
- 이기문(1972/1998), 『국어사개설』, 태학사.
- 이승녕(1961), 『중세국어 문법』, 을유문화사.
- 이현희(1994), 『중세국어 구문 연구』, 신구문화사.
- 이현희(1999), 「'똥다' 구문에 대한 통시적 연구」, 진단학보 87, 71-126면, 진단학회.
- 하길중(1997), 「현대 한국어 동등비교구문의 의미 연구」, 한국어학 5, 229-266면, 한국어학회.
- 하길중(1999), 『현대 한국어 비교구문 연구』, 박이정.
- 허웅(1975/1989), 『우리 옛말본』, 샘문화사.
- 황미향(1996), 「'갈-'구문의 통사와 의미」, 국어교육연구 28, 201-224면, 국어교육학회.

Chinese '(shi)' plus focuses and Search constraint at Determinacy

Ick-Hee Ihm
(Hankuk University of Foreign Studies)

KALS winter conference 16nd/Feb/2023

1. Introduction

- **Goals:**

- ✓What is the definition of phases?
- ✓What are constraints of Search at Determinacy?

2. Chomsky (2019) and problems

Chomsky (2019: 279) defines a workspace as a set of objects, and Chomsky (2019: 280) states two points: First, a sequence Σ (1) can be determined from that workspace.

$$(1) \Sigma = (X_1, X_2, X_3, \dots, X_n)$$

He states that this Σ has the following properties: it's the shortest sequence in which (i) X_i is accessible (ii) Σ exhausts the workspace. Second, he defines as MERGE an operation that replaces the members of the sequence with a set (MERGE denotes a function and Σ is an input item of that function), as illustrated in (11) below.

$$(2) \text{MERGE}(\Sigma) = \{\{X_1, X_2\}, X_3, \dots, X_n\}$$

2. Chomsky (2019) and problems

To see more precisely, let us consider four syntactic objects *a*, *b*, *c*, and *d*. For example, following the definition of MERGE: (i) in (3a), the EM of *a* and *b* and MERGE yield (3b); and (ii) in (4a), the IM of *c* and $\{a, \{b, c\}\}$ and MERGE produce (4b).

$$(4) \text{ a. } WS_1 = [a, b, c, d], \Sigma_1 = (a, b, c, d) \rightarrow \text{ b. } \text{MERGE}(\Sigma_1) = WS_2 = [\{a, b\}, c, d]$$

$$(5) \text{ a. } WS_1 = [\{a, \{b, c\}\}, d], \Sigma_1 = (c, \{a, \{b, c\}\}, d) \rightarrow \text{ b. } \text{MERGE}(\Sigma_1) = WS_2 = [\{c, \{a, \{b, c\}\}\}, d]$$

According to the property, this computation continues until Σ exhausts the workspace. In other words, in (5), *a*, *b*, *c*, *d*, $\{c, d\}$, $\{b, \{c, d\}\}$ and $\{a, \{b, \{c, d\}\}\}$ are all accessible to further operations. However, Chomsky (2019: 275) mentions that Merge should adhere to the principle of Determinacy (6), preventing ambiguous copy application.

The square brackets indicate the WS, but the WS is just a set. Therefore, it should be noted that the square brackets are equivalent to the curly brackets that are used for unordered sets, although I do not use the latter for the WS to avoid the notational complexity.

2. Chomsky (2019) and problems

(6) If the structural condition for a rule holds for some workspace, then the structural change must be unique, it must be determinate, unlike, say, phrase structure grammar.

We can interpret (6) as (i) (7) below and (ii) an ambiguous application of one element should not arise for the subsequent derivation.

(7) Determinacy applies at the output of MERGE.

To see more precisely, let us consider (8), where MERGE takes WS_1 as its input and maps it to WS_2 by applying IM to c.

- (8) a. $WS_1 = [\{a, \{b, c\}\}, d]$
 b. $WS_2 = [\{c, \{a, \{b, c\}\}\}, d]$

2. Chomsky (2019) and problems

But the following two concerns for the MS-approach can be pointed out: First, the Determinacy violation problem can be resolved by the Phase Impenetrability Condition (PIC) (i.e., an independently motivated condition on Transfer) with no appeal to the MS rule. Second, regarding accessibility, there seems to be a redundancy between MS and PIC: that is, MS makes the effects of PIC vacuous when we decide whether or not an element is accessible. If the MS rule decides the accessibility to a syntactic object, PIC is made vacuous.

3. Goto & Ishii (2020) and problems

To address the issues raised in the previous subsection, Goto & Ishii (ibid.) proposes a rule for Determinacy application timing (9) below.

(9) Determinacy applies at the input of MERGE.

Specifically, a Determinacy rule violation will occur if an ambiguous application to a syntactic object in a WS_x is made at the present stage of the MERGE derivation.

To see more precisely, let us first examine the following (10)(=8) case, where MERGE takes WS₁ as its input and maps it to WS₂ and the rule is not violated.

- (10) a. WS₁ = {{a, {b, c}}, d}
 b. WS₂ = {{c, {a, {b, c}}}, d}

3. Goto & Ishii (2020) and problems

According to the rule (17), in WS₁, there is only one copy of *c*, and so there is only **one option** to create WS₂: to move *c* in the base position. This means that the syntactic object *c* is an unambiguous one in WS₁. Thus, no Determinacy violation occurs in (18). Let us next examine the following (19) case where MERGE takes the WS₂ as its input and maps it to WS₃ and the rule is violated.

- (11) a. WS₂ = {{c, {a, {b, c}}}, d}
 b. WS₃ = {c, {c, {a, {b, c}}}, d}

As we can see, in (19a) WS₂, there are two copies of *c* at the input of MERGE: *c* and *c*. In other words, there are two options to create WS₃, i.e., either to move *c* or to move *c*. This means that *c* is accessible ambiguously in the WS₂. Thus, the rule violation is induced (a Determinacy violation arises).

3. Goto & Ishii (2020) and problems

Goto & Ishii (*ibid.*: (10)) show how successive-cyclic IM is ensured under Input Determinacy without relying on MS (a notion that plays an important role in Output Determinacy), taking the example in (20) below (the gray indicates PIC domain).

- (12) What did you say that John bought t?
- a. [RP what [R(buy) what]]
 - b. [CP what [C [TP John [T [VP John [v [RP what [R what]]]]]]]]
 - c. [VP you [v [RP what [R(say) [CP what [C-that [TP John [... [RP what
 - d. [CP what [C(that) [TP you [T [VP you [v [RP what [R [CP what [...

Goto & Ishii (*ibid.*: (11)) show that the Input Determinacy can derive the subject island effect such as (13) (Chomsky 1973, Huang 1982). The derivation of (13a) is (13b):

- (13) *a. Who did [pictures of t] please you?
- b. [CP who [C-did [TP [pictures of who] [T [VP [pictures of who] [v [...

3.1. The problem of phase definition

To support the Goto & Ishii (2021) mechanism, a standard for phase should be clearly established. However, this mechanism is not sufficiently defined.

Chomsky (2005) states that in finite clauses, the phase head C possesses an Agree feature and a Tense feature at the stage where it is introduced into the narrow syntax. Based on this background, Kanno (2008) propose (5) below as a condition that determines the phasehood of CPs (Kanno 2008: (5)).

- (5) The phasehood of CPs is determined by the combination of an Agree feature and a Tense feature; if a CP possesses both Agree and Tense features, it is a phase; if a CP does not have one or both of the two features, it is not a phase.

3.1. The problem of phase definition

Under his proposal, finite complements have both features and are thus considered to be phases, while control, ECM, and raising complements lack one or both features, thus not being phases. However, this view is problematic; to illustrate, let us look at the examples (15) and (16).

- (15) a. O que (.que) perguntaste como/quando arranjar?
 what wondered-you how/when to-fix
 'What did you wonder how/when to fix?'
 b. * O que (.que) perguntaste como/quando arranjares?
 what wondered-you how/when to-fix.2sg
 'What did you wonder how/when to fix?'

- (16) a. A different student wanted to read every book. ($\forall > \exists$) (Johnson (2000: 199))
 b. A different student wanted for you to read every book. ($*\forall > \exists$) (ibid.)

3.2. Proposal

Concerning phases, I adopt the Chomsky's (2015: 5) statement below.

- (17) Languages have unvalued features, assigned values in certain structural positions. These features mark phases, a particular execution of strict cyclicity, well-motivated on grounds computational efficiency.

Extending this suggestion, I argue that an unvalued phi feature constitutes a phase, which can successfully capture the empirical facts in (5b) and (6b). In (5b), C of the inflected control complement does constitute a phase because it bears an unvalued phi feature, while in (5a) C of the non-inflected control complement does not bear an unvalued phi feature and no phases are constituted. So, the former disallows extraction and the latter allows it. Regarding (6b), C of control complement has an unvalued phi feature, which is made evident by the existence of DP in Spec TP. A phase that prohibits extraction is constituted. So, inverse scope is disallowed.

3.5. Proposal

- (21) a. Zhangsan song-GEI-le Lisi yi-ben-shu, Xiaoming song-GEI-le Laowang yi-ge qianbihe.
 Zhangsan offer-give-PERF Lisi a-CL-book, Xiaoming offer-give-PERF Laowang a-CL pencil case
 ‘Zhangsan gave Lisi a book and Xiaoming gave Laowang a pencil case.’
 b. *Zhangsan song-GEI-le Li si yi-ben-shu, Xiaoming △/shi Laowang yi-ge qianbihe.
 Zhangsan gave to Li si a-CL-book, Xiaoming did to Laowang a-CL pencil case
 ‘Zhangsan gave Li si a book and Xiaoming did Laowang a pencil case.’
- (22) a. Zhangsan song-le yi-ben-shu gei Li si, Xiaoming song-le yi-ge qianbihe gei Laowang.
 Zhangsan gave a-CL-book to Li si, Xiaoming gave a-CL pencil case to Laowang
 ‘Zhangsan gave a book to Li si and Xiaoming gave a pencil case to Laowang.’
 b. *Zhangsan song-le yi-ben-shu gei Li si, Xiaoming △/shi yi-ge qianbihe gei Laowang.
 Zhangsan gave a-CL-book to Li si, Xiaomingdid a-CL pencil case to Laowang
 ‘Zhangsan gave a book to Li si and Xiaoming did a pencil case to Laowang.’

3.5. Proposal

- (23) a. [_{TopicP} Xiaoming_i [_{FocusP} Laowang_j [_{IP} t_i [_{AspP} song-GEI-le t_j ... yi-ge qianbihe de]]]]
 (topic movement and the object movement after the topic)
 b. [_{TopicP} Xiaoming_i [_{FocusP} Laowang_j [_{FocusP} yi-ge qianbihe_k [_{IP} t_i [_{AspP} song-GEI-le t_j ... t_k de]]]]
 (topic movement and the object movement after the topic)
 c. [_{TopicP} Xiaoming_i [_{FocusP} shi Laowang_j [_{FocusP} yi-ge qianbihe_k [_{IP} t_r [_{AspP} song-GEI-le t_j ... t_k de]]]]
 (a direct focus marker *shi* insertion, and IP deletion at PF)

- Search should obey focus constraint.

4. Conclusion

- An unvalued phi feature constitutes a phase.
- Focus constraint should be obeyed in the typical Chinese DOC gapping construction.

References

- Cole, Peter and Li-May Sung (1994) "Head Movement and Long-Distance Reflexives," *Linguistic Inquiry* 25, 355-406.
- Huang, C. T. James. (1982) *Logical relations in Chinese and the theory of grammar*, Doctoral dissertation, MIT.
- Huang, Cheng-Teh James, and Yen-Hui Audrey Li (1996) "Recent generative studies in Chinese syntax," *New horizons in Chinese linguistics*, 49-96. Dordrecht: Kluwer.
- Ishii, Toru. 1997. *An asymmetry in the composition of phrase structure and its consequences*. Doctoral dissertation, University of California, Irvine.
- Ishii, Toru. 2011. The subject condition and its crosslinguistic variations. In *Proceedings of western conference in linguistics (WECOL) 2011*: 407-418.
- Kayne, Richard. 1983. *Connectedness and binary branching*. Dordrecht: Foris.
- Otsuka, Tomonori (2014) "An Extension of Feature-Inheritance," *English Linguistics* 31, 509-544.
- Lasnik, Howard and Mamoru Saito. 1992. *Move a: Conditions on its application and output*. Cambridge, MA: MIT Press.
- Richards, Marc D. (2012) "On Feature Inheritance, Defective Phases, and the Movement-Morphology Connection," *Phases: Developing the Framework*, ed. by Ángel J. Gallego, 195-232, De Gruyter Mouton, Boston, MA.
- Saito, Mamoru and Keiko Murasugi. 1999. Subject predication within IP and DP. In *Beyond principles and parameters*, ed. by Kyle Johnson and Ian Roberts. Dordrecht: Kluwer.
- Saito, Mamoru. 2007. Notes on East Asian argument ellipsis. *Language Research*. 43: 203-227.
- (Saito (2017a: 1))
- Saito, Mamoru (2017b) "A Note on Transfer Domain," *Nanzan Linguistics* 12, 61-69.

코퍼스

Keyword analysis of the IMO Standard Marine Communication Phrases

February 16th, 2023

Meihan Wang
(Korea Maritime and Ocean University)

KASL winter program

CONTENTS

1 Introduction


2 Literature Review

3 Data and Methodology

4 Results and Analysis

5 Conclusions

1 Introduction

 The standardized maritime communication language used between and within ships is contained in the Standard Maritime Communication Phrase (SMCP) (IMO, 2001). The SMCP (commonly referred to as the Standard Phrase) is a revised, updated and enhanced version of the Standard Marine Navigational Vocabulary, which was developed after the fire on the Scandinavian Star ferry that cost the loss of 158 human lives.

➤ As IMO explains in the introduction to the SMCP, the purpose of compiling these phrases is to help improve the safety and conductivity of ships and also to standardize the verbal language used in communication at sea, in ports, on waterways and on board ships.

1 Introduction

Research questions

- 1 What communicative characteristics of the IMO Standard Marine Communication Phrases can be found in the keyword analysis?
- 2 What are the issues that should be paid attention to when translating the language of maritime communication?

2 Literature Review

Pritchard and Kalogjera (2000:185) define Maritime English (or English for Maritime Communications) for VHF communication purposes as "a specific, narrow-scope realization of Maritime English", a highly restricted subspecies of Maritime English characterized by "a limited vocabulary and simple grammatical structure to suit the specific requirements of interpersonal communication and interaction".

Approximately 20% (De la Campa, 2003) or even 30-40% (Trenkner, 2000) of shipping accidents resulting in emergencies, injuries to the marine environment and loss of human life are due to the failure, poor or careless use of basic Maritime English for communication.

2 Literature Review

"Poor language abilities contribute to unsafe acts by seafarers, violations of good practice and established rules, work underperformance and, consequently, to carelessness and confusion which may embody tragic results. The consequences of such inadequate English language knowledge and skills may range from mere annoyance to normal operations to all sorts of occupational accidents, personal injury, death, cargo loss, damage to property or irreversible environmental impact" (BocanegraValle, 2010: 158).

3 Data and Methodology

The IMO SMCP was the target corpus containing 34865 running words in total. BNC Baby (spoken) was chosen as a reference corpus. The resulting refined BNC Baby used in this study contained 799730 running words in total.

Table 3.1 TTR and STTR of IMO SMCP and refined BNC Baby (spoken corpus)

	IMO SMCP	Refined BNC Baby (spoken corpus)
tokens	32221	746960
types	2703	9902
type/token ratio (TTR)	14.07%	9.95%
standardized TTR	29.43%	30.61%
STTR std. dev.	64.17	63.26

3 Data and Methodology

WordSmith Tools was developed by the British linguist Mike Scott at the University of Liverpool. The core areas of the software package include three modules:

- (1) Concord is used to create concordances, so all the hits from a search within a previously defined body text.
- (2) WordList lists all the Words or on word forms that are included in the selected corpus and statistical data are different from the text corpus.
- (3) KeyWord creates a list of all those words and word forms according to certain statistical criteria in the text corpus significantly occur rarely or frequently.

4 Results and Analysis

4.1 Standardized verbal message patterns

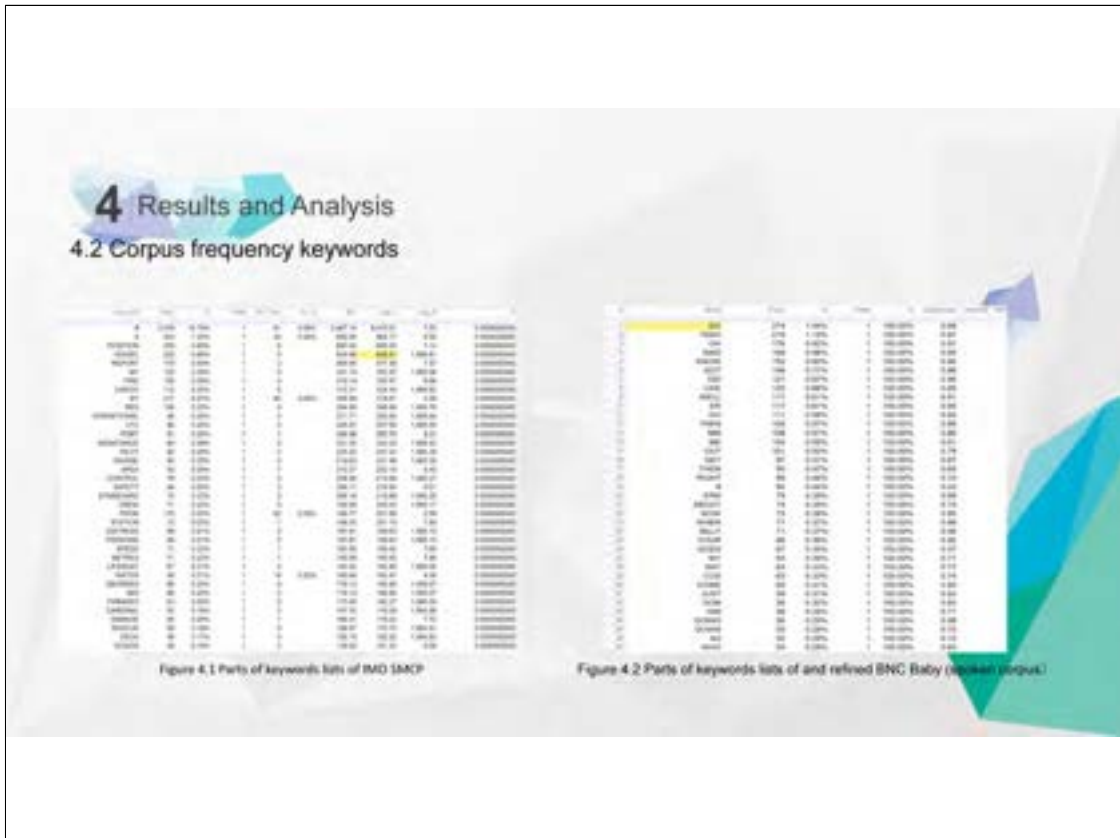
- Two main patterns:
 - > 1. Special-purpose messages
 - a. Distress messages
 - b. Urgency messages
 - c. Safety messages
 - > 2. Routine messages
 - a. Instruction
 - b. Advice
 - c. Warning
 - d. Information
 - e. Question
 - f. Answer
 - g. Request
 - h. Intention

4 Results and Analysis

4.1 Standardized verbal message patterns

Table 4.1 Examples of SMCP phrases as classified by message patterns and types

Message pattern ¹⁾	Message type ²⁾	Message marker ³⁾	Example ⁴⁾	SMCP phrase no. ⁵⁾
Special purpose ⁶⁾	Distress	Moylay Moylay Moylay	I require tug assistance... ...	A1.1.1.2.3 ⁷⁾ ...
	Urgency	Pan-Pan Pan-Pan Pan-Pan	I have lost dangerous goods of IMO-Class... ...	A1.2.2.1 ⁷⁾ ...
	Safety	Securite Securite Securite	Navigation is dangerous in area around...due to floating ice... ...	A1.3.1.2.2 ⁷⁾
Routine ⁶⁾	Instruction	Instruction	Do not tie the boat cables... Stand by on VHF Channel... ...	A1.5.1.10 ⁷⁾ A1.2.3.5 ⁷⁾ ...
	Advice	Advice	Warning: Uncharted rocks in position... ...	A1.1.1.4.4 ⁷⁾
	Warning	W	Warning: Uncharted rocks in position... ...	A1.1.1.4.4 ⁷⁾
	Information	Information	MV... helicopter is on the way to you... ...	A1.5.1.4 ⁷⁾
	Question	Question	Is the fire under control?... ...	A1.1.1.1.6 ⁷⁾
	Answer	Answer	Yes, fire is under control... ...	A1.1.1.1.6.1 ⁷⁾
	Request	Request	I require a helicopter... ...	A1.5.1.1 ⁷⁾
	Intention	Intention	I am ready to get underway... ...	A1.6.2.3.3.10.1 ⁷⁾



4 Results and Analysis

- Main characteristics of the IMO Standard Marine Communication Phrases
 - ✓ providing fully worded answers to "yes/no"-questions and basic alternative answers to sentence questions
 - ✓ using five tenses in maritime communication terms tenses: simple present tense, simple past tense, simple future tense, present progressive tense and present perfect tense.
 - ✓ omitting the articles the/a/an and the copula is/are
 - ✓ avoid using abbreviations
 - ✓ avoid using synonyms
 - ✓ mainly using the active voice in maritime communication
 - ✓ structuring the corresponding phrases according to the principle: *identical invariable plus variable*

5 Conclusions

- In order to avoid confusion and errors, maritime communication terms must be accurate, standard, concise and clear. However, in practice, many coastal workers and crews do not use and translate maritime communication language proficiently and accurately. One of the reasons is that they do not understand the linguistic characteristics of the maritime communication language and the issues that should be paid attention to when translating. Therefore, there is a need to study the linguistic characteristics and translation of marine communication language.
- It is especially important to raise awareness of the use of SMCP and to infuse it into the teaching of maritime English as well as to apply it correctly and effectively to the real-life scenarios of maritime communication.
- With regard to English for marine communications, much work remains to be done in analyzing standardized language use in the workplace.

Reference

- IMO (2001). Standard Marine Communication Phrases. Resolution A.918(22), 29 November 2001. London: International Maritime Organization.
- IMO (1977). Standard Marine Navigational Vocabulary. Resolution A.380 (X), 14 November 1977. London: International Maritime Organization
- Johnson, B. (1995). Some features of maritime telex service communication. *English for Specific Purposes*, 14(2), 127-136.
- Johnson, B. (1999). English in the Global Maritime Distress and Safety System. *World Englishes*, 18(2), 145-157.
- Pritchard, B. (2000). Maritime VHF communications: standards versus practice. In *Proceedings of the 2nd IMLA Workshop on Maritime English in Asia* (pp. 44-54). Dalian: Dalian Maritime University.

Reference



- Pritchard, B. (2003a). Maritime English syllabus for the modern seafarer: safety-related or comprehensive courses? *WMU Journal of Maritime Affairs*, 2(2), 149-166.
- Pritchard, B. (2003b). Maritime Communications and IMO SMCP 2001 (draft version). Retrieved from http://www.pfri.uniri.hr/~bopri/documents/MaritimeCommunicationsandSMCP_001.pdf
- Pritchard, B. & Kalogjera, D. (2000). On some features of conversation in maritime VHF communication. In M. Coulthard, J. Cotterill & F. Rock, F. (Eds.), *Dialogue Analysis VII: Working with Dialogue* (pp. 185-196). Tübingen: Max Niemeyer Verlag.
- Trenkner, P. (2005). The IMO Standard Marine Communication Phrases – Refreshing memories to refresh motivation. In *Proceedings of the IMLA 17th International Maritime English Conference* (pp. 1-17). France: Marseille
- Squire, D. (2006). The human element. The importance of effective communication. A presentation to the Maritime Communications & Technology Conference, London. Retrieved from <http://www.he-alert.com/documents/published/he00615.pdf>






An Exploration of Metaphorical Expressions in Maritime English from the Perspective of Keywords

Mengyu Xu & Se-Eun Jhang
(Korea Maritime & Ocean University)



Abstract

- This paper proposes a feasibly identifying metaphors way to figure out adjective-noun metaphors in a self-built maritime news corpus through text-dispersion keyword analysis. According to the metaphor identification procedure(MIP), linguistic metaphors are founded and divided by six keyword categories. Moreover, it is found that metaphorical expressions can exist in keywords of a specific corpus, and there is a collocation relationship between keywords and metaphors.
 - **Keywords:** maritime English; linguistic metaphor; metaphor identification; keyword; collocation
- 

CONTENTS	01 Introduction	02 Previous studies
	03 Data and methodology	04 Locating metaphors
	05 Results	06 Conclusion

1. Introduction

- Metaphor, as one of the figurative languages, has always been a hot topic in the field of Linguistics.

The essence of metaphor is to use one thing to describe and understand another.

1. Introduction(Cont'd)

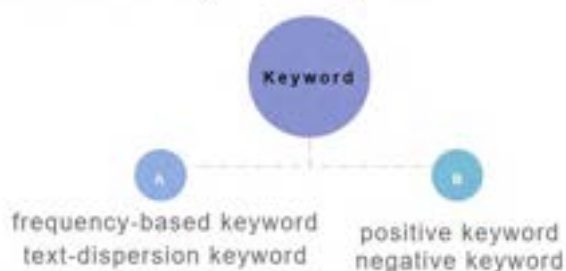


3 conceptual metaphor "This gadget will save you hours.", the conceptual metaphor is TIME IS MONEY.(Lakoff and Johnson, 2003:7)

Although the linguistic metaphor is the lowest of these three levels, it is the basis for studying the other two. Therefore, this paper will focus on this level.



1. Introduction(Cont'd)



This paper aims to find metaphorical expressions and explore the relationship between keywords and metaphors. This paper will use the MIP and keyword analysis to get metaphors in a specific corpus.



2.Previous studies

- Krishnakumaran and Zhu(2007) proposed three different kinds of metaphors with nouns.
- a. subject-linking verb-predicative structure --TIME IS MONEY
- b. verb, which is the focus of the metaphors representing the act of a subject noun on an object noun--The war ABSORBED his energy.
- c. adjective plus noun--SWEET girl
- Neuman, Assaf, Cohen et al.(2013) proposed three novel algorithms for identifying metaphors involving nouns based on the classification of metaphors with nouns.

2.Previous studies(Cont'd)

Input: An adjective-noun pair $\langle A, N \rangle$.

1. If the adjective A has a single dictionary definition then return LITERAL, else
2. If the noun N does not belong to any WordNet category, then return UNDECIDED, else
3. Let $\mathbf{N}(A)$ be the θ_{max} nouns most frequently collocated with A with mutual information of at least θ_{MI} .
4. Let \mathbf{N}^C be the k most concrete nouns in $\mathbf{N}(A)$.
5. Let $\mathbf{Cat}(A)$ be the set of all semantic noun categories containing at least θ_{cat} nouns in \mathbf{N}^C .
6. If N belongs to one of the categories in $\mathbf{Cat}(A)$, then return LITERAL, else return METAPHORICAL.

Figure 1. The adjective-noun algorithms proposed by Neuman, Assaf, Cohen et al.(2013)

- If an adjective only has one meaning in the dictionary, it can be regarded as a metaphor since metaphorical usage cannot exist in one sense only (Neuman, Assaf, Cohen et al., 2013).
- Wiktionary Dictionary
- Pariasa, Rajeg, and Sosiowati(2017) applied the method of the metaphor identification procedure(MIP) (Pragglejaz Group, 2007) and conceptual metaphor theory (CMT), getting several metaphors in Japanese poetry.

2. Previous studies (Cont'd)

- Question 1 : whether metaphors exist in the maritime English corpus ?
- Question 2 : whether keywords in the maritime English corpus can be seen as metaphors ?
- Question 3 : what is the relationship between keywords and metaphors?



3. Data and methodology

3.1 Data

- Table 1. the details about the study corpus and reference corpus

	Types	Tokens	Subjects	Averaged tokens	FTR	STTR
Study corpus	21,445	562,631	282	2,000	3.91%	43.31%
Reference corpus	995,937	1,012,785	500	2,000	4.80%	43.62%

Study corpus: Lloyd' s List, Maritime Reporter & Engineering News and The Journal of Navigation
Reference corpus: BE06 corpus

By setting the p-value as 0.000001 and min frequency as 5, **602** positive keywords and **143** negative keywords were obtained, totalling **745** keywords after using WordSmith 8.0.





3. Date and methodology(Cont'd)

- 3.2 Methodology

- 3.2.1 Text-dispersion keyword

- Text-dispersion keyword was proposed by Egbert and Biber (2019). Text-dispersion keyword analysis offered better results than frequency-based keyword analysis. WordSmith 8.0 (Scott, 2020) has concluded this kind of analytic method.

- 3.2.2 Metaphor identification procedure (MIP)

- Group proposed a metaphor identification procedure (MIP) in 2007.

- If the expression in the text is consistent with the most basic meaning of the word, we can consider the expression in the text to be literal. However, if the meaning in the text does not match the literal meaning of the word, then we can assume that the expression is metaphorical.



4. Locating metaphors(Cont'd)

- Step1: to make a text-dispersion keyword list with the help of WordSmith 8.0.

- Step2: to classify keywords according to specific wordlist in order to find metaphors' characteristic.

- Step 3: to find out all adjectives in keyword list and ensure the meaning number of each adjective, if an adjective only contains one meaning, the adjective is excluded, and vice versa.

- Step 4: to judge each <A,N>pair whether belongs to metaphorical expression based on the MIP

- Step 5: to calculate each metaphorical expression's MI value. If the value is more than or equals three, the expression is kept, and vice versa.



4. Locating metaphors (Cont'd)

- **Step1:** to make a text-dispersion keyword list with the help of WordSmith 8.0.
- p-value as 0.000001 and min frequency as 5, **602** positive keywords and **143** negative keywords were obtained, totaling **745** text-dispersion keywords.

• Table 2. Positive keywords and negative keywords of the study corpus

N	Keyword	Freq.	Term	BIC	Log L
1	WERE	1,141	244	478.87	485.63
2	MARITIME	931	229	398.63	405.29
3	NOW	586	204	397.73	404.39
...
603	WHATEVER	24	19	17.78	-24.25
604	EAT	5	5	17.60	-24.26
605	WATCHING	10	9	17.63	-24.29
...
743	HIS	80	40	91.85	-98.51
744	HIMSELF	9	8	98.86	-105.52
745	HIM	32	25	187.88	-144.54

4. Locating metaphors (Cont'd)

- **Step2:** to classify keywords according to specific wordlist in order to find metaphors' characteristic.
- Xu Lin (2021) provided 651 maritime science and technology words and 1684 maritime academic words.

• Table 3. The details about the keywords categories

Maritime Academic Keywords(Positive)	178
Maritime Science and Technology Keywords(Positive)	95
Other Positive Keywords	350
Maritime Academic Keywords(Negative)	9
Maritime Science and Technology Keywords(Negative)	1
Other Negative Keywords	133
Total	766

Some words are shown in the 651 maritime science and technology word list and 1684 maritime academic word list at the same time, so the total number of keywords 766 in the table 3 are more than the 745 keywords.

4. Locating metaphors(Cont'd)

- Step 3: to find out all adjectives in keyword list and ensure the meaning number of each adjective, if an adjective only contains one meaning, the adjective is excluded, and vice versa.
- Step 4: to judge each <A,N>pair whether belongs to metaphorical expression based on the MIP.

	Numbers	Adjectives	Adjectives having more than one meaning	Metaphors
Maritime Academic Keywords(Positive)	178	30	30	1
Maritime Science and Technology Keywords(Positive)	95	14	4	0
Other Positive Keywords	390	46	15	0
Maritime Academic Keywords(Negative)	9	0	0	0
Maritime Science and Technology Keywords(Negative)	1	1	1	0
Other	133	20	16	6



5. Results

- 5.1. Metaphors in maritime academic keywords (positive)
- Eg1. "We expect to see the first pre-commercial arrays - that's under 100MW by 2022 - and the first **commercial arrays** over 100MW by the middle of the decade," Lewis says. (Maritime Reporter & Engineering News•January 2021,page 9)
- basic meaning: a large group of people or things that are related in some way
- literal meaning: the floating offshore wind industry, where power generation devices are arranged in neat rows, looks like neat arrays.





5.Results(Cont'd)

- Eg2. That, in turn, assumes a significant shift in the **competitive landscape** of vessel ownership and, long term, we anticipate ESG capital requirements will increasingly favour the stability of consolidated fleets matched against long term cargo interests, eliminating much of the generational privilege felt by traditional ownership models.
- basic meaning: visible features of countryside or land
- literal meaning: commercial forms or patterns



5.Results(Cont'd)

- Eg3. Thirdly, a machine learning algorithm is utilised to predict the **critical passing distance** between vessels under different conditions through regression. (JN results)
- basic meaning: extreme value at which an object changes differently than it does below it.
- literal meaning: the minimum safe distance between two ships.





5.Results(Cont'd)

- Eg4. The chosen configuration of four generators, two large and two small, on an **integrated bus** provided the best overall reliability and fuel economy. (MR results)
- basic meaning: one kind of transportation
- literal meaning: integrated electronic circuits



5.Results(Cont'd)

- Eg5. It is entirely possible to read the runes and make a **strategic play** in the shipping market based on the supply and demand dynamics of 2020. (results001)
- basic meaning: script, game
- literal meaning: strategy





5.Results(Cont'd)

- Eg6. Even in the best of conditions, there are learning curves and **technical hiccups** when integrating new machinery across a fleet, and Reinauer was not exempt.
- basic meaning: gurk or burp
- literal meaning: problem



5.Results(Cont'd)

- Eg7. This is necessary to counter the stresses and strains from repetitive operations in tough sea conditions on a davit's moving parts, such as shock absorbers and winches, to ensure it remains fit for an **operational window** of up to sea state six.
- basic meaning: an actual object used for ventilation
- literal meaning: an operational page, or an operating platform





5.Results(Cont'd)

- There are seven metaphorical expressions in maritime academic keywords: commercial *arrays*, competitive *landscape*, *critical* passing distance, integrated *bus*, strategic *play*, technical *hiccups*, and operational *window*.
- Question 1 : Whether metaphors exist in the maritime English corpus ?
- **The answer is yes.** In the maritime English corpus,there exist the metaphorical expressions.



5.Results(Cont'd)

- Question 2 : whether keywords in the maritime English corpus can be seen as metaphors ?
- **The answer is also yes.** From the data above,it is concluded that although there one phrase can prove that keyword can be regarded as the metaphor, the answer is positive.
- Question 3 : What's the relationship between keywords and metaphors?
- From the six phrases above, it is proved that keywords and metaphorical expressions are often matched together, and adjective keywords can modify metaphorical expressions.



5. Results(Cont'd)

• 5.2. Metaphors in other negative keywords

- Table 5. Metaphorical expressions and their text meanings in corpus

N	Metaphorical expressions	Text meanings
1	dead letter	No valid warranty
2	dead space	Invalid area
3	red tape	Fussy and redundant things
4	young fleet	Ships with short service periods
5	black swan	An unusual event
6	live session	Real time content
7	live data	Real time content
8	live fleet	A ship in service

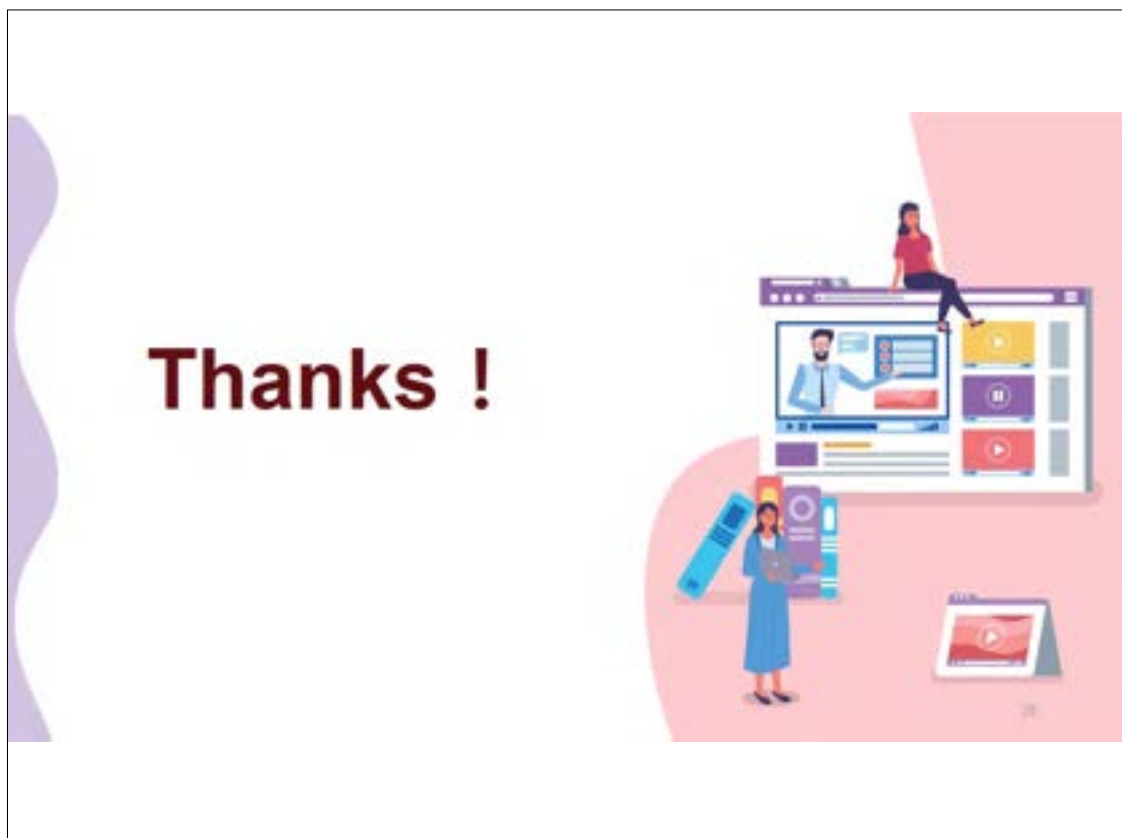
6. Conclusion

- This paper studied the metaphorical expressions in the maritime English corpus. According to Xu Lin's maritime vocabulary list, the text-dispersion keywords are classified, and finally the <A, N> metaphorical expressions in the corpus are retrieved according to MIP. We draw three conclusions. First, there are some specific metaphorical expressions in maritime English. Secondly, keywords in the corpus can be used as metaphorical expressions, but their frequency is relatively low. Finally, there is usually a certain collocation relationship between keywords and metaphorical expressions.



References

- Egbert, J., & Biber, D. (2019). Incorporating text dispersion into keyword analyses. *Corpora*, 14(1), 77-104.
- Group, Praggelax. (2007). MIP: A method for identifying metaphorically used words in discourse. *Metaphor and Symbol*, 22(1), 1-39.
- Keveszes, Z. (2010). *Metaphor: A Practical Introduction*. Oxford: Oxford University Press.
- Krishnakumaran S, Zhu X. (2007). Hunting elusive metaphors using lexical resources. *Proc. of the Workshop on Computational Approaches to Figurative Language*. Stroudsburg: Association for Computational Linguistics. 13-20
- Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. Chicago: University of Chicago Press
- Lakoff, G., & Johnson, M. (2003). *Metaphors We Live By*. Chicago: University of Chicago Press
- Neuman Y, Assaf D, Cohen Y, Lavi M, & Argamon S, et al. (2013). Metaphor Identification in Large Texts Corpora. *PLOS ONE* 8(4): e62343. doi:10.1371/journal.pone.0062343
- Pariss, I. O., I. M. R. & I. O. A. S. (2017). The Application of Metaphor Identification Procedure (MIP) and Conceptual Metaphor in Japanese Poetry. *Linguistica*, 24(1), 138-157.
- Philip, O. (2010). Metaphorical keyness in specialized corpora. *Keyness in Texts*. Amsterdam: John Benjamins Publishing Company. 181-206
- Searle, J. R. (1985). *Expression and meaning. Studies in the Theory of Speech Acts*. New York: Cambridge University Press.



Lexical bundles in the academic writing of Vietnamese EFL learners

February 16th, 2023

Quynh Thuy Bui

Department of English Language and Literature
Korea Maritime and Ocean University

Outline

1. Introduction

2. Literature Review

3. Research Design

4. Results & Discussions

5. Conclusion

I. Introduction

Formulaic language is widely used in academic writing and is considered as a useful measure of language development

Many researchers have showed that formulaic sequences play an important role in discourse and consist of natural language, both spoken and written (Sinclair, 1991; Wray, 2002; Hyland, 2012)

- Many studies on formulaic language have taken a data-driven and frequency-based approach to identifying special types of formulaic sequence, called lexical bundles.
- Lexical bundles are groups of three or more words that frequently recur in a register (Biber, Johnsson, Leech & Finngan, 1999)
- Numerous studies have analyzed lexical bundles in academic prose such as academic text and student writing (Cortes, 2004; Hyland, 2008; Lehmann, 2013)

- There are a few studies on lexical bundles in the academic writing by Vietnamese EFL learners. This study is conducted to identify the structural types and function discourse used in the academic writing by Vietnamese EFL learners.
- Using structural taxonomy of Biber et al., (1999,2004) and functional framework by Hyland (2008), this study compares the use of lexical bundles in the argumentative essays composed by Vietnamese undergraduate students and American undergraduate students

Research questions

- 1) What are the differences between lexical bundles used by L1-English and L2-Vietnamese undergraduate students in terms of structural types?
- 2) What are the differences between lexical bundles used by L1-English and L2-Vietnamese undergraduate students in terms of functional types?

2. Literature Review

- Biber et al. (1999) make the first significant step into corpus study in investigating the frequent word combination and divide into two categorizations: functional and structural types.
- These word combinations are considered as lexical bundles and defined as “recurrent expressions that usually co-occur in natural language use, regardless of their idiomaticity and their lexical status”.

- Biber & Barbiera (2007), lexical bundles are “important building blocks in discourse”
- Biber & Corad (1999): lexical bundles are not easy to master since they are both complex and not fixed.
 - ⇒ EFL learners have difficulty in using the lexical bundles in their writing and speech.
 - ⇒ Numerous studies on structural and functional features of lexical bundles have been conducted recently (Bychkovska & Lee, 2017; Chen & Baker, 2010; Hyland, 2008; Pan et al., 2015; Yang & Fang, 2021)

3. Research design

Groups	Data source	Number of essays	Number of words
VE	Argumentative essays written by Vietnamese students	100	32029
AE	Argumentative essays written by American students – extracted from LOCNESS	74	32336

3.1. Identification of lexical bundles

- Antconc is used to retrieve 4 lexical bundles, frequency threshold: 30 per million words, and dispersion threshold is set at 3 different texts.
- After retrieval of these bundles, context-dependent or topic-specific bundles are excluded from the analysis (Chen & Baker, 2010)
- Manually check bundle overlaps by using concordance analyses

Table 1: Number of types and frequency of lexical bundles in the two corpora

Groups	No. of types	Raw Frequency	No. of words	Normalized frequency (pmw)
VE	138	666	32029	20794
AE	19	72	32336	2227

- These four-word lexical bundles are classified according to their structures and functions or meanings, using structural taxonomy by Biber et al., (2004) and functional framework by Hyland (2008)

4. Results and discussions

4.1 Comparison of structural types of lexical bundles across two corpora

Table 2: Distribution of structural categories in the two corpora

	Subcategories	Types		Tokens	
		VE	AE	VE(pwm)	AE(pwm)
NP-based	1a. Noun phrase with of-phrase fragment (my point of view)	19	6	75(2341.6)	25(773.1)
	1b. Noun phrase with other post-modifier fragment (my opinions with examples)	8	0	24(749.3)	0
	1c. Other noun phrases (a second major reason)	6	0	37(1155.2)	0
	Total	33	6	136(4246.1)	25(773.1)
PP-based	2a. Prepositional phrase with embedded of-phrase fragment (with the development of)	5	2	19(593.2)	6(185.6)
	2b. Other prepositional phrase fragment (in the long term)	20	5	163(5089.1)	21(649.4)
	Total	25	7	182(5682.3)	27(835)
VP-based	3a. Anticipatory <i>it</i> +verb phrase/adjective phrases (it is true that)	17	1	76(2372.8)	3(92.8)
	3b. Copula be +noun phrase/adjective phrase (is a valid point)	4	0	12(374.7)	0
	3c. (Verb phrase) + that clause fragment (believe that they are)	4	2	14(437.1)	6(185.6)
	3d. (verb/adjective+) to-clause fragment (do not need to)	3	2	9(281.0)	7(216.5)
	3e. Adverbial clause fragment (when it comes to)	1	0	3(93.7)	0
	3f. Prououn/noun phrase + be VP (+...) (There is no doubt)	20	1	105(3278.3)	4(123.7)
	3g. VP with active verb (support my opinion with)	25	0	107(3340.7)	0
Total	74	6	326(10178.2)	20(618.5)	
Other	Others (more and more popular)	6	0	22 (686.9)	0
Total		138	19	666(20793.7)	72(2226.7)

Table 3 Proportional distribution of main structural categories in the two corpora

Categories	Types		Tokens	
	VE(%)	AE(%)	VE(%)	VE(%)
NP-based	33(24%)	6(32%)	136(20.5%)	25(34.7%)
PP-based	25(18%)	7(36%)	182(27.3%)	27(37.5%)
VP-based	74(54%)	6(32%)	326(48.9%)	20(27.8%)
Others	6(4%)	0	22(3.3%)	0
Total	138(100%)	19(100%)	666(100%)	72(100%)

- As can be seen in the Table 4, Vietnamese students use structural types of lexical bundle more frequently, over seven times than American students.
- American students are flexible to use three types of lexical bundles equally, Vietnamese students tend to rely on VP-based bundles, at the highest proportion of 54%

NP-based bundles

- NP-based bundles, there are 19 and 6 types of *noun phrase with of -phrase fragment* in VE and AE respectively
- In VE, *NP with of -phrase fragment* accounts for the highest proportion. 8 types of *noun-phrase post modifier fragment, such as my opinions with examples*. This type is not shown in VE

PP-based bundles

- The proportion of PP-based in AE is higher than in VE.
- Both groups prefer to use *other prepositional phrase fragment* than *Prepositional phrase with embedded of -phrase fragment*.

VP-based bundles

- American students do not use *Copula be +noun phrase/adjective phrase, Adverbial clause fragment*, while Vietnamese students tend to use all types of VP-based fragments
- In VE, the category of *VP with active verb* accounts for the most with 25 types, followed by *Pronoun/noun phrase + be/VP* with 20 types. The expressions *I completely disagree with, I firmly believe that, I do not agree with* are repeatedly used in Vietnamese students' essays, it is assumed that these students lack of lexical resources when presenting their ideas

4.2 Comparison of functional type of lexical bundles

- Table 4 Distribution of functional categories in the two corpora

Categories	Subcategories	Types		Tokens	
		VE	AE	VE (pmw)	AE(pmw)
1. Research-oriented bundles	A1. Location: indicating times/place	8(9.5%)	4(21.1%)	52(1623.5)	17(525.7)
	A2. Quantification	12(8.6%)	0	45(1405.0)	0
	A3. Description (one of which is)	2(1.5%)	1(5.2%)	13(405.9)	3(92.8)
	A4. Topic	16(11.6%)	5(26.3%)	56(1748.4)	18(556.7)
2. Text-oriented bundles	B1. Transition signals	11(8.0%)	1(5.3%)	89(2778.7)	5(154.6)
	B2. Resultative signals	10(7.3%)	1(5.3%)	37(1156.2)	3(92.8)
	B3. Structuring signal	6(4.3%)	0	53(1654.8)	0
	B4. Framing signals	6(4.3%)	2(10.5%)	24(749.3)	6(185.6)
3. Participant-oriented bundles	C1. Stance features	56(40.6%)	5(26.3%)	261(8148.9)	20(618.5)
	C2. Engagement features	10(7.3%)	0	36(1123.9)	0
Total	Total	138(100%)	19(100%)	666(20793.7)	72(2226.6)

Table 5: Distribution of main functional categories in the two corpora

Categories	Type (%)		Token(%)	
	VE (%)	AE(%)	VE(%)	AE(%)
Research-oriented bundles	39(28.3%)	10(52.6%)	166(24.9%)	38(52.8%)
Text-oriented bundles	33(23.9%)	4(21.1%)	203(30.5%)	14(19.4%)
Participant-oriented bundles	66(47.8%)	5(26.3%)	297(44.6%)	20(27.8%)
Total	138(100%)	19(100%)	666(100%)	72(100%)

- Vietnamese learners show a considerably more frequent use of functional bundles, over seven (types) and ten times (tokens) higher than the American students.
- While Vietnamese students strongly prefer to use participant-oriented bundles (47.8% of types, 44.6% of tokens), American students have tendency to employ research-oriented bundles (52.6% of types, 52.8% of tokens)

Research-oriented bundles

- American students employ more the topic-related bundles than Vietnamese learners, almost three times (11.6% and 26.3% of types) respectively.
- Vietnamese students use more bundles related to quantification (8.6%), such as *a great number of*, *a huge amount of*, *a large number of*, etc. According to previous studies (Biber et al., 1999, 2004; Bychkovska, Lee, 2017), these bundles are more frequently used in conversations than academic prose, making VE more colloquial and informal

Text-oriented bundles

- In terms of types, text-oriented bundles are the least commonly occurring ones in two groups' essays, 23.9% in VE and 21.1% in AE. In terms of tokens, this type of bundles make up a high proportion in VE, around 30.5%, and lower percentage in AE, at 19.4%.
- In AE, transition signals account for the highest percentage, at 13.4% as expected

Text-oriented bundles

- The lexical bundle *on the other hand* is used the most frequently, followed by "*on the one hand*". An underlying explanation for employing the transition signals the most commonly in VE might be that these students technically apply the format of argumentative essays which were instructed by their instructors

Participant-oriented lexical bundles

- Vietnamese students show a considerably more frequent use of participant-oriented bundles which, at 47.8% of types and 44.6% of tokens, is well nearly almost double the number by American students.
- Looking back the Table 4, stance features make up a dramatically significant proportion, around 40% of types and tokens while there are merely a few engagement features, at 7.3% of types and 5.4% of tokens

Participant-oriented lexical bundles

- Vietnamese students have tendency to express their idea explicitly in their essays with the first-person pronoun "I" (17 times"; which is opposed to the implicit means by American students.
- According to Hyland's (2005) Stance and Engagement framework of interpersonal discourse, the expressions such as *It is obvious that*, *it is clear that*, *it is apparent that* can be categorized as boosters

Participant-oriented lexical bundles

- The frequent use of boosters in VE may indicate the writer's attempt to show they have mastered a specific discourse pattern

5. Conclusion

- There are differences between Vietnamese students and American students in using structural types and functional types of lexical bundles.
- Vietnamese students use a wider range of four-word bundles, showing a overuse of structural and functional types.
- In terms of structural types, Vietnamese students highly rely on VP-based lexical bundles, why American students tend to employ three types of lexical bundles equally.

5. Conclusion

- In terms of functional types, in VE, participant-oriented bundles account for the most frequently, while in AE, research-oriented bundles are employed at the highest.

**From Gender-Biased to Gender-Specific and Gender-Inclusive Language:
A Corpus-Based Study on Maritime Conventions and Contemporary English**

February 6, 2023

LIU, Cuilin & JHANG, Se-Eun
(Korea Maritime & Ocean University)

The 2023 KASL Winter Program

Abstract

In this paper, we are interested in **some gender-inclusive and gender-specific words used in the maritime community** which is contemporarily still extremely male-dominated. The specific purposes of the present study are to gain insight into the trends in using SEAMAN and SEAFARER in maritime English and general English, to track changing features in using gendered and gender-inclusive personal nouns in maritime English, and to know the popular gender-specific words used in contemporary English. For those purposes, we will diachronically examine the frequency of some pre-determined terms (e.g. *seaman*, *seafarer*, *he or she*) as well as the context when necessary by relying on the corpus-based approach. Gender-inclusive words like *seafarer/s* and *he or she* has been dominantly used in maritime English as English for Specific Purposes. In contrast, gender-biased *seaman/woman* is still more frequently used than *seafarer/s*. Gender-specific words to refer to women began to appear in 2000s in maritime English and various forms (like *seawomen*, *female seafarers*) have appeared in contemporary English with compound versions more popular. Comparatively, gender-biased occupational nouns tend to gain more attention than gendered pronouns, which make gender-inclusive occupational nouns adopted earlier than pronouns. Moreover, gender-biased nouns tend to habitually co-occur with gendered pronouns, but it is not true of the behavior of all gender-inclusive occupational nouns, some of which still co-occur with gendered generic pronoun *he*.

Keywords: gender-biased, gender-inclusive, occupational reference, frequency, diachronic, ESP, corpus-based

Introduction

Gender-markers in English

Sexist languages

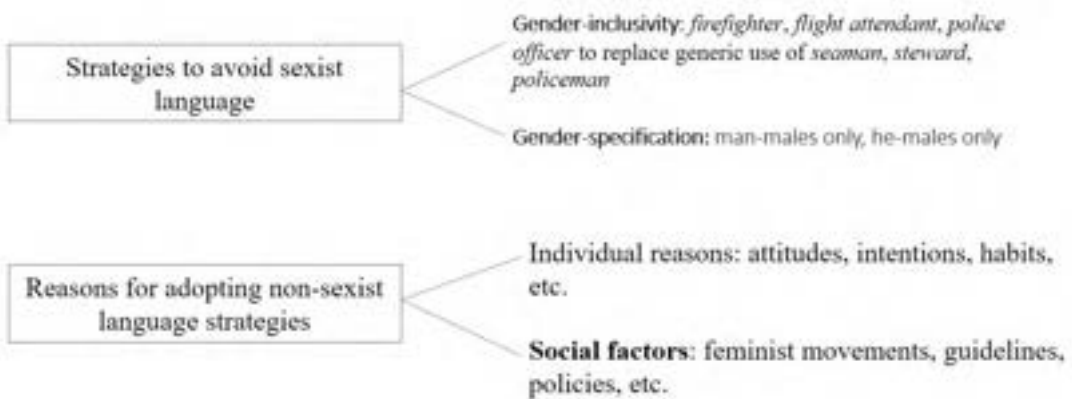
The present study

Gender markers are embedded in the **pronouns and possessives** (*he, she, her* and *his*); **a few nouns** referring to specifically masculine or feminine (*actor/actress, waiter/waitress*) and forms of address (*boys/girls, Mr/Miss/Mrs*). Another case needs special detailed attention is the **generic use of *man* or *he*** to refer to both genders. Gender in English is primarily a semantic category with important social implications (Hellinger, 2001).

Classification of linguistic sexism and examples

Linguistic sexism is the specific form of sexism underlying the above-mentioned forms or giving expressions to most of them.

Stereotypes	Common Sexist Forms	Examples
Invisibility of women (exclusive language)	generic use of <i>he-woman</i> language (eg. Janice Moulton 1981, Wendy Martyna 1978, Deborah Cameron 1985)	Man: gender-neutral meaning and gender-specific meaning -man compound words: mankind, chairman, fireman, policeman, mailman Personal pronoun: he/his/him
	Gender-marked terms (with sexist stereotypes)	family man but no family woman
Markedness for women (with sexist attitudes towards women)	Sexist suffixes	-ess, -ette, -trix: master-mistress, steward-stewardess, poet-poetess; usher-usherette; administer-administrix
	Sexist metaphorical reference terms	bitch, stud
	Sexist address terms	Miss; girls; Mrs
	Gender-marked women but not men	career woman, lady doctor



The present study

Among the corpus studies on gender and language, however, only a limited number have investigated gender-related representations from a diachronic perspective and with even less attention on English for Specific Purposes (ESP). Maritime English is one of the branches of English for Specific Purposes (Hong and Jhang, 2010) and is used as an official language within the international maritime community, which is traditionally male-dominated with women seafarers in deck and engine departments representing only 1% of all seafarers (BIMCO & ICS, 2016).

Hence, it will be especially interesting to **investigate the gender-related words and examine the implementation of the strategies to use gender-inclusive English in such a male-dominated industry.**

Research Questions

1. What are trends of using SEAMAN and SEAFARER in maritime English legal texts and general English over time?
2. How are the gendered and gender-inclusive personal pronouns used in maritime English over time?
3. To make symmetric reference to two genders, what gender-specific versions are adopted and popularly used in contemporary general English?

Purpose of the present study

- to gain insight into the trends in using SEAMAN and SEAFARER in maritime English(COMC) and general English(COHA)
- to track changing features in using gendered and gender-inclusive personal nouns **in maritime English(COMC)**
- and to know the popular gender-specific words used in contemporary English(NOW)

Procedures of the present study

Stage 1

the changing features of **occupational nouns** and **personal pronouns**-

- (1) frequencies of both the traditional generic form *seaman/men* and gender-inclusive form *seafarer/s* are firstly searched in COHA and COMC
- (2) frequencies of personal pronouns: nominative pronouns (*he, she, he or she*), objective pronouns (*him, her, him or her*) as well as possessive pronouns (*his, her, his or her*) in COMC
- (3) concordances of personal pronouns to get evidence for some assumptions

Stage 2

The use of gender-specific lexical terms males regarding the position of seafarers-

- (1) predetermined gender-specific terms (like *seawomen, male seafarers*) were examined by using the NOW Corpus.
- (2) Concordances were extracted for close observation

11

Data and Methods

- Corpus-based analysis:** The present study to larger extent falls in the corpus-based approach, which hypothesize that gender-inclusive words have been increasingly used and the idea to explore the formation of gender-specific and gender-inclusive words are also based on the widely-accepted gender-inclusive instances in general English.
- Corpora:** the Corpus of Historical American English (COHA) is selected for a diachronic investigation, the News on the Web (NOW Corpus) is used to represent the contemporary English, and the self-compiled Corpus of Maritime Conventions (COMC) contains all seafarer-related conventions and recommendations from 1920s to 2010s is the target ESP corpus.

Table 1. A list of maritime conventions for the ESP corpus

Appendix 1

Year	Year ID	Title
1920s	1921	Peace of Geneva Convention, 1921
	1922	Compulsory Arbitration (Geneva) Convention, 1922
	1923	Medical Examinations of Steam Power (Gene Convention), 1923
	1924	Labour Discipline (Geneva) Recommendations, 1924
	1925	Seaman's Articles of Agreement Convention, 1925
	1926	Regulations of Officers Convention, 1926
	1927	Statutes with By-Laws Convention, 1927
	1928	Peace of Stockholm and Geneva (Gene Convention), 1928
	1929	Agreement, Liability, Deck and Squared Riggers Convention, 1929
	1930	Seaman's Welfare in Ports Recommendations, 1930
1930s	1931	Seaman's Welfare (Gene Convention), 1931
	1932	Port and Seaway (Gene) Convention, 1932
	1933	Contributions of Ship's Cooks Convention, 1933
	1934	Seal Service (London) Convention, 1934
	1935	Seaman's Welfare Convention, 1935
	1936	Port Visitation (London) Convention, 1936
	1937	Medical Examinations (London) Convention, 1937
	1938	Contributions of Ship's Stewards Convention, 1938
	1939	Accommodation of Crews Convention, 1939
	1940	Peace, Order of Trade and Shipping (Gene Convention), 1940
1940s	1941	Port Visitation (London) Convention, Revised, 1941
	1942	Accommodation of Crews Convention, Revised, 1942
	1943	Medical Examinations of Steam Power (Gene Convention), Revised, 1943
	1944	Labour Discipline (Geneva) Recommendations, Revised, 1944
	1945	Regulations of Officers (Geneva) Convention, Revised, 1945
	1946	Seaman's Welfare (Gene) Convention, Revised, 1946
	1947	Seaman's Articles of Agreement (Gene) Convention, Revised, 1947
	1948	Seaman's Welfare (Gene) Convention, Revised, 1948
	1949	Seaman's Welfare (Gene) Convention, Revised, 1949
	1950	Seaman's Welfare (Gene) Convention, Revised, 1950
1950s	1951	Seaman's Welfare (Gene) Convention, Revised, 1951
	1952	Seaman's Welfare (Gene) Convention, Revised, 1952
	1953	Seaman's Welfare (Gene) Convention, Revised, 1953
	1954	Seaman's Welfare (Gene) Convention, Revised, 1954
	1955	Seaman's Welfare (Gene) Convention, Revised, 1955
	1956	Seaman's Welfare (Gene) Convention, Revised, 1956
	1957	Seaman's Welfare (Gene) Convention, Revised, 1957
	1958	Seaman's Welfare (Gene) Convention, Revised, 1958
	1959	Seaman's Welfare (Gene) Convention, Revised, 1959
	1960	Seaman's Welfare (Gene) Convention, Revised, 1960
1960s	1961	Seaman's Welfare (Gene) Convention, Revised, 1961
	1962	Seaman's Welfare (Gene) Convention, Revised, 1962
	1963	Seaman's Welfare (Gene) Convention, Revised, 1963
	1964	Seaman's Welfare (Gene) Convention, Revised, 1964
	1965	Seaman's Welfare (Gene) Convention, Revised, 1965
	1966	Seaman's Welfare (Gene) Convention, Revised, 1966
	1967	Seaman's Welfare (Gene) Convention, Revised, 1967
	1968	Seaman's Welfare (Gene) Convention, Revised, 1968
	1969	Seaman's Welfare (Gene) Convention, Revised, 1969
	1970	Seaman's Welfare (Gene) Convention, Revised, 1970
1970s	1971	Seaman's Welfare (Gene) Convention, Revised, 1971
	1972	Seaman's Welfare (Gene) Convention, Revised, 1972
	1973	Seaman's Welfare (Gene) Convention, Revised, 1973
	1974	Seaman's Welfare (Gene) Convention, Revised, 1974
	1975	Seaman's Welfare (Gene) Convention, Revised, 1975
	1976	Seaman's Welfare (Gene) Convention, Revised, 1976
	1977	Seaman's Welfare (Gene) Convention, Revised, 1977
	1978	Seaman's Welfare (Gene) Convention, Revised, 1978
	1979	Seaman's Welfare (Gene) Convention, Revised, 1979
	1980	Seaman's Welfare (Gene) Convention, Revised, 1980
1980s	1981	Seaman's Welfare (Gene) Convention, Revised, 1981
	1982	Seaman's Welfare (Gene) Convention, Revised, 1982
	1983	Seaman's Welfare (Gene) Convention, Revised, 1983
	1984	Seaman's Welfare (Gene) Convention, Revised, 1984
	1985	Seaman's Welfare (Gene) Convention, Revised, 1985
	1986	Seaman's Welfare (Gene) Convention, Revised, 1986
	1987	Seaman's Welfare (Gene) Convention, Revised, 1987
	1988	Seaman's Welfare (Gene) Convention, Revised, 1988
	1989	Seaman's Welfare (Gene) Convention, Revised, 1989
	1990	Seaman's Welfare (Gene) Convention, Revised, 1990
1990s	1991	Seaman's Welfare (Gene) Convention, Revised, 1991
	1992	Seaman's Welfare (Gene) Convention, Revised, 1992
	1993	Seaman's Welfare (Gene) Convention, Revised, 1993
	1994	Seaman's Welfare (Gene) Convention, Revised, 1994
	1995	Seaman's Welfare (Gene) Convention, Revised, 1995
	1996	Seaman's Welfare (Gene) Convention, Revised, 1996
	1997	Seaman's Welfare (Gene) Convention, Revised, 1997
	1998	Seaman's Welfare (Gene) Convention, Revised, 1998
	1999	Seaman's Welfare (Gene) Convention, Revised, 1999
	2000	Seaman's Welfare (Gene) Convention, Revised, 2000
2000s	2001	Seaman's Welfare (Gene) Convention, Revised, 2001
	2002	Seaman's Welfare (Gene) Convention, Revised, 2002
	2003	Seaman's Welfare (Gene) Convention, Revised, 2003
	2004	Seaman's Welfare (Gene) Convention, Revised, 2004
	2005	Seaman's Welfare (Gene) Convention, Revised, 2005
	2006	Seaman's Welfare (Gene) Convention, Revised, 2006
	2007	Seaman's Welfare (Gene) Convention, Revised, 2007
	2008	Seaman's Welfare (Gene) Convention, Revised, 2008
	2009	Seaman's Welfare (Gene) Convention, Revised, 2009
	2010	Seaman's Welfare (Gene) Convention, Revised, 2010

Two corpus-analysis programs for general English and maritime English data

For general English data analysis, we used the web program of COHA and NOW available at <https://www.english-corpora.org/coha> from 26 September, 2022 to 26 December, 2022 for frequency of the data and extract concordances of gender-specific lexical terms for the position of seafarers.



Figure 1. A Screen Shot of a COHA web program

For maritime English data analysis, **Wordsmith Tools 8.0** was used to make a wordlist for frequency of the data and to extract concordances of *seamen* and *seafarers*, and personal pronouns used in COMC.

Table 1 Basic information of the Corpus of Maritime Conventions (COMC)

Dataset	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s	2010s	Total
Texts	6	5	12	3	1	5	5	7	2	1	47
Tokens	7,240	10,396	33,853	6,890	9,113	54,295	11,745	61,986	60,461	118,726	374,705

Table 2 Basic information of the Corpus of Historical American English (COHA)

Data	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s	2010s
Tokens	25,700,	27,707,	27,399,	28,661,	29,122,	28,829,	29,851,	33,149,	34,821,	35,452,
	422	879	750	274	676	225	580	318	812	806
Genres	4	5	5	5	5	5	5	5	5	5

Results and Discussion

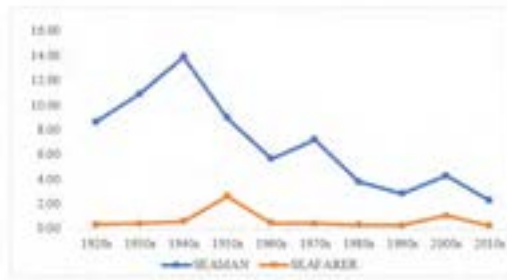


Figure 2 Frequencies of SEAMAN and SEAFARER in COHA (per million)

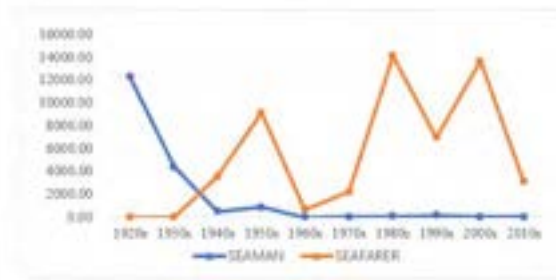


Figure 3 Frequencies of SEAMAN and SEAFARER in COMC (per million)

First, the use of the gender-exclusive generic form SEAMAN is still more popular than the gender-inclusive word SEAFARER in general English. In contrast, the gender-inclusive word SEAFARER has, to great extent, replaced the gender-exclusive form SEAMAN in maritime conventions, or more broadly in maritime legal texts. Second, in both general English and maritime conventions the use of SEAFARER had a great increase in 1950s but no increase for SEAMAN, which indicates that in this decade gender-inclusive words are more prevailing.

Table 5 Frequencies of *able seaman* and *able seafarer* in COMC and COHA

COMC	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s	2010s
able seaman	0	0	10	5	0	0	0	0	0	0
able seafarer	0	0	0	0	0	0	0	0	3	26
COHA	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s	2010s
able seaman	15	4	5	8	0	0	0	0	6	0
able seafarer	0	0	0	0	0	0	0	0	0	0

Table 5 shows two interesting points which can give strong evidence to the assumption mentioned above. One is that *able seafarer* only appears in the maritime English corpus, a specific English corpus whereas never appears in COHA, a general English corpus. The other point is that *able seafarer* only appears in 2000s and 2010s which has much to do with the influence of the official's propose of using *able seafarer* instead of *able seaman* in ILO (2005).

Table 6 Frequencies of personal pronouns in COMC

Type	1920s	30s	40s	50s	60s	70s	80s	90s	2000s	10s	
gendered pronouns	he	19	15	25	2	5	67	1	11	2	0
	him	1	7	43	8	2	21	13	1	0	0
	his	31	18	24	6	11	97	0	10	0	1
	she	0	0	1	0	0	0	0	0	0	0
	her	0	0	0	0	0	0	0	0	0	0
gender- inclusive pronouns	he or she	0	0	0	0	0	0	6	0	3	0
	him or her	0	0	0	0	0	0	0	1	2	0
	his or her	0	0	0	0	0	0	6	1	13	3

Table 6 shows that from 1920s to 1970s, all the personal pronouns are gendered forms and all are male personal pronouns, despite one exceptional instance *she* shown in the subcorpus of 1940s, which actually refers to the ship rather than people (the concordance of *she* is not listed). In terms of the gender-inclusive personal pronouns, they started to appear in 1980s and seemed to be kept to use since then.

Table 7 Context of the form *he* in 1980s to 2000s in COMC

No.	Context	Subcorpus	Gendered pronouns
1	Whenever a engineer is assigned for any reason to the territory of a Member, the competent authority should, if <i>he</i> is engaged, immediately inform...	1980s	-
2	That of the radio officer in radio-telephone operation, except as provided by the relevant Radio Regulations, for which <i>he</i> does not hold the appropriate certificate.	1980s	-
3	If such a person holds an appropriate certificate, <i>he</i> shall be required to give a test accepted by the Administration.	1980s	-
4	It should be remembered that it may take the master a few minutes to reach the bridge before <i>he</i> can take over the operation.	1980s	-
5	It may take the master... some discussion need to be taken before <i>he</i> reaches the bridge.	1980s	-
6	If any person working in a space deck or any area adjacent thereto should give a pre-arranged signal.	1980s	-
7	The person is enabled so that <i>he</i> can be located out.	1980s	-
8	If a person working in the space indicates that <i>he</i> is being affected by the atmosphere.	1980s	-
9	All occupational persons should be removed from the space as quickly as possible, unless <i>he</i> is specially trained.	1980s	-
10	Should the responsible officer notice the space above the sea surface, <i>he</i> should notify the bridge and inform.	1980s	-
11	The bridge duty officer who should arrange to check on the well-being frequently and at specific times as long as <i>he</i> remains awake.	1980s	-
12	Some aircraft should have a clear view of the runway deck. They should use the deck from a uniform altitude.	1980s	-
13	All language activities... the same, however, that cannot printing is very important.	2000s	-
14	All... (English... also indicated... the circumstances that the female pronouns refer to be...	2000s	-

For one thing, it reveals that instances of *he* in subcorpus of 1980s and 1990s are still pertaining to the generic usage with only two instances in 2000s used to refer to specific male persons, which to some degree suggests that the gender-inclusive pronouns has not gained a completely established position until the 2000s.

For another, it reveals some interesting results on how gender-inclusive occupational nouns are connected with pronouns. The first, second, eleventh, and twelfth concordance lines reveal that although the occupational nouns are gender-inclusive lexical terms (like *seafarer*, *radio officer*), the denotational pronouns are gender-biased *he*, even when the nouns are plural forms, like in the twelfth concordance line (plural *drivers* denoted by singular *he*).

From a sociolinguistic perspective, it indicates that gendered personal pronouns seem not to have gained as much attention as gendered occupational nouns so that gender-inclusive personal pronouns were adopted much later than gender-inclusive occupational nouns.

Concordance list from the 1926 version:

1. Any *seaman* who is landed during the term of his engagement or on its expiration shall be entitled to be taken back to his own country, or to the port at which he was engaged...
2. A *seaman* shall be deemed to have been duly repatriated if he has been provided with suitable employment...
3. A *seaman* shall be deemed to have been repatriated if he is landed in the country to which he belongs...
4. The expenses of repatriation shall not be a charge on the *seaman* if he has been left behind by reason of...
5. When a *seaman* is repatriated as member of a crew, he shall be entitled to remuneration for work done during the voyage.

Concordance list from the 1987 version:

6. The *seafarer* shall have the right to choose from among the prescribed destinations the place to which he or she is to be repatriated.
7. ...the moment the *seafarer* leaves the ship until he or she reaches the repatriation destination.
8. ...pay and allowances from the moment he or she leaves the ship until he or she reaches the repatriation destination.
9. the *seafarer* is to be repatriated or the State of which he or she is a national may arrange for his or her repatriation.
10. A *seafarer* shall be deemed to have been duly repatriated when he or she is landed at a destination.

We had another assumption, from the linguistic perspective, that the use of gender-biased generic use of **-man nouns** might strengthen the use of **male-biased** personal pronouns, because it will be very possibly followed by the gender-biased generic word *he*, meanwhile occupational **gender-inclusive nouns** might strengthen the use of **gender-inclusive pronouns** to maintain the consistency of both gender-inclusive styles and meanings.

Table 8 Frequency of gender-specific terms in NOW Corpus

gender-specific terms	seawoman/seawomen	male seafarer's	female seafarer's	man/men seafarer's	woman/women seafarer's
frequency	36	13	65	1	63

1. There's an immediate need for our *seamen and seawomen* to be trained and equipped with a standard protocol... (Philippines)

2. ...raised funds for a permanent memorial to the *seamen and seawomen* from the Dundee area... (Great Britain)

3. There was, in fact, a bar with the sign, "*Seamen and Seawomen* Welcome" but there was no mention of how compulsory the seawomen had to be... (Canada)

4. They handed skills to later generations of *seamen and seawomen*, and this role became the proof of their necessity... (Great Britain)

5. ...identifying all possible factors that contribute to the current gender disparity between *seamen and seawomen*... (Great Britain)

6. ...the lifestyle reflected in signs like '*Seamen and Seawomen*'... (Gambia)

From Table 8, it can be noted that, in terms of female references to this occupation, the frequency of *female seafarers* and *woman/women seafarers* is higher than that of *seawoman/seawomen*. The use of *male seafarers* is more frequent than *men seafarers*.

- Traditionally generic occupational words tend to reduce to gender-specific words.
- morphological rules
- natural, articulate or aesthetically pleasing

Conclusion

- Gender-exclusive (male-biased) word SEAMAN is still more prevailing in general English but in maritime English legal texts gender-inclusive word SEAFARER have taken the dominant place since 1940s. The occupation of women in maritime industry tends to have less influence on the adoption of gender-inclusive word SEAFARER compared with feminist language policies and guidelines in this field.
- Gender-inclusive personal pronouns take the form of *he or she* and *he/she* (*him or her* and *him/her*, *his or her*) in maritime English corpus (COMC). And it seems that attention on gender-inclusive adoption of pronouns is less paid than occupational nouns among maritime community. It is not until in 1980s that gender-inclusive pronouns began to appear and established its position in maritime legal texts around in 2000s.
- In terms of gender-specific words, although females are still gender-marked while males maintained unmarked positions, female seafarers have become visible in contemporary English with the compound style (female/women seafarers) more popularly adopted than the derivational form (seawoman/women).

Conclusion

Further assumptions and arguments

- Gender-inclusive policies affect people within the maritime community more greatly than those not.
- Gender-inclusive occupational nouns tend to habitually co-occur with gender-inclusive personal pronouns, while male-biased nouns (*-man*) seem to be grammatically related to male-biased pronouns (*he/him/his*).
- The success of certain newly-adopted gender-specific words seem to follow some morphological rules (e.g. regarding lexeme or phonology) to grammatically right or to keep aesthetic, which however, needs further exploration.

Thank you very much!

Questions are welcome!

28



A Hybrid Approach to Keywords Analysis of Aviation Accident Investigation Reports Corpus

16th February, 2023

Yang Qi, Homin Park , & Se-Eun Jhang

(Dalian Maritime University & Korea Maritime and Ocean University)



Contents

1. Introduction
 2. Literature Review on keywords analysis
 3. Data and methods
 4. Results and Discussion
 5. Conclusion
- References



1. Introduction

1.1 Purpose of the present study

1.2 Research Questions

1.3 Aviation Accident Investigation Reports



1.1 Purpose of the present study

- The purpose of the study is to compare three different ways of keyword lists of an Aviation Accident Reports Investigation Corpus (AAIRC) as a study corpus through three different approaches to keywords analysis : corpus frequency-based keywords analysis (Scott 2016), text dispersion based keywords analysis (Egbert and Biber 2019; Scott 2020) and Gries'(2021) hybrid approach.
- We focus on the cross-comparison of **aviation-related keywords** out of the top 100 keywords extracted by using three different methods and then provide **heuristic explanations** in order to investigate which method is the best.
- We also discuss **two/three dimensional diagrams** for keyness of words in AAIRC as a study corpus and the COCA Academic 2011-2012 corpus and a reference corpus as general English, as positive **evidence for a hybrid approach to keywords analysis as the best way.**



1.2 Research Questions

- (1) First, how do we extract a list of keywords by using a hybrid approach to keywords analysis proposed by Gries (2021)?
- (2) Second, what are significant differences between Scott's (2016) frequency based keywords analysis, Egbert and Biber's (2019) text dispersion based keywords analysis, and a hybrid approach of DKL-frequency and DKL-dispersion proposed by Gries (2021)?

5



1.3 Aviation Accident Investigation Reports

- Accident Reports are one of the main products of an National Transportation Safety Board (NTSB) investigation.
- NTSB is an independent U.S. government investigative agency responsible for civil transportation accident investigation.
- Reports provide details about the accident, analysis of the factual data, conclusions and the probable cause of the accident, and the related safety recommendations.
- Most reports focus on a single accident, though the NTSB also produces reports addressing issues common to a set of similar accidents.
- Accident Investigation Reports include Aviation Accident Reports, Hazardous Materials Accident Reports, Highway Accident Reports, Marine Accident Reports, Pipeline Accident Reports and Railroad Accident Reports.

6



2. Literature Review on Keywords Analysis

2.1 Traditional approach: Corpus Frequency Keywords

2.2 New Approach: Text Dispersion Keywords (Egbert and Biber 2019)

2.3 A hybrid Approach (Gries 2021)



2.1 Traditional Approach: Corpus Frequency Keywords

- Scott (2016, 2020) : a keyword is a very important word and the frequency of its first occurrence in a given text is very unusual, but it doesn't mean that its high frequency is abnormally high.
- Baker (2004) : the keywords themselves carry some important background information, or the implied information of the article, as well as the corresponding cultural and knowledge content.
- Scott (1997); Baker (2004); Scott & Tribble (2006) : as a quantitative perspective, keywords are those that have statistical significance in terms of the frequency of occurrence in the text or corpus compared with the standard set by the reference corpus.



2.2 New approach: Text Dispersion Keywords by Egbert and Biber (2019)

- According to Egbert and Biber (2019), corpus frequency keywords are frequently used in a corpus.
- However, they are not dispersed broadly through each text of the corpus. That is to say, such corpus frequency keywords cannot really reflect the characteristics of the discourse in this field, nor can they describe the characteristics of this corpus in a scientific and objective way.

9



2.2.1 Text dispersion keyness for keyword analysis

- Egbert & Biber (2019) use log-likelihood or G² to evaluate these figures arithmetic.
- This formula for G² is:

$$G^2 = 2 \sum_i O_i \ln \left(\frac{O_i}{E_i} \right)$$

	Corpus frequency keyness	Text dispersion keyness
Definition	Words that are statistically more frequent in a target corpus when compared with a reference corpus	Words that occur in statistically more texts in a target corpus when compared with a reference corpus
Variable	Frequency in corpora	Dispersion across texts
Formula	Log-likelihood O = observed word frequency E = expected word frequency	Log-likelihood O = observed word dispersion (number of texts) E = expected word dispersion (number of texts)
Requirements	– At least one text in target and reference corpora – Software (e.g., AntConc, WordSmith) or specialised program	– Many texts in target and reference corpora – Specialised program

10



2.2.2 Egbert and Biber (2019)

- The two goals of Egbert and Biber (2019) are to introduce text dispersion keyness and also tell us the importance of text dispersion.
- This method focuses on dispersion and generates key wordlist.
- There is **no** related word frequency, and this measure is a kind of range.
- Both the sizes and corpus parts are **not** considered in this method.

11



2.3 A hybrid approach (Gries 2021)

- Egbert and Biber's (2019) text dispersion analysis for keyword research can still be upgraded a little bit, and we can consider both frequency and text dispersion at the same time. Then we calculate this keyness value.

12

2.3.1 Overview of a hybrid approach to keywords analysis



- This study focuses on a hybrid approach to keywords analysis proposed by Gries (2021), providing **heuristic explanations** and top 100 keywords examples to show how we extract a keyword list by using this hybrid approach.
- A majority of the keywords analyses on educational and genre studies are a blend of the two ways. They are frequency lists and association statistics. Two steps are needed: First, we establish a frequency list of a study corpus and another frequency list of a reference corpus. Second, we can create a contingency table (2 plus 2 table) which is usually used in collocation statistics, as shown in Table 1,

Table 1. A Contingency Table

	target corpus	reference corpus	sum
target word	a	b	a+b
other words	c	d	c+d
sum	a+c	b+d	N

14

2.3.1 Overview of the hybrid approach to keywords analysis (Cont'd)



- The LLR log-likelihood ratio for this contingency table is computed accordingly.
- The expected frequencies and log-likelihood score are used in the followed equation,

$$(1) a_{expected} = \frac{(a+b) \times (a+c)}{N}$$

$$(2) LLR/G^2 = 2 \times \left(a \times \log \frac{a}{a_{expected}} + b \times \log \frac{b}{b_{expected}} + c \times \log \frac{c}{c_{expected}} + d \times \log \frac{d}{d_{expected}} \right)$$

14



2.3.2 DKL frequency: New frequency-based keyness measure by Gries (2021)

- A dispersion measure beyond range such as Deviation of Proportions (DP)
- Gries' (2021) first enhancement is a new keyness measure with frequency information.
- Gries' (2021) first method is less associated with frequency.
- This measure is called "Kullback-Leibler (KL) divergence".
- It is written as **DKL** (posterior/data/priority /theory)" (Gries 2021, p.14)

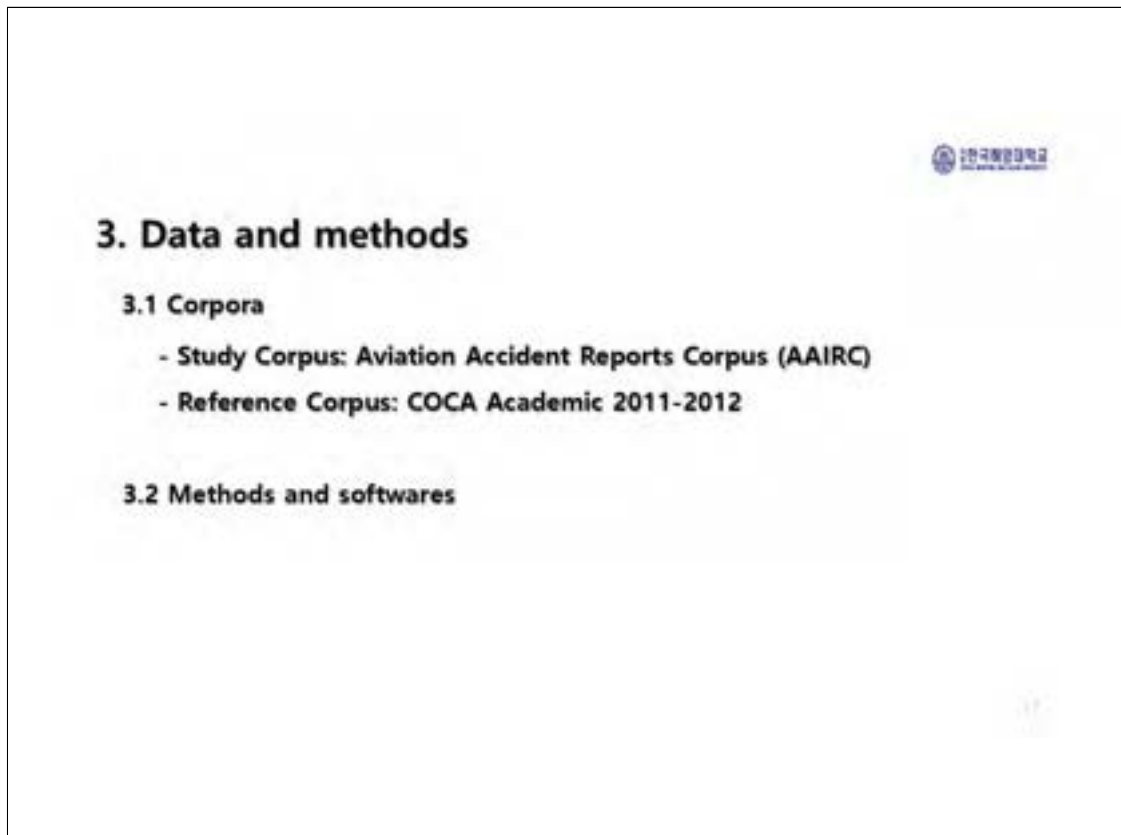
15



2.3.3 DKL dispersion : Application of dispersion-based keyness measure by Gries (2021)

- Gries' (2021) second enhancement is a new keyness measure with frequency information and dispersion information.
- Gries' (2021) second method is to add dispersion to the keyword analysis, so that we can try to solve a task of which method is the best out of several keyness measures.

16





3.2 Methods and softwares

- **WordSmith Tools 8.0** was used to extract corpus frequency keywords and text dispersion keywords.
- **Python** was used to process the data to get DKL frequency, DKL target corpus dispersion and DKL reference corpus dispersion for a hybrid keyness.
- DKL dispersion are calculated in both the AAIRC as a study (=target) corpus and COCA Academic 2011-2012 as a reference corpus.
- After reading the Book Corpus Data Processing using Python (with Professor Leilei and co-author Homin Park), we wrote some code that used Python programs to extract the corresponding keywords. First of all, in the aspect of data processing of corpora, we carried out sentence and word segmentation. Second, we carried out the magnetic princess, extraction of keywords, extraction of words fast, delete the stop words, and calculate the collocation strength of individual words.

79



4. Results and Discussion

- 4.1 Comparison of three keyword lists
- 4.2 Comparison of aviation-related keywords between K1, K2 and K3 keywords analyses
- 4.3. Gries' hybrid approach to keywords analysis

80



4.1 Comparison of three keyword lists

- **K1 keywords analysis:** Corpus Frequency (CF) keywords as a traditional analysis by not employing the text dispersion function in keywords setting of WordSmith Tools 8.0 (Scott 2016, 2020).
- **K2 keywords analysis:** Text Dispersion (TD) keywords as a new analysis by employing the text dispersion function in keywords setting of WordSmith Tools 8.0 (Scott 2020)
- **K3 keywords analysis:** keywords of DKL frequency and DKL dispersion as a hybrid approach to keywords analysis
- The top 100 keywords by using three different keywords analyses were compared
- There appeared 17 shared keywords.

21

Table 2. 17 shared keywords

No	SHARED KEYWORDS (IT)	K1 RANKING	K2 RANKING	K3 RANKING
1	accident	2	1	3
2	aircraft	6	2	2
3	airplane	1	6	6
4	airport	17	11	13
5	airspeed	26	29	13
6	aviation	19	8	7
7	cfr	45	12	16
8	cockpit	22	18	9
9	crew	10	3	17
10	crewmembers	100	26	22
11	crj	28	32	14
12	fax	7	1	1
13	flight	1	3	2
14	stab	5	1	1
15	pilot	3	7	10
16	press	9	13	9
17	takeoff	23	16	12

Table 3. 83 unshared keywords in K1

No.	RANKING	K1 KEYWORDS	No.	RANKING	K1 KEYWORDS	No.	RANKING	K1 KEYWORDS
1	3	safety	29	43	section	17	72	blade
2	11	helicopter	26	44	joining	18	73	swivel
3	12	captain	25	46	information	19	74	terrain
4	13	engine	32	47	inspection	49	75	pid
5	14	report	33	48	emergency	61	76	battery
6	15	operation	34	49	show	62	77	certification
7	16	passenger	23	50	require	63	79	company
8	18	train	26	51	right	64	80	door
9	20	air	27	52	grind	65	81	indicate
10	21	operator	28	53	fuel	66	82	peak
11	24	officer	29	54	checklist	67	83	foot
12	25	recommendation	40	55	heat	68	84	certificate
13	26	land	41	56	seat	69	85	door
14	27	weather	42	57	airline	70	86	board
15	28	airline	43	58	fan	71	87	issue
16	29	fly	44	59	part	72	88	operational
17	30	about	45	60	ice	73	89	fire
18	31	controller	46	61	leaves	74	90	condition
19	32	operate	47	62	record	75	91	descent
20	33	crates	48	63	attempt	76	92	requirement
21	34	altitude	49	64	position	77	93	sh
22	35	speed	50	65	aft	78	94	load
23	37	incident	51	66	before	79	95	ground
24	38	it	52	67	recorder	80	96	regulation
25	39	alert	53	68	structure	81	97	alarm
26	40	approach	54	69	guidance	82	98	personnel
27	41	station	55	70	stall	83	99	extra
28	42	thrust	56	71	wind			

22

Table 4. 83 unshared keywords in K2

No.	KAWING	K2 KEYWORDS	No.	KOCTN	K2 KEYWORDS	No.	KAWING	K2 KEYWORDS
1	14	admitt	39	40	fatal	37	74	thrust
2	15	operation	30	47	descent	38	75	manoeuvr
3	17	file	31	48	checklist	39	76	crewmember
4	18	operator	32	49	depart	40	77	ais
5	19	altitude	33	50	subje	41	78	officer
6	20	airaccident	34	51	incident	42	79	radar
7	21	operate	35	52	ts	43	80	departur
8	22	passenger	36	53	altitude	44	81	manufactur
9	23	air	37	54	ac	45	82	fatigue
10	24	engine	38	55	check	46	83	adviser
11	25	captain	39	56	reflight	47	84	helicopter
12	27	land	40	57	recorder	48	85	board
13	29	operational	41	59	airman	49	86	airworthiness
14	30	airline	42	58	accordion	50	87	visibility
15	31	certificate	43	60	inspection	51	88	carrier
16	33	manual	44	61	controller	52	89	maintenance
17	34	test	45	62	retrospective	53	90	surpass
18	35	alert	46	63	flag	54	91	probable
19	36	aft	47	64	certification	55	92	wing
20	37	weather	48	65	crash	56	93	ai
21	38	altitude	49	66	transportation	57	94	configuration
22	39	equip	50	67	agl	58	95	door
23	40	atr	51	68	landlight	59	96	rewind
24	41	regulation	52	69	terrain	60	97	descent
25	42	speed	53	70	structure	61	98	fuelage
26	43	emergency	54	71	atr	62	99	asst
27	44	personnel	55	72	airman	63	100	route
28	45	housing	56	73	climb			

Table 5. 83 unshared keywords in K3

No.	KAWING	K3 KEYWORDS	No.	KAWING	K3 KEYWORDS	No.	KAWING	K3 KEYWORDS
1	11	possibility	29	46	swain	57	74	evacuate
2	18	landlight	30	47	accr	58	75	handcraft
3	19	agl	31	48	altitude	59	76	deice
4	20	compass	32	49	stage engine	60	77	altitude
5	21	atr	33	50	proceed	61	78	uncontrolled
6	22	agl	34	51	firefighting	62	79	announcer
7	24	atr	35	52	volan	63	80	acc
8	25	reviewing	36	53	acc	64	81	inspect
9	26	acc	37	54	acc	65	82	ctrl
10	27	agl	38	55	acc	66	83	control
11	28	accident	39	56	accident	67	84	altitude
12	29	flyover	40	57	acc	68	85	open
13	30	dispatcher	41	58	ground-unacceptable	69	86	afk
14	31	flightpath	42	59	altitude	70	87	announcement
15	32	announcement	43	60	acc	71	88	announce
16	33	acc	44	61	altitude	72	89	afk
17	34	altitude	45	62	acc	73	90	altitude
18	35	altitude	46	63	acc	74	91	altitude
19	36	altitude	47	64	acc	75	92	altitude
20	37	acc	48	65	acc	76	93	altitude
21	38	acc-unacceptable	49	66	acc	77	94	acc
22	39	acc	50	67	acc	78	95	altitude
23	40	altitude-unacceptable	51	68	acc-unacceptable	79	96	altitude-critical
24	41	acc	52	69	altitude-critical	80	97	acc
25	42	altitude-unacceptable	53	70	altitude	81	98	altitude
26	43	altitude-critical	54	71	afk	82	99	altitude
27	44	acc	55	72	altitude	83	100	altitude
28	45	altitude-unacceptable	56	73	altitude			

4.2 Comparison of aviation-related keywords between K1, K2 and K3 keywords analyses



- More meaningful words are found in K3 of the present target corpus based on the observation below:
- The AAIRC usually contains information on the following details: Type of accident, accident location and time, airplane information (name of the airplane, official number, owner, nationality, damage, etc.), summary of the accident, pre-accident condition, the process of the accident, measures taken during the accident, results, recommendations and so on.



Table 6. A total of 100 keywords belonging to three categories of keywords, as follows (Lee and Jhang 2021):

1. Function words
2. Aviation-related words
3. Theme specific content words
4. General content words

No.	KEYWORDS CLASSIFICATION	SHARED KEYWORDS
1	Function words	0
2	Aviation-related words	10
3	Theme specific content words	4
4	General content words	3
	Total	17

Table 7. 17 shared keywords in three keywords analyses

NO	SHARED KEYWORDS (17)	K1 RANKING	K2 RANKING	K3 RANKING	KEYWORDS CLASSIFICATION
1	accident	2	4	5	General content words
2	aircraft	6	7	8	Aviation-related words
3	airline	8	9	9	Aviation-related words
4	airport	12	11	15	Aviation-related words
5	airport	28	28	22	Aviation-related words
6	aviation	29	9	7	Aviation-related words
7	air	40	22	26	Aviation-related words
8	weight	22	18	4	Aviation-related words
9	crew	28	9	17	Theme specific content words
10	crewmembers	108	26	23	Theme specific content words
11	air	28	22	18	Aviation-related words
12	air	28	22	18	Aviation-related words
13	flight	1	1	2	Aviation-related words
14	air	2	1	1	Theme specific content words
15	air	1	1	18	Aviation-related words
16	crisis	9	13	8	General content words
17	takeoff	22	18	12	General content words

➔ A heuristic explanation: Most of the keywords generated by K3 analysis have higher ranking than other analyses.

25



4.3. Gries' hybrid approach to keywords analysis

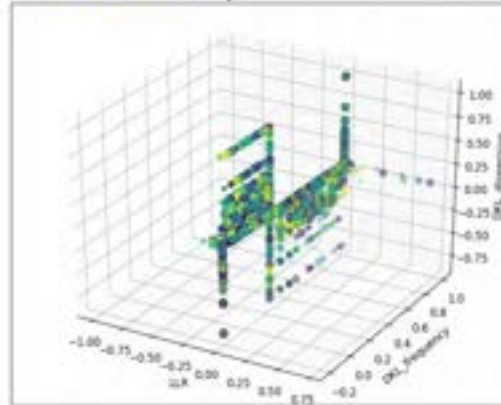
- To study a hybrid approach to keywords analysis mixed with frequency and dispersion, we investigated some plots in 2-dimensions and 3-dimensions.
- The x-axis represent the words' frequency or DKL-frequency or DKL-dispersion. For example, in case that values of (-1,1) appear on the plot, the range (-1,0) represents words whose frequency or DKL-frequency or DKL-dispersion distribution decides them to be AAIRC keywords.
- The more deviates, the stronger a word's frequency preference for either AAIRC or COCA academic. That is to say, the strongest AAIRC words in terms of frequency or DKL-frequency or DKL-dispersion will be far on the left. At the same time, the strongest COCA Academic words in terms of frequency or DKL-frequency or DKL-dispersion will be far on the right.

26

4.3.1 Three dimensional diagram illustrating DKL-frequency and DKL-dispersion

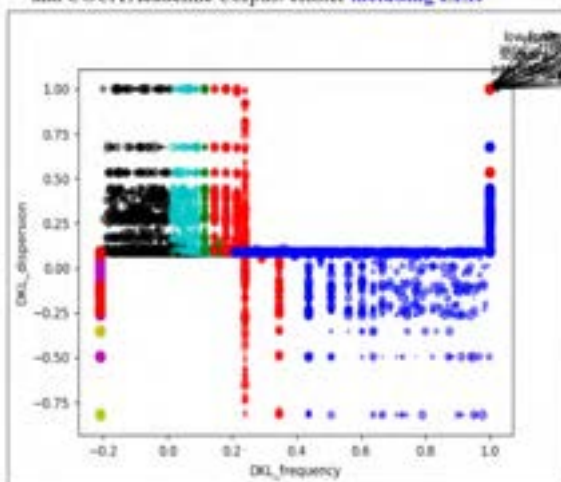
Three dimensional keyness of words diagram illustrates both the relationship between LLR, DKL-Frequency, and DKL-Dispersion, as shown in Figure 2.

Figure 2. Three dimensional keyness of words in AAIRC and COCA Academic Corpus



27

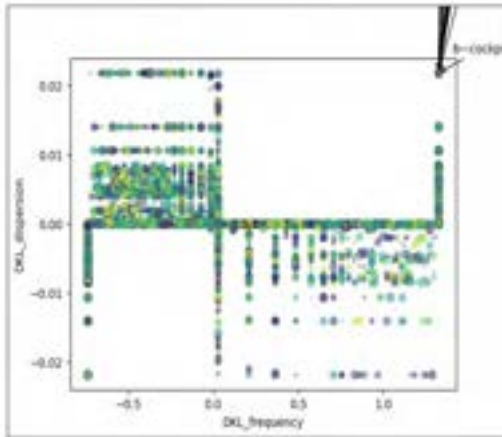
Figure 3. Two-dimensional keyness of words in AAIRC and COCA Academic Coepus: cluster including LLR



- This two-dimensional keyness of words diagram illustrates both the relationship between DKL-Frequency and DKL-Dispersion.
- The cluster-cluster diagram is actually a variant representation of a three-dimensional graph. The x axis is DKL-frequency, and the y axis is DKL-dispersion. This information is reflected in color. We cluster all the values of LLR into 7 categories, and then use a color for each category. So that we can clearly see that LLR is a category of data in which DKL-frequency and DKL-dispersion are not aggregated, which indicates that LLR is not a representative indicator for finding keywords and key keywords.
- So it's not very accurate to judge by LLR in keywords analysis, compared with DKL-frequency and DKL-dispersion. As is illustrated in Gries' examples in Clinton-Trump, just like the graph of clustering that Gries (2021) mentioned in his paper.

28

Figure 4. Two-dimensional keyness of words in AAIRC and COCA Academic Corpus-cluster not including I.I.R



- This diagram does not consider and add the information of LLR on the basis of cluster diagram, and it is separately taken out for LLR.
- The x-axis is DKL-frequency, and the y-axis is DKL-dispersion, and then we have a graph that looks like this.

29

Table 8. Part of the top N wordlist

Word	Value
seen—most	1.4142135623730951
wing—that	1.4142135623730951
wing—	1.4142135623730951
low-force	1.4142135623730951
gemfibrozil	1.4142135623730951
hydrochlorot	1.4142135623730951
overnights	1.4142135623730951
commutable	1.4142135623730951
company-related	1.4142135623730951
sabena	1.4142135623730951
midstate	1.4142135623730951
ellensburg	1.4142135623730951
unprompted	1.4142135623730951
male-dominated	1.4142135623730951
passenger-type	1.4142135623730951
stress-free	1.4142135623730951
medium-range	1.4142135623730951
turbopropellers	1.4142135623730951
six-bladed	1.4142135623730951
pre-stall	1.4142135623730951

- Top-n actually corresponds to the distance between all points in Figure 3 and the far point.
- The farther the distance is, the higher the key degree is, the more critical the word is.
- It is deduced that this word is a relatively key word in the AAIRC corpus. So, the distance between the left and right words from the origin is calculated here.
- That's the distance between all the points in the legend, and then we sort it from top to bottom, and then we open up the graph top-n, and we can see, on the left is the word, and on the right is the value of an index, and that's the sorting result as keywords.

30



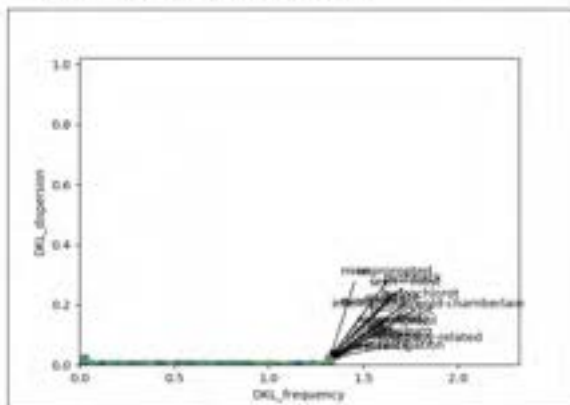
Table 9. Part of the top N-ed wordlist

WORD	VALUE
rule	1.4142135623730951
feet	1.4142135623730951
induction	1.4142135623730951
sterile	1.4142135623730951
approach-to-stall	1.4142135623730951
alpha	1.4142135623730951
atc	1.4142135623730951
arct	1.4142135623730951
map	1.4142135623730951
minneapolis-st	1.4142135623730951
nm	1.4142135623730951
go-around	1.4142135623730951
vsr	1.4142135623730951
ie	1.4142135623730951
eastern	1.4142135623730951
q-series	1.4142135623730951
c	1.4142135623730951
phase	1.4142135623730951
yawn	1.4142135623730951
sooner	1.4142135623730951

- Top-n-ed graph is just cut the upper right corner, the values greater than 0, and then sort a keyword. Then the corresponding result is the figure 4 diagram.
- It essentially calculates the order of all the words or points in Figure 4, one distance from the origin, and doesn't take into account the values that are less than zero, so the results are slightly different between the two.



Figure 5. Two-dimensional keyness of words in AAIRC and COCA Academic Corpus

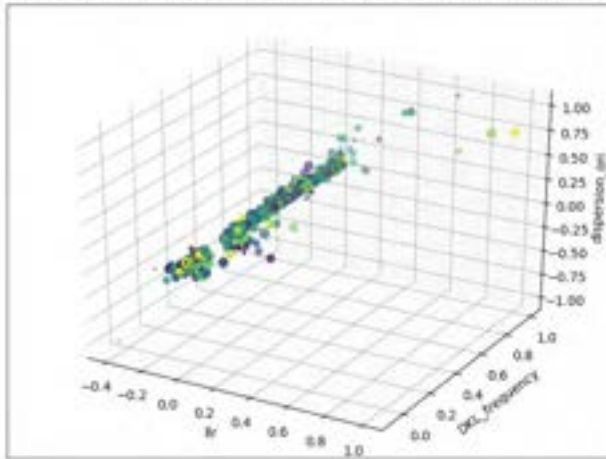


The ZOOM diagram in Figure 5 is a magnified version of the upper right corner of Figure 3, which is also the same as diagramed in Gries (2021), the upper right corner commonly used in his paper is enlarged.

4.3.2 Results and the best out of three keywords analyses



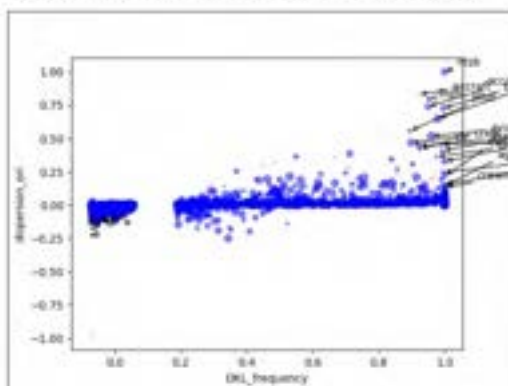
Figure 6. Three dimensional diagram of DKL_Frequency_Dispersion_Ori_LLRL



- The 3D_DKL_Frequency_Dispersion_Ori_LLRL diagram between -1 and 1 based on the five values we calculated with a normalization.
- The 3D has three dimensions, so we took three of the five results and drew this picture. The purpose is to show the relationship between them.
- DKL Frequency is one dimension, the dispersion is one dimension, and LLR is one dimension. Thus a three dimensional diagram was drawn.
- This is the result of sorting the words, using the combination of the two methods, Gries (2021) and Egbert and Biber (2019).
- It means that we need to calculate the first and second values separately, Then we calculate the square and the square root of the two values, and we can find the relationships in AAIRC, as shown in Figure 6.

33

Figure 7. 2D-Cluster_DKL_Frequency_Dispersion_Ori



- Cluster_DKL_Frequency_Dispersion_Ori is the cluster diagram. The *ordinate* is the dispersion value calculated by Egbert and Biber's (2019) method.
- The color is also divided into 7 classes according to the logarithmic likelihood ratio.
- This is a clustering diagram, which is also used by three parts. LLR, DKL-frequency (Gries 2021), dispersion Ori (Egbert and Biber 2019).
- LLR is used for clustering analysis.
- We found the left lower part and upper right part are blue dots, and they represent different colors.
- LLR cannot distinguish good and different words, and they are also unable to extract keywords.
- DKL-frequency is good to differentiate.
- In the dimension of DKL-frequency, the dots to the right are more important.

34





Table 10. A list of keywords in K3 analysis

File	Dispersion	url	top 17 list	記録本
文件名	編碼列	地址(C)	查看(N)	検索(P)
ntsb	1.4142135623730951			
flight	1.242092607865094			
aircraft	1.240999355294017			
faa	1.2338787180266004			
accident	1.2041809760603248			
airplane	1.1749074084486355			
aviation	1.0907256205526714			
cockpit	1.0880567793779312			
runway	1.0728666625099044			
pilot	1.0537530175456808			
postaccident	1.0523313389844002			
takeoff	1.0411841976478942			
airspeed	1.0249233573632979			
cvr	1.023767097568569			
airport	1.022450763469201			
cfr	1.0152243512745809			
crew	1.0131719734282223			

- The best way out of three keyword analyses is the combination of Egbert and Biber (2019) and Gries (2021).
- Gries (2021) proposes an improved version of Egbert and Biber (2019) that is DKL-dispersion, which has no better effect than Egbert's method on the AAIRC-COCA Academic dataset.
- Gries (2021) alone is not generally applicable to all the texts because of its limitations.
- As shown in Figure 10, top 17 words are obtained by using the combination of the two methods of Egbert and Biber (2019) and Gries (2021).
- It means that we need to calculate the first and second values separately. Then we calculate the square and the square root of the two values.





5. Conclusion

5.1 Summary

5.2 Pedagogical Implications





5. Conclusion

5.1 Summary

- We have investigated which keywords analysis can generate more appropriate keywords of AAIRC as a study corpus through **heuristic explanations** of multilateral comparison of top 100 keywords using three different approaches to keywords analysis as well as **two/three dimensional diagrams** for keyness of words in AAIRC and COCA Academic Corpus.
- The best way out of three keyword analyses is the combination of Egbert and Biber (2019) and Gries (2021), which is K3 analysis, because this analysis generated more significant and meaningful words in AAIRC such as type of accident, accident location and time, airplane information (name of the airplane, official number, owner, nationality, damage, etc.), summary of the accident, pre-accident condition, the process of the accident, measures taken during the accident, results, recommendations and so on.

17



5.2 Pedagogical Implications

- The appropriate use of English for aviation can save human lives and avoid accidents and incidents in an aviation field.
- Understanding important words of English for aviation can make L2 writers more efficient to improve ESP writing such as aviation accident investigation reports. To do this, one of the best guides for the teaching and learning of Aviation English writing in an aviation field is likely to provide L2 writers with a list of more appropriate keywords extracted from a corpus of aviation accident investigation reports.

18

References

- Baker, P. (2004). Querying keywords: Questions of difference, frequency, and sense in keyword analysis. *Journal of English Linguistics*, 3, 346-359.
- Egbert, J. and Biber, D. (2019). Incorporating text dispersion into keyword analyses. *Corpora*, 14(1), 77-104.
- Gries, S. T. (2021). A new approach to (key) keywords analysis: Using frequency, and now also dispersion. *Research in Corpus Linguistics*, 9(2), 1-33.
- Lee, S.H. and Jhang, S.E. (2021) Keyword analyses of the 19th century English journal log-book corpus: Vancouver and Broughton, a manuscript presented at 2021 KASELL Spring Conference on English Linguistics on 5th of June, 2021, hosted by The Korean Association for the Study of English Language and Linguistics.
- Scott, M. (2016). *WordSmith Tools version 7*, Stroud: Lexical Analysis Software.
- Scott, M. (1997). PC analysis of key words – and key key words. *System*, 25(2), 233-245.
- Scott, M. (2020). *WordSmith Tools version 8*, Stroud: Lexical Analysis Software.
- Scott, M. and Tribble, C. (2006). *Textual patterns, key words and corpus analysis in language education*. Amsterdam, The Netherlands: John Benjamins.



Thank You Very Much For Listening.

Lexical bundles Analyses across Maritime English Legal Genres

February 16, 2023

Guandong Zhang (Korea Maritime & Ocean University)

Contents

1. Introduction (purpose/research questions)
 2. Literature Review
 3. Data and Methodology
 4. Results and Discussion
 5. Conclusion
- References

1. Introduction

- The purpose of the current study
 - This paper attempts to identify frequently used lexical bundles across four legal genres of a self-built Maritime English Law Corpus, namely case law, legal document, legislation and academic law articles.
 - This paper also conducts a deep analysis by using the category **criteria** proposed by Biber's (2006) to analyze the structural and functional characteristics of 4-word lexical bundles.

Research Questions

- (1) What are the most frequent lexical bundles across four legal genres?**
→ Comparison of general statistics for 4-8 word bundles in four genres will be analyzed through WordSmith Tools 8.0 (Scott, 2020) across genres.
- (2) How do we differentiate one genre from the others? In other words, does each of the four maritime English legal genres have its own unique characteristics?**
→ Comparison of 4-word lexical bundles lists will be further analyzed to explore characteristics of bundles of each genre.

2. Literature Review

- Biber et al.(1999) first proposed the term 'lexical bundle' to describe frequently recurring sequences of words regardless of the nature of any links between them.
- Breeze (2013) investigated 4-word lexical bundles in four legal corpora: academic law, case law, legislation, and documents.
 - two million words are composed of each subcorpus containing about 500,000 words.
 - major differences and the roles between syntactic and discourse patterns of these bundles across the four legal genres.

2. Literature Review

- Csomay (2004) explored 4-word lexical bundles based on an academic article corpus, and he found that these lexical bundles can be categorized into noun phrases, prepositional phrases, "it+be+adjective" fragments and "verb+complement" fragments.
- Based on spoken data, Nesi & Basturkmen (2006) concluded that the frequent bundles encountered in university lectures mainly had a discourse-signalling function.



3. Data and Methodology

-Study corpus:

This study consists of about 1 million words from the area of commercial law in English, divided into 4 sub-corpus of approximately 250,000 words for each genre.

- Case law (admiralty and maritime cases, admiralty court opinion, law reports)
- Document (charter contracts)
- Legislation (IMO Conventions)
- Academic (maritime law journal papers)

1



General statistics of the two corpora

Corpora	Genres	Tokens	Types	Type/Token ratio (TTR)	Standardized TTR	STTR std. dev.
Study Corpus (Maritime English Law Corpus)	Case Law	252,311	10,990	4.57	34.02	64.71
	Document	253,728	4,872	1.95	29.36	69.63
	Legislation	250,018	6,494	2.70	28.38	70.88
	Academic	254,255	14,430	5.85	38.86	59.76
	Total	1,010,312	21,568	2.21	32.77	66.33

2

Information of the Maritime English Law Corpus



Corpora	Genres	Source of Data	Texts	Tokens	Types	Type/Token ratio (TTR)	Standardized TTR
Study Corpus (Maritime English Law Corpus)	Case Law	Maritime Cases	51	100,710	5,278	5.43	32.13
		Court Opinion	70	103,556	7,454	7.65	36.27
		Law Reports	25	48,045	4,098	8.91	33.85
	Document	Charter Contracts	125	253,728	4,872	1.95	29.36
		SOLAS	31	100,127	4,204	4.32	29.23
	Legislation	MARPOL	22	52,640	2,745	5.46	26.76
		STCW	18	33,809	2,022	6.22	31.48
		COLREG	7	12,131	1,315	11.18	25.53
		FAL	5	9,113	1,229	13.79	31.13
		LL	9	20,425	1,660	8.66	24.30
		SAR	4	6,641	949	14.98	26.48
		SUA	2	5,548	802	15.07	27.98
		CSC	4	6,418	952	15.34	28.17
	Academic	SFV	1	3,166	575	18.71	26.67
		Maritime Law Journal Papers	135	254,255	14,430	5.85	35.86
	Total	Maritime English Law	509	1,010,312	21,568	2.21	32.77

Methodology



- WordSmith Tools 8.0 (Scott, 2020)
 - (1) WordSmith Tools Version 8 was used to profile the four corpora in terms of types, tokens and type/token ratio.
 - (2) WordSmith Tools was then used to identify all the 4-8 bundles across four genres.
 - (3) 4-word lexical bundles in four genres were then categorized in terms of their structure and function.

4. Results and Discussion

- Number of different 4-8 word bundles in the four sub-corpora

	Case law	Documents	Legislation	Academic
4-word	101	542	269	70
5-word	29	460	147	17
6-word	11	422	81	6
7-word	2	407	55	3
8-word	0	396	41	2

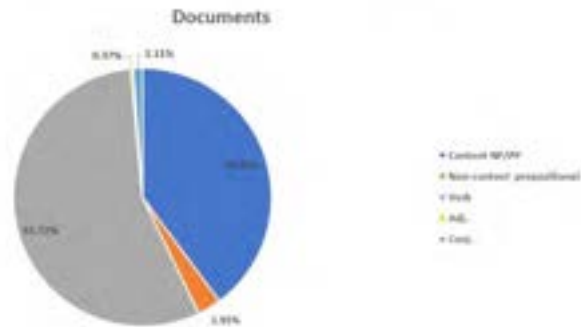
4.1 Categories and distribution of 4-word bundles

- Types of 4-word bundle in case law (N=101)



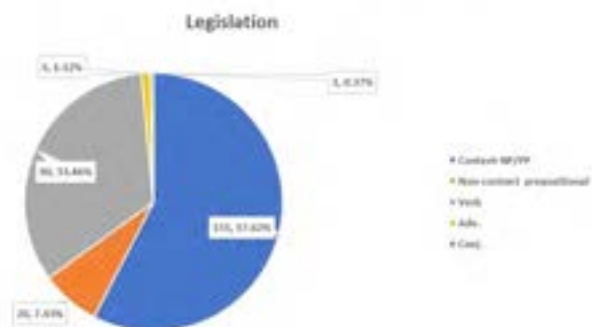
4.1 Categories and distribution of 4-word bundles

- Types of 4-word bundle in documents (N=542)



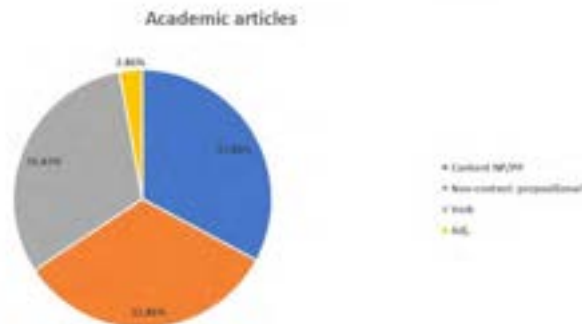
4.1 Categories and distribution of 4-word bundles

- Types of 4-word bundle in legislation (N=269)



4.1 Categories and distribution of 4-word bundles

- Types of 4-word bundle in academic articles (N=70)



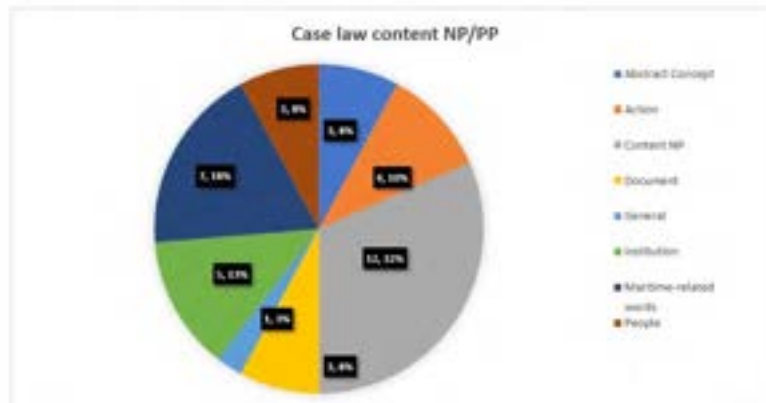
4.2 Categories and distribution of content noun phrase bundles

- Frequent content-related noun phrases in the four corpora

Case law	Documents	Legislation	Academics
<p>Institutions court of appeal in parliament and the council european parliament and the</p> <p>People course for your smart master of your smart teachers and teacher workers</p> <p>Documents teacher workers compensation act foreign affairs and security affairs and security policy</p> <p>Abstract concepts early and substantial action damage caused by the facts of this case</p> <p>Actions advice of the economic judgment of the supreme decision of the court</p> <p>Content prepositional phrase on the high seas beyond a marine vessel off the coast of</p> <p>Maritime-related bundles side of the channel delimited side of the shore of any state</p>	<p>Institutions maritime arbitrators association london maritime arbitrators association court of competent jurisdiction</p> <p>People arbitrator as sole arbitrator owners and the charterers party stated in law</p> <p>Documents this charter party the provisions of this clause laws of the place</p> <p>Abstract concepts dispute arising out of their obligations under this performance of this charter</p> <p>Actions requirement of any further ascertainment of a case receipt of such notice</p> <p>Content prepositional phrase under this charter party on board the vessel during the charter period</p> <p>Law-related bundles law and any dispute mediator's costs and expenses costs of the arbitration</p> <p>Dates when the arbitration proceedings time when the arbitration days of receipt of</p>	<p>Institutions authority of the government maritime safety committee of united nations for registration</p> <p>People secretary general of the members of the organization</p> <p>Documents international convention for the convention for the safety international regulations for preventing</p> <p>Abstract concepts satisfaction of the administration facts of the present opinion of the administration</p> <p>Actions verification acceptance approval or acceptance approval or acceptance</p> <p>Content prepositional phrase with the provisions of accordance with the provisions by the secretary general</p> <p>Maritime-related bundles prevention of pollution from side of the ship protection of the marine</p> <p>Dates date of entry into months after the date at least six months</p> <p>Numbers two thirds majority of one third of the</p>	<p>Documents liability for oil pollution and liability for oil law of the sea</p> <p>Abstract concepts incitation of liability for</p> <p>Content prepositional phrase under english law this on the marine environment for the implementation of</p> <p>Maritime-related bundles compensation for oil pollution protection of the marine sustainable use of marine</p>

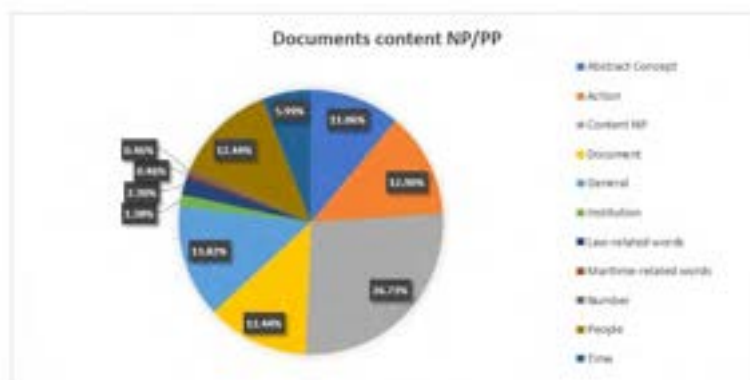
4.2 Categories and distribution of content noun phrase bundles

- Types of content phrase in case law (N=38)



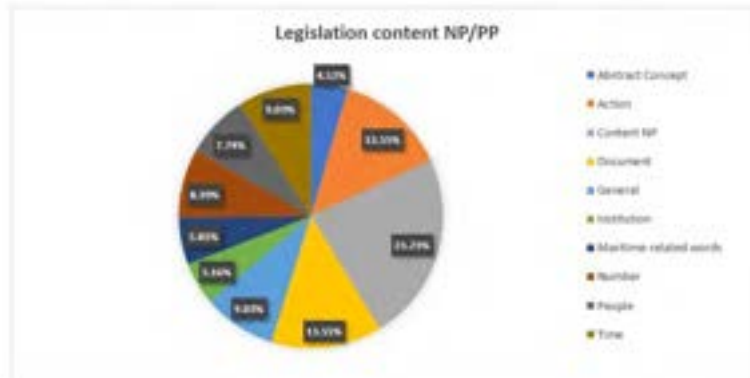
4.2 Categories and distribution of content noun phrase bundles

- Types of content phrase in legislation (N=217)



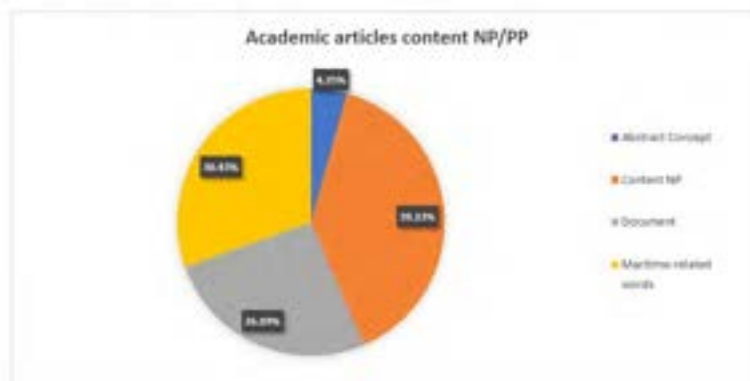
4.2 Categories and distribution of content noun phrase bundles

- Types of content phrase in documents (N=155)



4.2 Categories and distribution of content noun phrase bundles

- Types of content phrase in academic articles (N=23)



4.2 Categories and distribution of non-content bundles

- Frequency of non-content prepositional phrase bundles in the 4 sub-corpora

F/Case law	F/Documents	Bundles	F/Legislation	F/Academic
44	66	for the purposes of	37	13
36	0	as a matter of	0	10
35	115	for the purpose of	130	6
34	67	at the time of	38	12
30	90	as a result of	0	19
25	0	on the basis of	12	16
18	0	with regard to the	0	6
13	46	on behalf of the	0	0
12	0	with respect to the	12	11
12	51	on the part of	0	0
12	0	on the other hand	0	31
11	0	at the same time	11	25
11	0	as well as the	15	40
9	0	with a view to	15	6
9	53	for the benefit of	0	0
7	0	at the end of	20	11

4.2 Categories and distribution of non-content bundles

- Prepositional phrase bundles in the 3 sub-corpora

Case law	Documents	Legislation
by the supreme court	by the charterers to	by the secretary general
by counsel for shippers	by the charterers at	by the administration or
by the district court	by the parties and	by their respective governments
by counsel for ever	by the owners and	by the administration for
	by the other party	by the administration in
	by the two so	
	by the tribunal when	
	by each of the	
	by any court of	
	by the arbitration tribunal	

4.3 Categories and distribution of verb bundles

• Frequent verb bundles across 3 genres

Case law	Documents	Legislation	Academics
<p>Regulatives rendered the vessel unseaworthy be subject to the it is to be is the subject of ought to have been or ought to have it may be that</p> <p>Textual navigation having regard to the is consistent with the stance I do not consider supreme court held that court found that the not consider that the it is well established it seems to me I do not accept consider that the court held that a</p> <p>Emphatics it is true that it is necessary to it is useful to it is clear that</p>	<p>Regulatives arbitration shall be conducted charterers shall have the charterers shall indemnify the owners shall not be sole arbitrator shall be this clause shall apply owners shall provide the parties may agree the may agree the arbitration parties may agree at tribunal may take the tribunal may designate for party may advise the have the right to governed by and construed</p> <p>Textual navigation arising out of or conducted in accordance with including but not limited construed in accordance with notice that it has comply with the requirements notice to the other give notice that it except as otherwise provided</p> <p>Stance not be deemed to not agree to mediate does not agree to considers necessary to protect confirm that they agree</p>	<p>Regulatives this regulation shall be means shall be provided approval or accession shall arrangements shall be made provision shall be made administration may permit the period as may be may be necessary to may become parties to subject to the provisions is entitled to fly entitled to fly the ship is entitled to administration shall ensure that denunciation shall take effect</p> <p>Textual navigation comply with the requirements referred to in paragraph having regard to the hereinafter referred to as referred to as the complying with the provisions unless expressly provided otherwise issued under the provisions comply with the provisions</p> <p>Stance deemed to have been</p> <p>Emphatics it is necessary to it is possible to it is important to</p>	<p>Regulatives it should be noted that taking into account the can be used to will need to be</p> <p>Textual navigation was based on the is related to the is a need for can be found in is consistent with the is beyond the scope has the potential to</p> <p>Emphatics it is necessary to it is possible to it is important to</p>

4.3 Categories and distribution of verb bundles

• Frequent postmodifiers in verb bundles across 3 genres

Case law	Documents	Legislation
laid down in article	contained in this charter	approved by the administration
	agreed by the parties	adopted by the organization
	specified in the mediation	effected by the deposit
	set by the mediator	communicated by the secretary
	incurred in the mediation	adopted by a two
	paid for by the	

5. Conclusion

- This paper analyzed the types of 4-word lexical bundles across four legal genres. Each genre has its own characteristic in terms of the nature and function of lexical bundles.

(1) What are the most frequent lexical bundles across four legal genres?

- In terms of the length of lexical bundles, 4-word lexical bundles accounts for the largest proportion, and the number of bundles decreases as the length increases, which means that the two have a negative correlation.

25

(2) How do we differentiate one genre from the others? In other words, does each of the four maritime English legal genres have its own unique characteristics?

- In case law genre, it has same proportion between content noun bundles and verb bundles. Noun phrase bundles relating to institutions, people, documents and actions, as well as many noun prepositional phrases function as framing attributes.
- Document genre contains a lot of verb bundles, which have deontic or referential function.
- In legislation genre, content noun bundles account for the largest proportion, and bundles refer to adverb is an unique feature across 4genres.
- As for academic articles genre, content noun bundles, verb bundles as well as non-content prepositional bundles almost have the same proportion.

References

- Breeze, Ruth. 2013. Lexical bundles across four legal genres. *International Journal of Corpus Linguistics* 18(2), 229-253.
- Biber, D. 2006. *University Language: A Corpus-based Study*. Amsterdam/New York: John Benjamins.
- Bhatia, V.K. 1993. *Analysing Genre: Language Use in Professional Settings*. London: Longman.
- Biber, D. & Clark, V. 2002. "Historical shifts in modification patterns with complex noun phrase structures: How long can you go on without a verb?". In T. Fanego, M.J. López Couso & J. Pérez Guerra (Eds.), *English Historical Syntax and Morphology*. Amsterdam/New York: John Benjamins, 43-66.
- Biber, D. & Conrad, S. 1999. "Lexical bundles in conversation and academic prose". In H. Hasselgard & S. Oksefell (Eds.), *Out of Corpora: Studies in Honour of Stig Johansson*. Amsterdam: Rodopi, 181-189.
- Biber, D., Conrad, S. & Cortes, V. 2004. "If you look at...: Lexical bundles in university teaching and textbooks". *Applied Linguistics*, 25 (3), 371-405.
- Cortes, V. 2004. "Lexical bundles in published and student disciplinary writing: Examples from history and biology". *English for Specific Purposes*, 23 (4), 397-423.
- Nesi, H. & Basturkmen, H. 2006. "Lexical bundles and discourse signalling in academic lectures". *International Journal of Corpus Linguistics*, 11 (3), 283-304.
- Scott, Mike. 2020. WordSmith tools version 8: Lexical Analysis Software. Retrieved April 15, 2020, from <https://www.lexically.net/wordsmith/downloads/>
- Vázquez Orta, I. 2010. "A genre-based view of judgments of appellate courts in the common law system". In M. Gotti & C. Williams (Eds.), *Legal Discourse Across Languages and Cultures*. Bern: Peter Lang, 263-284.

27

THANK YOU VERY MUCH
for listening my presentation.



Any Questions?

28

코퍼스/응용

코퍼스 기반 한국수어의 기초 어휘 비교 연구

2023. 02. 16

장세은 • 박호민
(한국해양대학교)

한국언어과학회 겨울학술대회
부산대학교 온라인 ZOOM

Abstract and Key words

- The purpose of this present study is to discuss general vocabulary of Korean Sign Language (KSL) compiled by Chosun University in order to suggest desired vocabulary to be used in spoken language textbooks of Korean and English for the first grade to the ninth grade deaf students. To do this, first, we will make a word list tagged with with target parts-of-speech(POS) including nouns, verbs, adjectives, adverbs, and others. Second, we will illustrate the first to the 18th 100 signs by using **both** cumulative coverage **and** balanced numbers of each POS. The results show that the first 100 signs are the most frequent words accounting for around 41.03% coverage of all the **18,515 signs with 268,421 frequency** and the most frequent 1000 words account for 76.9%. The most frequent 1800 words account for around 83.84% and its frequency is 14 cutting off 1889 words. Third, this classification will be compared with the KSL online dictionary compiled by the National Institute of Korean Language. Finally, we will point out some problems to be considered in general vocabulary research for sign language when deaf students learn spoken languages such as Korean and English by using KSL as their native language. It is hoped that this KSL general vocabulary research can contribute to developing language textbooks for deaf elementary and middle school students.
- Key words: deaf, general vocabulary, Korean Sign Language, corpus, high frequency vocabulary, accumulated coverage, language textbooks, elementary school, middle school

Contents

1. Introduction (purpose/research questions)
 2. Literature Review
 3. Data and Methodology
 4. Results and Discussion
 5. Conclusion and More issues to be considered
- References

1. Introduction

- The **purpose** of the current study
to discuss general vocabulary of Korean Sign Language (KSL) compiled by Chosun University in order to suggest desired vocabulary to be used in spoken language textbooks of Korean and English for the first grade to the ninth grade deaf students.

1. Introduction (Cont'd)

▪ To do this,

(1) making a **word list tagged with target parts-of-speech(POS)** including nouns, verbs, adjectives, adverbs, and others.

(2) illustrating **the first to the 18th 100 signs** (100 signs x 9 grades x 2 semesters) for the nine grades from six-year elementary school to three-year middle school by using **both cumulative coverage and balanced numbers of each POS**. The results show that the first 100 signs are the most frequent words accounting for around 41.03% coverage of all the **18,515 signs with 268,421 frequency** and the most frequent 1000 words account for 76.9%. The most frequent 1800 words account for around 83.84% and its frequency is 14 cutting off 1889 words.

(3) making a comparison of the KSL online dictionary compiled by the National Institute of Korean Language.

(4) pointing out some problems to be considered in general vocabulary research for sign language when deaf students learn spoken languages such as Korean and English by using KSL as their native language.

5

1. Introduction (Cont'd)

▪ An Expected Effect

It is hoped that this KSL general vocabulary research can contribute to developing language textbooks for deaf elementary and middle school students.

▪ Research Questions

(1) How many words should comprise general vocabulary in Korean Sign Language?

(2) Which words are proper to create spoken language textbooks of Korean and English for the first grade to the ninth grade deaf students ?

6

2. Literature Review

Q: How many words should comprise general vocabulary in English as a spoken language?

- Nation (2001)

- A **traditional way** to categorize vocabulary has been as **general (i.e. high-frequency) vocabulary, academic vocabulary, technical vocabulary, and low-frequency vocabulary.**

- Using this categorization, general vocabulary has typically been set at around **2000 word families.** This is the number of word families in the original **General Service List (GSL)**, which is the foundation for the **Academic Word List (AWL)**, which remains the most commonly used list of academic words. While more recent general word lists appear to contain more words (the **NGSL** has 2801 words which comprise the most important high-frequency words in English), these are lemma-based lists, i.e. words and inflected forms, not word families. The **NGSL** has 2368 word families close to the 2000 word family range.

2. Literature Review (Cont'd)

Schmitt and Schmitt (2014)

- A more **recent way** to categorize vocabulary is as **high-frequency, mid-frequency, and low-frequency vocabulary.**

- This categorization expands the **high-frequency range** to the most frequent **3000 words**, with **mid-frequency** words representing words from **3001 to 9000** frequency levels, and **low-frequency** representing words from **9001 and lower.**

Key issues: Counting units

□ Types

- Each different word is counted as a type in a corpus.
- Types are suggested to word lists for beginners (spelling, recognizing words) (Carroll et al., 1971; Zeno et al. 1995).
- Word types are suggested in technical word lists (Nation, 2016).

□ Lemmas

- A lemma is a base word with part of speech and its form with inflectional suffixes.
- Lemmas (and above Level 2 word families) are suggested for receptive knowledge (reading and listening) (Nation, 2016).
 - ✓ Display the important information for vocabulary learning--the part of speech.
 - ✓ Allow learners to clearly focus on the vocabulary items in order to ease the burden. (Gardner & Davies, 2014)

□ Word families

- A word family is a base word with its inflectional forms and derivational forms.
- Bauer and Nation (1993) Word family levels.
- Learners have difficulties in derivational forms (Level 6 word families) (Schmitt & Zimmerman, 2002; Nagy & Tomsend, 2012; McLean, 2018).

Influential general and academic vocabulary lists

Vocabulary types	Vocabulary lists	Brief description
General vocabulary	General Service List, GSL (West, 1953)	2,000 word families
	New General Service List, New-GSL (Brezina & Gablasova, 2015)	2,494 lemmas
Academic vocabulary	Academic Word List, AWL (Coxhead, 2000)	570 word families
	Academic Vocabulary List, AVL (Gardner & Davis, 2014)	3,015 lemmas
	Academic Spoken Word List, ASWL (Dang et al., 2017)	1,741 word families

11

Lexical coverage

- One of the topics in second language acquisition that attract researchers' interests is **how many words indeed should be familiar with by the learners to successfully understand a text.**
- Based on the previous studies (e.g. Liu & Nation, 1985; Laufer, 1989; Laufer & Ravenhorst-Kalovski, 2010) and combing the results, Laufer and Ravenhorst-Kalovski (2010) suggested the two lexical coverage of 95% and 98% may represent the possible lower and upper boundaries, that is, **95% is acceptable for a minimal comprehension of texts, and 98% is preferred for an optimal comprehension.**

12

Some studies of vocabulary size needed for reasonable comprehension

Genres	95% coverage	98% coverage	Representative Studies
Novels and news	4,000	8,000-9,000	Nation (2006)
TV programs	4,000	8,000-9,000	Webb & Rodgers (2009a)
Movies	3,000	6,000	Webb & Rodgers (2009b)
Business textbooks	3,500	5,000	Hsu (2011)
Engineering textbooks	5,000	10,000	Hsu (2014)

33

A Case Study of Sign Language

JSL Colloquial Corpus



2011.04 ~ 2014 동안 연구를 위해 제작된 일본 수어 말뭉치

British Sign Language Corpus의 수어 제작자인 Adam Schembri 교수의 Corpus 구축 방법론을 조언받아 제작 (일본 국립 정보학 연구소)

일본 내 지역별, 연령별, 성별 수어 대화 영상을 촬영하여 ELAN 프로그램으로 Annotation을 진행하였음




34

[American Sign Language \(ASL\)](#)

[100+ First ASL signs - HandSpeak](#)
Learn the most used American Sign Language vocabulary for your ASL classes from ASL 101 beginners to 301 advanced and Deaf culture studies.

<https://www.handspeak.com/word/most-used/>



The screenshot shows a webpage titled "100+ First ASL words" with a dark blue background. It includes a list of 100 common ASL signs arranged in a grid. The signs are numbered 1 through 100. Below the list, there is a note: "The word list is subject to change over time, it doesn't have to be exactly 100, but it can be a bit more or less, too."

3. Data and Methodology

3.1 자료제공자 (Informants)

- 연령층 남녀별: 평균연령 49.61세
- 남 16명, 여 10명

나이	참여인원
20대	1
30대	2
40대	12
50대	6
60대	3
70대	1
80대	1
총 인원	26

- 지역별/ 주제별 통계

지역	세부	인원	지역별 참여 인원
전남	강진	1	23
	광양	2	
	광주	7	
	나주	1	
	남악	3	
	목포	3	
	무안	2	
	보성	1	
	순천	1	
	회순	2	
전북	원주	1	2
	전주	1	
서울/경기	인천	1	1

- 주제별 총 참여 인원

주제	참여인원
연간	18
살	24
식생활	18
의생활	14
주생활	14
사회생활	18
경제생활	14
교육	12
종교	12
문화	18
정치와 행정	14
자연	18
동식물	18
계년	12
국가별 및 지역	16
기타	18
누적인원	258

3. Data and Methodology (cont'd)

3.2 Data

3.2.1 The Original Plan of the Chosun KSL Corpus

- A small corpus of Chosun KSL was compiled from December 2020 to April 2022 and was transcribed from January 2021 to May 2022 by the Language Convergence Institute, Chosen University, Gwangju, South Korea, which has been supported by The Five-year Institute Project of Research Foundation of Korea in September 2020.

- Original Plan: It was firstly planned that the Chosun KSL Corpus would consist of **300,000 gloss words**.

3.2.2 Corpus

- The **study corpus** consists of 70 files containing **251,205** tokens and 17,652 types (after **retaining content words tagged with** target parts-of-speech(POS) including nouns, verbs, adjectives, adverbs, and others and deleting other POS from the second corpus) as in Table 3.

→ This study corpus was analyzed to categorize the first to 18th 100 words **for this current study**

Table 1. and Table 2. Top 20 and Bottom 10 words

Table 1. The First Corpus					Table 2. The Second Corpus						
#	Gloss	POS	Types	Types(20MF%)	#	Gloss	POS	Types	Types(20MF%)		
1	이	NP	3,691	2,630	2,630	1	이	NP	3,432	2,738	2,738
2	는	NP	4,448	1,621	4,322	2	가	NP	4,371	1,678	4,426
3	가	VA	3,995	1,294	5,724	4	고	VA	3,856	1,224	5,080
4	고	VA	3,475	1,215	6,611	5	을	VA	3,244	1,979	6,867
5	을	VV	3,388	1,989	8,128	6	와	VV	3,182	1,912	8,111
6	와	VA	2,912	1,820	9,143	7	로	VA	2,821	1,826	9,168
7	로	VA	2,696	1,936	10,277	8	에	NP	2,676	1,957	10,145
8	에	C	2,624	1,921	10,989	9	이	C	2,569	1,867	11,022
9	이	NP	2,602	1,914	11,911	10	아	NP	2,555	1,811	12,091
10	아	VA	2,571	1,912	12,808	11	고	VA	2,499	1,906	12,997
12	고	MAA	2,546	1,898	13,844	12	고	MAA	2,399	1,895	13,792
13	고	VA	2,714	1,785	14,449	13	고	VA	2,306	1,738	14,121
14	고	VN	1,981	1,886	15,126	14	고	VN	1,864	1,828	15,248
15	고	NP	1,771	1,828	15,711	15	고	NP	1,754	1,643	15,859
16	고	VV	1,610	1,840	16,351	16	고	NP	1,512	1,610	16,479
17	고	NP	1,579	1,752	16,866	17	고	VV	1,470	1,580	16,952
18	고	NP	1,367	1,549	17,416	18	고	VV	1,425	1,528	17,478
19	고	VV	1,498	1,524	17,944	19	고	C	1,426	1,486	17,980
20	고	NP	1,411	1,509	18,458	20	고	NP	1,378	1,508	18,504
21	고	NP	1,440	1,504	18,950	21	고	VV	1,374	1,501	19,025
2000	이	VV	1	1,000	19,999	1915	이	VV	1	1,000	19,997
2001	고	NP	1	1,000	19,997	1916	고	NP	1	1,000	19,991
2002	고	NP	1	1,000	19,997	1917	고	NP	1	1,000	19,994
2003	고	NP	1	1,000	19,997	1918	고	NP	1	1,000	19,995
2004	고	NP	1	1,000	19,998	1919	고	NP	1	1,000	19,996
2005	고	NP	1	1,000	19,998	1920	고	NP	1	1,000	19,998
2006	고	NP	1	1,000	19,998	1921	고	NP	1	1,000	19,999
2007	고	NP	1	1,000	19,999	1922	고	NP	1	1,000	19,999
2008	고	NP	1	1,000	19,999	1923	고	NP	1	1,000	19,999
2009	고	NP	1	1,000	19,999	1924	고	NP	1	1,000	19,999
2010	고	NP	1	1,000	19,999	1925	고	NP	1	1,000	19,999
2011	고	NP	1	1,000	19,999	1926	고	NP	1	1,000	19,999
2012	고	NP	1	1,000	19,999	1927	고	NP	1	1,000	19,999
2013	고	NP	1	1,000	19,999	1928	고	NP	1	1,000	19,999
2014	고	NP	1	1,000	19,999	1929	고	NP	1	1,000	19,999
2015	고	NP	1	1,000	19,999	1930	고	NP	1	1,000	19,999

251,205 tokens and 20,614 types

173,086 tokens and 19,183 types

Table 3. Top 20 and Bottom 10 words in the study corpus

Top 20 in The Study Corpus							Bottom 20 in The Study Corpus						
Ranking	Gloss	POI	Type	types%	tokens/DJM	한국수어사전 존재유무							
1	나	NP	2,452	2.762358	2,762,358	○							
2	저기	NP	4,571	1,7029219	4,479157741	○							
3	없다	VA	3,876	1,4216473	5,90080379	○	18496	없음어	EC	1	0.0003725	99.99968627	X
4	있다	VA	3,244	1,2085491	7,199354335	○	18497	없어계	EC	1	0.0003725	99.99105882	X
5	있다	VV	1,182	1,5954152	8,294805548	○	18498	주어	EC	1	0.0003725	99.99143137	X
6	없다	VA	2,820	1,0505884	5,345393989	○	18499	한천치(는)기계	EC	1	0.0003725	99.99180382	X
7	공	NNG	2,670	0,9122531	10,31774732	○	18500	종자	EC	1	0.0003725	99.99217647	X
8	목	IC	2,589	0,9570786	11,27482574	○	18501	자유종계	EC	1	0.0003725	99.99257157	X
9	내	NP	2,559	0,9537531	12,22873887	○	18502	마스.러	EC	1	0.0003725	99.99329412	X
10	있다	VA	2,459	0,9160982	13,14427709	○	18503	무어하게	EC	1	0.0003725	99.99386667	X
11	보다	VV	2,400	0,8540178	14,03803491	○	18504	대우요계	EC	1	0.0003725	99.99403921	X
12	없다	VA	2,004	0,7465884	14,78498329	○	18505	없나	EC	1	0.0003725	99.99441176	X
13	아니다	VCM	1,864	0,6944335	15,4794148	○	18506	어디나/특정해	EC	1	0.0003725	99.99515686	X
14	-면	EC	1,754	0,6534311	16,13289591	○	18507	아래편	EC	1	0.0003725	99.99552941	X
15	공	NNG	1,532	0,5707452	16,70381112	○	18508	아니라	EC	1	0.0003725	99.99627451	X
16	생각하다	VV	1,470	0,5476472	17,25128838	X	18509	충치계	EC	1	0.0003725	99.99703961	X
17	할하다	VV	1,425	0,5308825	17,78274074	X	18510	뒤고	EC	1	0.0003725	99.99774471	X
18	어떤	IC	1,416	0,52775295	18,33867026	X	18511	높어	EC	1	0.0003725	99.9985008	X
19	해	NNG	1,376	0,5126276	18,82229781	○	18512	그렇지	EC	1	0.0003725	99.99904235	X
20	할하다	VV	1,374	0,5118825	19,33498026	○	18513	괜찮아	EC	1	0.0003725	99.9992549	X
							18514	고개고마아래.그렇지	EC	1	0.0003725	99.99962745	X
							18515	-만아나라	EC	1	0.0003725	100	X

268,421 tokens and 18,515 types

3.2.3 전사 파일 리스트 (총 104 개: 34 개 & 70개)

출처/파일명	파일명	출처/파일명	파일명
21120	SAU_005.026_SAM_3_1.2	21920	SAU_005.026_SAM_3_3.1.1
21121	SAU_005.026_SAM_3_1.3	21921	SAU_005.026_SAM_3_3.2.2
21122	SAU_005.026_SAM_3_1.4	21922	SAU_005.026_SAM_3_3.1.1.1
21923	SAU_005.026_SAM_3_3.1.2	21923	SAP_04_211.026_SAM_3_3.1.2
21924	SAU_005.026_SAM_3_3.2	21924	SAP_04_211.026_SAM_3_3.1.3
21925	SAU_005.026_SAM_3_3.3	21925	SAP_04_211.026_SAM_3_3.2.1
21926	SAU_005.026_SAM_3_3.4	21926	SAP_04_211.026_SAM_3_3.1.1.1
21927	SAU_005.026_SAM_3_3.5	21927	SAP_04_211.026_SAM_3_3.2
21928	SAU_005.026_SAM_3_3.6	21928	SAP_04_211.026_SAM_3_3.1.2
21929	SAU_005.026_SAM_3_3.7	21929	SAP_04_211.026_SAM_3_3.1.3
21930	SAU_005.026_SAM_3_3.8	21930	SAP_04_211.026_SAM_3_3.2.1
21931	SAU_005.026_SAM_3_3.9	21931	SAP_04_211.026_SAM_3_3.1.1.1
21932	SAU_005.026_SAM_3_4.1	21932	SAP_04_211.026_SAM_3_3.1
21933	SAU_005.026_SAM_3_4.2	21933	SAP_04_211.026_SAM_3_3.1.2
21934	SAU_005.026_SAM_3_4.3	21934	SAP_04_211.026_SAM_3_3.1.3
21935	SAU_005.026_SAM_3_4.4	21935	SAP_04_211.026_SAM_3_3.2
21936	SAU_005.026_SAM_3_4.5	21936	SAP_04_211.026_SAM_3_3.1.1.1
21937	SAU_005.026_SAM_3_4.6	21937	SAP_04_211.026_SAM_3_3.1.2
21938	SAU_005.026_SAM_3_4.7	21938	SAP_04_211.026_SAM_3_3.1.3
21939	SAU_005.026_SAM_3_4.8	21939	SAP_04_211.026_SAM_3_3.2
21940	SAU_005.026_SAM_3_4.9	21940	SAP_04_211.026_SAM_3_3.1.1.1
21941	SAU_005.026_SAM_3_5.1	21941	SAP_04_211.026_SAM_3_3.1
21942	SAU_005.026_SAM_3_5.2	21942	SAP_04_211.026_SAM_3_3.1.2
21943	SAU_005.026_SAM_3_5.3	21943	SAP_04_211.026_SAM_3_3.1.3
21944	SAU_005.026_SAM_3_5.4	21944	SAP_04_211.026_SAM_3_3.2
21945	SAU_005.026_SAM_3_5.5	21945	SAP_04_211.026_SAM_3_3.1.1.1
21946	SAU_005.026_SAM_3_5.6	21946	SAP_04_211.026_SAM_3_3.1.2
21947	SAU_005.026_SAM_3_5.7	21947	SAP_04_211.026_SAM_3_3.1.3
21948	SAU_005.026_SAM_3_5.8	21948	SAP_04_211.026_SAM_3_3.2
21949	SAU_005.026_SAM_3_5.9	21949	SAP_04_211.026_SAM_3_3.1.1.1
21950	SAU_005.026_SAM_3_6.1	21950	SAP_04_211.026_SAM_3_3.1
21951	SAU_005.026_SAM_3_6.2	21951	SAP_04_211.026_SAM_3_3.1.2
21952	SAU_005.026_SAM_3_6.3	21952	SAP_04_211.026_SAM_3_3.1.3
21953	SAU_005.026_SAM_3_6.4	21953	SAP_04_211.026_SAM_3_3.2
21954	SAU_005.026_SAM_3_6.5	21954	SAP_04_211.026_SAM_3_3.1.1.1
21955	SAU_005.026_SAM_3_6.6	21955	SAP_04_211.026_SAM_3_3.1.2
21956	SAU_005.026_SAM_3_6.7	21956	SAP_04_211.026_SAM_3_3.1.3
21957	SAU_005.026_SAM_3_6.8	21957	SAP_04_211.026_SAM_3_3.2
21958	SAU_005.026_SAM_3_6.9	21958	SAP_04_211.026_SAM_3_3.1.1.1
21959	SAU_005.026_SAM_3_7.1	21959	SAP_04_211.026_SAM_3_3.1
21960	SAU_005.026_SAM_3_7.2	21960	SAP_04_211.026_SAM_3_3.1.2
21961	SAU_005.026_SAM_3_7.3	21961	SAP_04_211.026_SAM_3_3.1.3
21962	SAU_005.026_SAM_3_7.4	21962	SAP_04_211.026_SAM_3_3.2
21963	SAU_005.026_SAM_3_7.5	21963	SAP_04_211.026_SAM_3_3.1.1.1
21964	SAU_005.026_SAM_3_7.6	21964	SAP_04_211.026_SAM_3_3.1.2
21965	SAU_005.026_SAM_3_7.7	21965	SAP_04_211.026_SAM_3_3.1.3
21966	SAU_005.026_SAM_3_7.8	21966	SAP_04_211.026_SAM_3_3.2
21967	SAU_005.026_SAM_3_7.9	21967	SAP_04_211.026_SAM_3_3.1.1.1
21968	SAU_005.026_SAM_3_8.1	21968	SAP_04_211.026_SAM_3_3.1
21969	SAU_005.026_SAM_3_8.2	21969	SAP_04_211.026_SAM_3_3.1.2
21970	SAU_005.026_SAM_3_8.3	21970	SAP_04_211.026_SAM_3_3.1.3
21971	SAU_005.026_SAM_3_8.4	21971	SAP_04_211.026_SAM_3_3.2
21972	SAU_005.026_SAM_3_8.5	21972	SAP_04_211.026_SAM_3_3.1.1.1
21973	SAU_005.026_SAM_3_8.6	21973	SAP_04_211.026_SAM_3_3.1.2
21974	SAU_005.026_SAM_3_8.7	21974	SAP_04_211.026_SAM_3_3.1.3
21975	SAU_005.026_SAM_3_8.8	21975	SAP_04_211.026_SAM_3_3.2
21976	SAU_005.026_SAM_3_8.9	21976	SAP_04_211.026_SAM_3_3.1.1.1
21977	SAU_005.026_SAM_3_9.1	21977	SAP_04_211.026_SAM_3_3.1
21978	SAU_005.026_SAM_3_9.2	21978	SAP_04_211.026_SAM_3_3.1.2
21979	SAU_005.026_SAM_3_9.3	21979	SAP_04_211.026_SAM_3_3.1.3
21980	SAU_005.026_SAM_3_9.4	21980	SAP_04_211.026_SAM_3_3.2
21981	SAU_005.026_SAM_3_9.5	21981	SAP_04_211.026_SAM_3_3.1.1.1
21982	SAU_005.026_SAM_3_9.6	21982	SAP_04_211.026_SAM_3_3.1.2
21983	SAU_005.026_SAM_3_9.7	21983	SAP_04_211.026_SAM_3_3.1.3
21984	SAU_005.026_SAM_3_9.8	21984	SAP_04_211.026_SAM_3_3.2
21985	SAU_005.026_SAM_3_9.9	21985	SAP_04_211.026_SAM_3_3.1.1.1
21986	SAU_005.026_SAM_3_10.1	21986	SAP_04_211.026_SAM_3_3.1
21987	SAU_005.026_SAM_3_10.2	21987	SAP_04_211.026_SAM_3_3.1.2
21988	SAU_005.026_SAM_3_10.3	21988	SAP_04_211.026_SAM_3_3.1.3
21989	SAU_005.026_SAM_3_10.4	21989	SAP_04_211.026_SAM_3_3.2
21990	SAU_005.026_SAM_3_10.5	21990	SAP_04_211.026_SAM_3_3.1.1.1
21991	SAU_005.026_SAM_3_10.6	21991	SAP_04_211.026_SAM_3_3.1.2
21992	SAU_005.026_SAM_3_10.7	21992	SAP_04_211.026_SAM_3_3.1.3
21993	SAU_005.026_SAM_3_10.8	21993	SAP_04_211.026_SAM_3_3.2
21994	SAU_005.026_SAM_3_10.9	21994	SAP_04_211.026_SAM_3_3.1.1.1

220317	SC_BS_042,043_CW,PS_50,50_FF_3_1,2,3,4,5
220317	SC_BS_042,043_CW,PS_50,50_FF_3_6,7,8,9,10
220317	SC_BS_042,043_CW,PS_50,50_FF_3_11,12,13
220317	SC_BS_042,043_CW,PS_50,50_FF_3_14,15,16
220406	GY,GY_044,045_GY,YS_40,40_M,M_3_1,2,3,4,5
220406	GY,GY_044,045_GY,YS_40,40_M,M_3_6,7,8,9
220406	GY,GY_044,045_GY,YS_40,40_M,M_3_10,11,12
220406	GY,GY_044,045_GY,YS_40,40_M,M_3_13,14,15,16

The third refined corpus contains 70 files from **210413** to 220406 after the deletion of other 34 files (from the first file of 201228 to 34th files of 210322).

23

- Naming of Video + Annotation Files
Current staying region_ Informants number_ Birth place_ Age group_ Sex_ Video topics

WJ,GS_017,018_JJ,JJ_40,40_F,M_3-3,3-4,4-1,5-1

① ② ③ ④ ⑤ ⑥

① WJ : Wan-Ju(완주), GS : Gun-San(군산)

② 017, 018: Informants

③ JJ : Birth place (Jeon-Ju, 전주)

④ 40: forties year old

⑤ Female, Male: Sex

⑥ Topics(주제)

3-3 (Food Life: most favorite food that you recently ate),

3-4 (Food Life: any food that you want to eat in future),

4-1 (Clothing Life: Tell me how to purchase clothes and shoes),

5-1 (Housing Life: your preferred place and time to meet with you're a few)

*Currently 16 major categories , 88 subcategories, 199 video topics

24

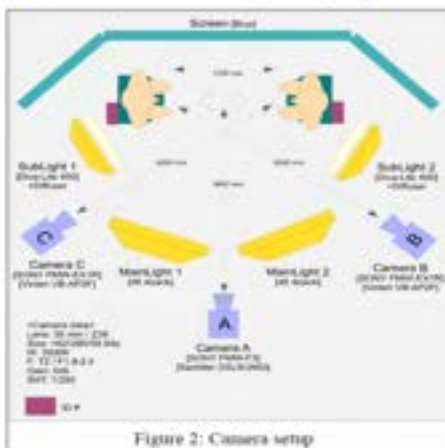
3.2.4 촬영 주제 분류표 (16 대범주, 88 소범주, 199 주제코드)

- Four types of video data collection (BSL + JSL)
 1. Conversations with another deaf person
 2. Answering interview questions
 3. Telling stories
 4. Lexical elicitation

25

영상수집 방법

· C (video recording of right person), A (overall), B (video recording of left person)



26

예:
 촬영일
 폴더명
 전사파일
 220404

대범주	개수
1 인간	7
2 삶	58
3 식생활	16
4 의생활	0
5 주생활	1
6 사회생활	0
7 경제생활	0
8 교육	0
9 종교	2
10 문화	0
11 정치와 행정	1
12 자연	1
13 동식물	0
14 개념	0
15 나라명 및 지명	0
16 기타	2

23

3.3 Methodology

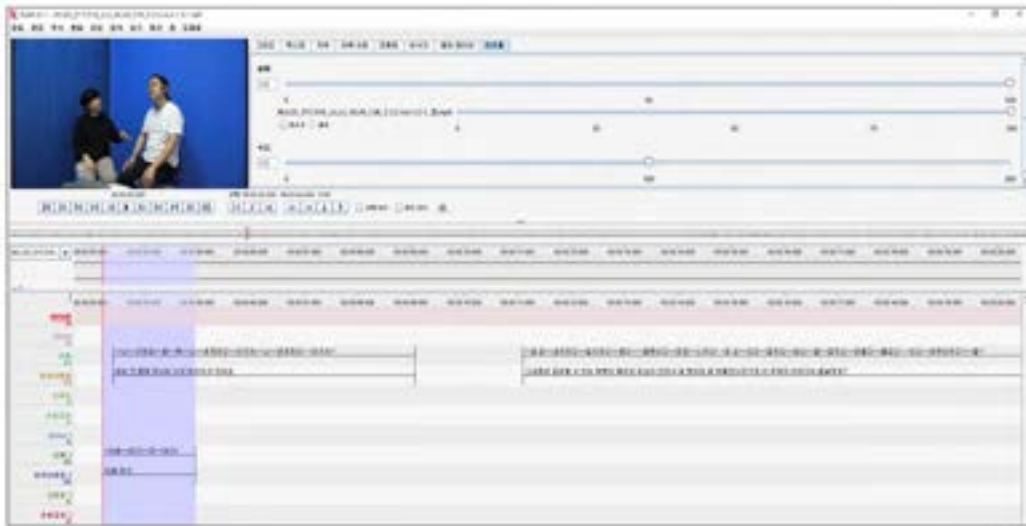
3.3.1 Description of Video Clips + ELAN

•전사와 분석방법

녹화한 동영상 자료를 ELAN (EUDICO Linguistic Annotator) 프로그램을 사용하여 정밀전사와 어구 주해 (Gloss) 분석으로 개별 어휘와 관용 어휘가 포함된 단어휘에 사용되는 어휘 비수지번호, 품사태깅 (세종말뭉치 품사태그셋) 주석 (annotation)으로 “조선대 KSL 코퍼스”구축

24

Video Clips + ELAN



29



30

3.3.2 세종 말뭉치 태그셋 (Sejong Tagset)

(A) 세종 말뭉치 태그셋 (Sejong Tagset)

태그명	표현	설명
100. 4문	4문 표시	4문 표시
	4문 대문자	4문 대문자
	4문 소문자	4문 소문자
	4문 숫자	4문 숫자
100. 5문	5문 표시	5문 표시
	5문 대문자	5문 대문자
	5문 소문자	5문 소문자
	5문 숫자	5문 숫자
100. 6문	6문 표시	6문 표시
	6문 대문자	6문 대문자
100. 7문	7문 표시	7문 표시
	7문 대문자	7문 대문자
100. 8문	8문 표시	8문 표시
	8문 대문자	8문 대문자
100. 9문	9문 표시	9문 표시
	9문 대문자	9문 대문자
100. 10문	10문 표시	10문 표시
	10문 대문자	10문 대문자
100. 11문	11문 표시	11문 표시
	11문 대문자	11문 대문자
100. 12문	12문 표시	12문 표시
	12문 대문자	12문 대문자
100. 13문	13문 표시	13문 표시
	13문 대문자	13문 대문자
100. 14문	14문 표시	14문 표시
	14문 대문자	14문 대문자
100. 15문	15문 표시	15문 표시
	15문 대문자	15문 대문자
100. 16문	16문 표시	16문 표시
	16문 대문자	16문 대문자
100. 17문	17문 표시	17문 표시
	17문 대문자	17문 대문자
100. 18문	18문 표시	18문 표시
	18문 대문자	18문 대문자
100. 19문	19문 표시	19문 표시
	19문 대문자	19문 대문자
100. 20문	20문 표시	20문 표시
	20문 대문자	20문 대문자
100. 21문	21문 표시	21문 표시
	21문 대문자	21문 대문자
100. 22문	22문 표시	22문 표시
	22문 대문자	22문 대문자
100. 23문	23문 표시	23문 표시
	23문 대문자	23문 대문자
100. 24문	24문 표시	24문 표시
	24문 대문자	24문 대문자
100. 25문	25문 표시	25문 표시
	25문 대문자	25문 대문자
100. 26문	26문 표시	26문 표시
	26문 대문자	26문 대문자
100. 27문	27문 표시	27문 표시
	27문 대문자	27문 대문자
100. 28문	28문 표시	28문 표시
	28문 대문자	28문 대문자
100. 29문	29문 표시	29문 표시
	29문 대문자	29문 대문자
100. 30문	30문 표시	30문 표시
	30문 대문자	30문 대문자
100. 31문	31문 표시	31문 표시
	31문 대문자	31문 대문자
100. 32문	32문 표시	32문 표시
	32문 대문자	32문 대문자
100. 33문	33문 표시	33문 표시
	33문 대문자	33문 대문자
100. 34문	34문 표시	34문 표시
	34문 대문자	34문 대문자
100. 35문	35문 표시	35문 표시
	35문 대문자	35문 대문자
100. 36문	36문 표시	36문 표시
	36문 대문자	36문 대문자
100. 37문	37문 표시	37문 표시
	37문 대문자	37문 대문자
100. 38문	38문 표시	38문 표시
	38문 대문자	38문 대문자
100. 39문	39문 표시	39문 표시
	39문 대문자	39문 대문자
100. 40문	40문 표시	40문 표시
	40문 대문자	40문 대문자
100. 41문	41문 표시	41문 표시
	41문 대문자	41문 대문자
100. 42문	42문 표시	42문 표시
	42문 대문자	42문 대문자
100. 43문	43문 표시	43문 표시
	43문 대문자	43문 대문자
100. 44문	44문 표시	44문 표시
	44문 대문자	44문 대문자
100. 45문	45문 표시	45문 표시
	45문 대문자	45문 대문자
100. 46문	46문 표시	46문 표시
	46문 대문자	46문 대문자
100. 47문	47문 표시	47문 표시
	47문 대문자	47문 대문자
100. 48문	48문 표시	48문 표시
	48문 대문자	48문 대문자
100. 49문	49문 표시	49문 표시
	49문 대문자	49문 대문자
100. 50문	50문 표시	50문 표시
	50문 대문자	50문 대문자
100. 51문	51문 표시	51문 표시
	51문 대문자	51문 대문자
100. 52문	52문 표시	52문 표시
	52문 대문자	52문 대문자
100. 53문	53문 표시	53문 표시
	53문 대문자	53문 대문자
100. 54문	54문 표시	54문 표시
	54문 대문자	54문 대문자
100. 55문	55문 표시	55문 표시
	55문 대문자	55문 대문자
100. 56문	56문 표시	56문 표시
	56문 대문자	56문 대문자
100. 57문	57문 표시	57문 표시
	57문 대문자	57문 대문자
100. 58문	58문 표시	58문 표시
	58문 대문자	58문 대문자
100. 59문	59문 표시	59문 표시
	59문 대문자	59문 대문자
100. 60문	60문 표시	60문 표시
	60문 대문자	60문 대문자
100. 61문	61문 표시	61문 표시
	61문 대문자	61문 대문자
100. 62문	62문 표시	62문 표시
	62문 대문자	62문 대문자
100. 63문	63문 표시	63문 표시
	63문 대문자	63문 대문자
100. 64문	64문 표시	64문 표시
	64문 대문자	64문 대문자
100. 65문	65문 표시	65문 표시
	65문 대문자	65문 대문자
100. 66문	66문 표시	66문 표시
	66문 대문자	66문 대문자
100. 67문	67문 표시	67문 표시
	67문 대문자	67문 대문자
100. 68문	68문 표시	68문 표시
	68문 대문자	68문 대문자
100. 69문	69문 표시	69문 표시
	69문 대문자	69문 대문자
100. 70문	70문 표시	70문 표시
	70문 대문자	70문 대문자
100. 71문	71문 표시	71문 표시
	71문 대문자	71문 대문자
100. 72문	72문 표시	72문 표시
	72문 대문자	72문 대문자
100. 73문	73문 표시	73문 표시
	73문 대문자	73문 대문자
100. 74문	74문 표시	74문 표시
	74문 대문자	74문 대문자
100. 75문	75문 표시	75문 표시
	75문 대문자	75문 대문자
100. 76문	76문 표시	76문 표시
	76문 대문자	76문 대문자
100. 77문	77문 표시	77문 표시
	77문 대문자	77문 대문자
100. 78문	78문 표시	78문 표시
	78문 대문자	78문 대문자
100. 79문	79문 표시	79문 표시
	79문 대문자	79문 대문자
100. 80문	80문 표시	80문 표시
	80문 대문자	80문 대문자
100. 81문	81문 표시	81문 표시
	81문 대문자	81문 대문자
100. 82문	82문 표시	82문 표시
	82문 대문자	82문 대문자
100. 83문	83문 표시	83문 표시
	83문 대문자	83문 대문자
100. 84문	84문 표시	84문 표시
	84문 대문자	84문 대문자
100. 85문	85문 표시	85문 표시
	85문 대문자	85문 대문자
100. 86문	86문 표시	86문 표시
	86문 대문자	86문 대문자
100. 87문	87문 표시	87문 표시
	87문 대문자	87문 대문자
100. 88문	88문 표시	88문 표시
	88문 대문자	88문 대문자
100. 89문	89문 표시	89문 표시
	89문 대문자	89문 대문자
100. 90문	90문 표시	90문 표시
	90문 대문자	90문 대문자
100. 91문	91문 표시	91문 표시
	91문 대문자	91문 대문자
100. 92문	92문 표시	92문 표시
	92문 대문자	92문 대문자
100. 93문	93문 표시	93문 표시
	93문 대문자	93문 대문자
100. 94문	94문 표시	94문 표시
	94문 대문자	94문 대문자
100. 95문	95문 표시	95문 표시
	95문 대문자	95문 대문자
100. 96문	96문 표시	96문 표시
	96문 대문자	96문 대문자
100. 97문	97문 표시	97문 표시
	97문 대문자	97문 대문자
100. 98문	98문 표시	98문 표시
	98문 대문자	98문 대문자
100. 99문	99문 표시	99문 표시
	99문 대문자	99문 대문자
100. 100문	100문 표시	100문 표시
	100문 대문자	100문 대문자

Target POS
한국수어에 맞는 (수준별) 어휘 선정을 위해 선별된 품사

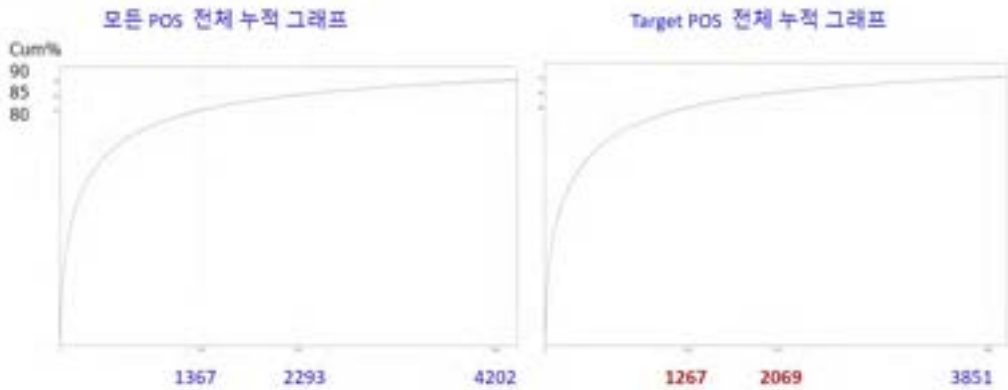
4. Results and Discussion

Cumulative %	80%	85%	90%	100%
All POS	1,367 / 219,127 (23)*	2,293 / 233,313 (11)	4,202 / 246,499 (5)	19,183 / 273,086
Target words	1,267 / 214,735 (24)	2,069 / 228,238 (12)	3,851 / 241,667 (5)	18,515 / 268,421

* Type number / Token number (cut-off minimum frequency)

전체 19,183개의 단어가 273,086번의 빈도수를 가지지만 **목표 단어**는 18,515개 단어로 268,421번 빈도수를 보여주어 **96.5%** 분포함

4. Results and Discussion (cont'd)



Top 30 어휘 목록

Ranking	Class	POS	type	types%	types(CUM%)	한국수어사전 존재유무
1	나	NP	2,452	2.762258	2.76225832	○
2	네가	NP	4,571	5.2029219	4.47915741	×
3	말다	VR	1,896	5.4276471	5.90680579	○
4	있다	VR	1,284	1.2095491	7.10935435	○
5	알다	VV	1,182	1.1854912	8.29484558	○
6	안다	VR	2,820	1.0593864	9.35423209	○
7	로	NGO	2,400	0.8922938	10.10794212	○
8	씩	C	2,369	0.9570796	11.27482518	○
9	이	NP	2,559	0.9513531	12.22617867	○
10	물다	VR	2,409	0.9160962	13.14227709	○
11	보다	VV	2,400	0.8941178	14.03639491	○
12	알다	VR	2,094	0.7863884	14.79898329	○
13	이(나)	VCN	1,884	0.6944115	15.4733948	○
14	..은	CC	1,754	0.6334015	16.10679631	○
15	집	NGO	1,532	0.5707912	16.70911112	○
16	생각하다	VV	1,470	0.5476472	17.25275838	×
17	알았다	VV	1,620	0.5388820	17.79264038	○
18	이해	C	1,476	0.5275295	18.39967026	×
19	해	NGO	1,376	0.5126276	18.92229791	○
20	알았다	VV	1,374	0.5119875	19.33428535	○
21	눈감	NGO	1,370	0.5103951	19.84567952	○
22	알았다	VV	1,364	0.4895295	20.33430903	×
23	무엇	NP	1,311	0.4884919	20.82251389	○
24	먹다	VV	1,290	0.4805861	21.30710222	○
25	현재	NGO	1,284	0.4789151	21.78141525	○
26	먹기	NP	1,238	0.4612758	22.24269103	○
27	해당하다	VCN	1,231	0.4609417	22.70313265	×
28	먹다	NP	1,220	0.4543099	23.15804215	×
29	먹다	VV	1,182	0.4403751	23.59841718	○
30	먹기	NP	1,168	0.4313099	24.01972707	○

인칭대명사: 나(7,452), 너(2,559), 당신(340), 그(6105), 우리(66), 그분(46), 그사람(6), 그녀(2) 10,582건의 빈도수를 가져 3.9%
 복원형대명사/지시대명사: 자기(4571), 여기(1,238), 저기(1,220), 저것(160), 여기(188), 이것(95), 그것(102), 그것을(5), 총 10,388건의 빈도수를 가져 3.8%를 차지함.
 → 기능어이지만 POS로는 내용어로 분류함
 의존사: 무엇(1,311), 방법(327), 왜(64), 누구(272), 언제(171), 어떻게(45), 누가(1) 등 POS에서는 내용어로 분류 중 3,406 빈도수 → 1.3%
 감탄사: 뭐(2,569), 어봐(1,416), -어(48), 아-(2), 오(18), 아하(74), 아(12), 글썽(11), 아참(8), 아몰라(5), 아(5) 등 95개의 type, 4,443 빈도수 → 1.7%
 [내용어]
 활용사: 갔다(3,816), 있다(3,244), 알다(2,820), 했다(2,459) 등 상위 6개 각각 3, 4, 6, 10위 차지하고 있고 전체 1,093개 type, 35,165 빈도수 → 12.1%
 동사: 알다(3,182), 보다(2,400), 생각하다(1,478), 알았다(1,425), 알았다(1,370), 안하다(1,314), 먹다(1,290), 거다(1,182) 등 전체 0.932개의 type, 30,651 빈도수 → 11%
 명사: 곳(2,610), 집(1,531), 해(1,376), 눈(1,370), 현재(1,284), 사람(1,159), 여자(96) 등 전체 6,306개의 type, 79,524 빈도수 → 29.8%
 부사: 일반적으로 → 무(85), 자(83), 조금(68) 등 전체 565개의 type, 11,649 빈도수 / 접속부사 → 그러나(1,008), 하지만(1,39), 그런데(74) 등 13개 type 1,779 빈도수 → 합 4.8%

Top 20 명사

일반명사			고유명사			의존명사			대명사			수사		
Ranking	Gloss	types	Gloss	types	Gloss	types	Gloss	types	Gloss	types	Gloss	types		
1	공	2610	한국	596	여러가지	604	나	1452	저나	988				
2	집	1532	말라	527	책	507	저기	4571	물	399				
3	책	1176	중	467	두번째	281	내	2559	첫째	328				
4	농민	1170	중국	243	-중	262	무엇	1311	-은	297				
5	현재	1284	서울	162	대로	238	대기	1238	올해	286				
6	사람	1159	미국	154	현실	119	저게	1220	영	96				
7	여자	969	부산	148	첫번째	115	이거	1560	생각	87				
8	방법	927	스	121	외	108	그것	1821	냉각	38				
9	영공역률	699	일본	156	달라	102	이것	957	냉	33				
10	군	664	북도	154	한번	95	저기	525	천	22				
11	친구	615	장	105	일년	95	달산	346	10만	22				
12	개	609	북문	104	1년	93	누구	272	100만	20				
13	남자	592	넷플릭스	94	10년	75	저것	560	2만	18				
14	대기	585	제주도	89	세번째	71	그거	529	3만	16				
15	나라	577	공주	85	물론	65	그야	505	8만	14				
16	청년	560	천주	83	2년	44	우리	66	5만	14				
17	문제	546	일	60	20년	44	12+	60	2억	12				
18	돈	535	재수	60	1위	44	12++	51	20만	12				
19	학교	531	영역	54	몇년	42	여기	48	어음전액	10				
20	선생님	530	가짜	51	13세	41	여기+	22	10만	10				

35

Top 20 품사

동사			형용사			일반부사			접속부사			감탄사		
Ranking	Gloss	types	Gloss	types	Gloss	types	Gloss	types	Gloss	types	Gloss	types		
1	있다	1182	같다	1816	도	887	그러나	1020	뭐	2581				
2	보다	2480	있다	1044	위반해	830	하지만	120	어제	1416				
3	생각하다	1470	같다	2820	조금	687	그런데	14	어제+	50				
4	일하다	1421	있다	2479	꽤	609	그러면	8	-은	48				
5	일하다	1374	있다	2004	저렴	604	다시/또	2	어제++	38				
6	일하다	1314	다르다	960	가끔	529	그러나+	2	어-	27				
7	하다	1290	일하다	950	대개	495	3.8~4.8	2	뭐?	23				
8	가다	1182	생겼다	791	항상	494	-은/그러면	2	어	18				
9	있다	981	종결다	623	모두	338	하지만(그)	5	어제+++	15				
10	일하다	834	가능하다	526	그럼	293	어제서	1	어제	14				
11	일하다	817	있다	494	바로	264	공책,그런	5	뭐++	18				
12	일하다	816	아깝다	363	그대로	217	그러나	5	어	12				
13	수행하다	814	적었다	250	계속	229			뭐+	11				
14	일하다	792	죽었다	327	일체	210			공책	11				
15	있다	791	불편하다	375	정간	196			공	10				
16	대화하다	786	일하다	311	대개	180			어	10				
17	있다	693	났다	299	죽었다	175			어디까지	8				
18	있다	642	계산하다	290	모래	171			가장	8				
19	하다	624	났다	271	연계	171			다들	8				
20	일하다	620	계속하다	246	죽	148			어디까지	7				

36

Top 20 품사

Ranking	긍정지시사		부정지시사		부사격조사		보조사		종결어미		연결어미	
	Gloss	types	Gloss	types	Gloss	types	Gloss	types	Gloss	types	Gloss	types
1	대응하다	1207	아니다	1064	대응해	176	부디	281	입니다?	17	-면	1754
2	재중하다	142	아니다*	22	강으로	378	-도	242	있다/주세요	14	어떻게	45
3	물안이다	106	아니다	10	안개	237	까지	152	아그렇구나	5	재중물론	17
4	전망하다	81	아니다...	7	부제	179	-부디	87	드려요	5	-까지	11
5	다행하다	64	아니다...	4	재한해	174	-까지	52	입니다	2	재중물론	6
6	보통하다	60	아니다*	4	재보다	161	도	39	오르려?	2	안녕하세요	6
7	재적하다	44	재중하다/재중신재하다	2	-보다	77	중/까지	18	가능할까?	2	영리	4
8	중속하다	21	재중하다	2	중보다	69	간	16	-입니다?	2	아니면	4
9	입니다	20	다중다/아니다	2	-중해	54	정거장/일부일하다	14	(중)계	2	영안/어떻게	4
10	물론하다	21	그재중하다...	2	보다+	17	부디...	12	알아채지않다	1	있다+	4
11	재중하다	20	-중간하다	2	-간해	26	중/안	11	질문/입니다?	1	없다/어떻게	3
12	간섭하다	19	없다/아니다	1	부제	21	죽도	11	물론/-입니다?	1	신속히/빠르게	3
13	형식하다	16	없다/아니다	1	-오해	21	아니면	10	중요하?	1	없다/중요하게	3
14	수제다	14	아니다	1	간해	20	수제하다	10	입니다	1	재도물론	2
15	없다	14	아니다+ag	1	그재해	20	오르려	9	입니다	1	오르려	2
16	재중하다	13	아니다+...+...	1	-간해	17	최중하다	8	중/안하	1	없게	2
17	재중재중하다	12	아니다(재중재)	1	보다+	16	지도/사해	6	아니하면	1	안가	2
18	재중하다	10	중재해/아니다	1	이적해	15	영안	6	아-그렇구나	1	재까지	2
19	노정하다	10	-중하다	1	-재중해	14	일부일하다	6	까지...	1	안녕하세요+	2
20	물안이다/안섭하다	8	-중하다	1	하중/강으로	11	어중하다	6	해라	1	물안이다	2

4. Results and Discussion (cont'd)

Cumulative %		EI 41%	84 13%
Target words		1,786 (15)*	1,889 (14)
명사	일반(NNG)	600	727
	고유(NNP)	26	39
	의존(NNE)	49	53
	대명사(ONP)	23	27
	수사(NR)	14	16
	소계	832	878
동사	VV	635	671
형용사	VA	162	189
부사	일반(MAG)	82	88
	접속(MAJ)	2	3
지형사	공형(NCP)	13	15
	부정(VCS)	3	3
접탄사	JC	9	11
격자	부사격조사	18	19
	보조사(CC)	8	9
	중속조사(IC)	1**	1
	종결어미(OP)	1***	2
	연결어미(CC)	1****	3

* Type number (not off maximum frequency)
 * ** <과(12), <과(6), <와(2), 나(고)막(2), 커(2) > *** <입니다(17), <알려주세요(14) > **** <면(1,74), <어떻게(45), <재중 물론(17)>

학년	명사	동사	형용사	부사/나머지	총계
1	94	72	20	14	200
2	94	72	20	14	200
3	94	72	20	14	200
4	94	72	20	14	200
5	94	72	20	14	200
6	94	72	20	14	200
7	94	72	20	14	200
8	94	72	20	14	200
9	94	72	20	14	200
소계	846	648	180	126	1,800

Q: **균등하게 배분? 또는 빈도수의 비율로 배분하기?**

수준별 어휘선정 숫자

POS	PREDI	SA	SP	NP	NP1	NP2	MA	MAG	VCP	VCN	VB	V	IC	IC	EC	EF	MM	TOTAL
1학년	88	3	11	8	39	4	1	23	1	1	2	2		1	1		1	109
2학년	101	3	1	7	16	4		36	1		1							168
3학년	101	4	1	5	16	5		3	1		1							133
4학년	119	3	3	9	11	1		4	1		1			1				143
5학년	88	1	3	5	37	6		1	1		1							144
6학년	87	3	3	11	36	7		8	1		1			1				156
7학년	87	3	3	6	16	1		10	1		4	1	1	1	1	1		125
8학년	91	3	1	11	36	1		8	1		1			1		1		134
9학년	80	1	1	7	36	1		8	1		1	1	1	1	1			120
TOTAL	887	23	31	68	211	34	1	87	17	1	19	3	3	6	4	2	6	1049

POS별 학년분포도 (누적비율로 재조정하였음)
 전체 POS → 14,026 vocabulary (어휘) 수, 269,586 단어수 (token 수) (약 200,000~300,000 단어로 구성된 한국수어교과서 구축을 목표
 → 목표 달성
 POS: 서술어(PRED), 수사(NP), 대명사(MP), 고유명사(NMP), 일반명사(NNG), 접속부사(MAJ), 일반부사(MAG), 긍정지시사(VCP), 부정지시사(VCN),
 부사격조사(MBB), 보조사(MB), 접속조사(JC), 감탄사(IC), 연결어미(ECL), 종결어미(EF), 관형사(MM)

★★고려해야하는 단어들: 한국어로는 다른 표현이지만 한국수어로는 같은 단어군)
 어휘목록: TOP 200 (1학년용), TOP 400 (2학년 용), TOP 600 (3학년용), TOP 800 (4학년 용),
 TOP 1000 (5학년용), TOP 1200 (6학년 용), TOP 1400 (7학년용), TOP 1600 (8학년 용), TOP 1800 (9학년용)

4. Results and Discussion (cont'd)

The First 100 Signs for 1st Grade

1. 명사 47개 (고병도 순) 94/2 = 47

나, 자기, 곧, 너, 집, 때, 농민, 무엇, 현재, 여기, 저거, 이거, 사람, 그것, 여자, 이것, 방법, 얼굴어름, 꽃, 친구, 개, 한국, 남자, 아가, 나라, 장영, 문제, 돈, 학교, 선생님, 거기, 꽃, 영화, 예전, 모양, 중, 시간, 법, 산, 협회, 수어, 취미, 옛날, 자, 얼굴, 아파트, 당신

※ 기능어: 지시대명사/인칭대명사/의문대명사/조동사(원로상 '곧', 진행형 '중') /지시대부사 /문법항태소 (부사? 명사? 현재사제 '현재/과거사제 '예전', '옛날', 접미사?? ('사람'))

※ 내용어: 개인적인 정보 ('집' '여자' '얼굴어름' '친구' '남자' '아가' '취미') /중요어는 것 ('개' '꽃' '영화' '산' /농사화 ('농민' '장마' '학교' '선생님' '협회' '수어') /경제어음('문제' '돈' '법' '차' '아파트', Y기타('사람' '한국' '나라' '모양' '시간'))

참고: 선생님 → 선생 + 님 (합성어 또는 파생어??)

고유명사: 한국, 중국, 미국, 일본, 서울, 부산, 목포, 광주, 전주, 독도, 제주도, 영어 등 어떤 단어를 First 100 Signs에 넣을까?
 의존명사: 여러가지 → 수어에서는 접미사/ ~중 → 진행조동사? (참고: 일반명사로 분류된 '곧' 은 수어에서는 완료조동사?)
 때문 (수어에서는 문사가 무엇일까?), 첫번째, 두번째, 세번째 등 [수사+때] 표현을 First 100 Signs에 넣을까?
 수사: 하나, 둘, 셋, 넷 ... 첫째, 둘째, 셋째, 넷째, 천, 만, 억 등 수사, 기수와 서수 및 숫자단위도 First 100 Signs에 넣을까?

4. Results and Discussion (cont'd)

The First 100 Signs for 1st Grade (cont'd)

2. 동사 36개 (고빈도 순) 72/2 =36

맞다, 보다, 생각하다, 말하다, 잘하다, 안하다, 먹다, 가다, 알다, 일하다, 원하다, 만들다, 수어하다, 못하다, 되다, 대화하다, 두다, 주다, 죽다, 모른다, 사용하다, 놀다, 느끼다, 만나다, 하다, 살다, 끝나다, 바꾸다, 여행하다, 사다, 가르치다, 모이다, 키우다, 경험하다, 받다

<유형별 동사구별>

- ※ 일치동사: 보다, 말하다, 수어하다, 대화하다, 주다, 만나다, 가르치다, 받다,
- ※ 공간동사: 가다, 여행하다
- ※ 일반동사: 맞다, 생각하다, 잘하다, 안하다, 먹다, 알다, 일하다, 원하다, 만들다, 못하다 등

<긍정과 부정 의미적 동사구분>

- ※ 긍정적인 의미를 갖는 동사: 맞다, 잘하다, 되다, 좋다
- ※ 부정적인 의미를 갖는 동사: 안하다, 못하다, 죽다, 모른다, 알다,

4. Results and Discussion (cont'd)

The First 100 Signs for 1st Grade (cont'd)

3. 형용사 10개 (고빈도 순)

같다, 있다, 없다, 좋다, 많다, 다르다, 필요하다, 괜찮다, 즐겁다, 가능하다

4. 부사/나머지 품사 7개 (고빈도 순)

뭐(2569), 아니다(1864), ~면(1754), 이봐(1416), 때문이다(1237), 그러나(1008), 또(887),

- ※ 기능어: 감탄사(뭐, 이봐), 접속사(~면, 그러나)
- ※ 내용어: 부정지시사(아니다), 긍정지시사 (때문이다), 부사(또)

4. Results and Discussion (cont'd)

The Second 100 Signs for 1st Grade

1. 명사 47 또는 48개 (고빈도 순)

전체, 걸국, 부모, 나이, 그림, 나무, 혼자, 물, 최고, 만약, 대부분, 학생, 이용, 바다, 가족, 누구, 오피, 저음, 교회, 병원, 보통, 경찰, 다음, 중국, 코로나, 진짜, 마음, 이유, 내용, 이들, 원래, 물, 부부, 생각, 입장, 휴대폰, 방식, 힘, 어머니, 이야기, 목적, 아들, 눈, 문장, 책, 술(175), 땀(174), 밤(174)

- ※ 가족관계 (부모, 가족, 부부, 어머니, 아들) / 대명사 (누구, 이들, 방식) 신체부분 (전체, 물, 눈)
- ※ 접속사: '만약'

4. Results and Discussion (cont'd)

The Second 100 Signs for 1st Grade (cont'd)

2. 동사 36개 (고빈도 순)

걸은하다, 돕다, 오다, 사라지다, 듣다, 사랑하다, 가지다, 더하다, 머무리하다, 가져오다, 쉬다, 잡다, 발전하다, 공부하다, 불나다, 다니다, 배우다, 계속하다, 배고하다, 합하다, 운동하다, 방문하다, 세우다, 버리다, 질문하다, 부르다, 구경하다, 떨어지다, 요리하다, 울다, 활동하다, 운전하다, 날다, 증가하다, 감사하다, 기다리다

<유형별 동사구별>

- ※ 일차동사: 걸은하다, 돕다, 가져오다, 잡다, 버리다, 질문하다, 부르다
- ※ 공간동사: 오다, 다니다, 방문하다
- ※ 일반동사: 그 외 동사들 모두

3. 형용사 10개 (고빈도 순) : 크다, 어렵다, 의미하다, 똑같다, 불편하다, 힘들다, 낮다, 깨끗하다, 싫다, 비슷하다

4. 부사/나머지 7개 (고빈도 순) :

자연히(830), 조금(687) 왜(669), 지금(604), 가장(529), 때문에(516), 이미(495),

- ※ 기능어: 의문사(왜), 부사격조사(때문에, 문법형태소 (부사) 명사) 현재시제(지금)
- ※ 내용어: 부사: '차연히', '조금' '가장' '이미'

부사/나머지 7개 Third 100 Signs for 2nd Grade 항상(494), 모두(338), 앞으로(318), 그냥(293), 부디(283), 서로(264), 그대로(251)
부사격조사(앞으로(318) → 문법형태소 미래시제:cf. 명사 미래(100)/, cf. 과거시제 부사격조사(전에(237)), cf. 후에(179))

4. Results and Discussion (cont'd)

국립국어원 한국수어사전과의 비교

- 한국어 대응표현
한국수어 lemma (표제어) 텍스트제공
→ 3,676개 (7,112 개 types)

- 한국수어 표제어 등록수 → 24,705개(??)

- 일련번호 1~ 20 예 (우측표)

번호	국어대안번호	한국수어표제어	수어 표제어 번호	WORD
1	2076	취미	1	1 취미
2	2076	취미	2	2 취미
3	2076	취미	3	3 취미
4	2076	취미	4	4 취미
5	2076	취미	5	5 취미
6	2076	취미	6	6 취미
7	2076	취미	7	7 취미
8	2076	취미	8	8 취미
9	2076	취미	9	9 취미
10	2076	취미	10	10 취미
11	2076	취미	11	11 취미
12	2076	취미	12	12 취미
13	2076	취미	13	13 취미
14	2076	취미	14	14 취미
15	2076	취미	15	15 취미
16	2076	취미	16	16 취미
17	2076	취미	17	17 취미
18	2076	취미	18	18 취미
19	2076	취미	19	19 취미
20	2076	취미	20	20 취미
21	2076	취미	21	21 취미
22	2076	취미	22	22 취미
23	2076	취미	23	23 취미
24	2076	취미	24	24 취미
25	2076	취미	25	25 취미
26	2076	취미	26	26 취미
27	2076	취미	27	27 취미
28	2076	취미	28	28 취미
29	2076	취미	29	29 취미
30	2076	취미	30	30 취미
31	2076	취미	31	31 취미
32	2076	취미	32	32 취미
33	2076	취미	33	33 취미
34	2076	취미	34	34 취미
35	2076	취미	35	35 취미
36	2076	취미	36	36 취미
37	2076	취미	37	37 취미
38	2076	취미	38	38 취미
39	2076	취미	39	39 취미
40	2076	취미	40	40 취미
41	2076	취미	41	41 취미
42	2076	취미	42	42 취미
43	2076	취미	43	43 취미
44	2076	취미	44	44 취미
45	2076	취미	45	45 취미
46	2076	취미	46	46 취미
47	2076	취미	47	47 취미
48	2076	취미	48	48 취미
49	2076	취미	49	49 취미
50	2076	취미	50	50 취미
51	2076	취미	51	51 취미
52	2076	취미	52	52 취미
53	2076	취미	53	53 취미
54	2076	취미	54	54 취미
55	2076	취미	55	55 취미
56	2076	취미	56	56 취미
57	2076	취미	57	57 취미
58	2076	취미	58	58 취미
59	2076	취미	59	59 취미
60	2076	취미	60	60 취미
61	2076	취미	61	61 취미
62	2076	취미	62	62 취미
63	2076	취미	63	63 취미
64	2076	취미	64	64 취미
65	2076	취미	65	65 취미
66	2076	취미	66	66 취미
67	2076	취미	67	67 취미
68	2076	취미	68	68 취미
69	2076	취미	69	69 취미
70	2076	취미	70	70 취미
71	2076	취미	71	71 취미
72	2076	취미	72	72 취미
73	2076	취미	73	73 취미
74	2076	취미	74	74 취미
75	2076	취미	75	75 취미
76	2076	취미	76	76 취미
77	2076	취미	77	77 취미
78	2076	취미	78	78 취미
79	2076	취미	79	79 취미
80	2076	취미	80	80 취미
81	2076	취미	81	81 취미
82	2076	취미	82	82 취미
83	2076	취미	83	83 취미
84	2076	취미	84	84 취미
85	2076	취미	85	85 취미
86	2076	취미	86	86 취미
87	2076	취미	87	87 취미
88	2076	취미	88	88 취미
89	2076	취미	89	89 취미
90	2076	취미	90	90 취미
91	2076	취미	91	91 취미
92	2076	취미	92	92 취미
93	2076	취미	93	93 취미
94	2076	취미	94	94 취미
95	2076	취미	95	95 취미
96	2076	취미	96	96 취미
97	2076	취미	97	97 취미
98	2076	취미	98	98 취미
99	2076	취미	99	99 취미
100	2076	취미	100	100 취미

45

4. Results and Discussion (cont'd)

조선대한국수어코퍼스 vs. 국립국어원한국수어사전과의 type 비교

	Chosen KSL corpus	KSL online dictionary compiled by the National Institute of Korean Language	Rate (%)
All POS	19,183	9,336	48.7
Target words	18,515	9,119	49.3
Content words (N, V, Adj, Adv)	17,652	9,003	51.0
Nouns	7,733	4,231	54.7
Verbs	8,073	3,541	43.9
Adjectives	1,142	768	67.3
Adverbs	704	436	61.9

아래의 경우를 고려한다면 비교 비율 같은 훨씬 더 높아짐

- (1) 동사 명사 비구별: (생각)하다, (일)하다, (수어)하다, (대화)하다, (사용)하다, (여행)하다, (경험)하다, (결혼)하다, (사랑)하다, (발전)하다, (공부)하다, (계속)하다, (비교)하다, (운동)하다, (질문)하다, (구경)하다, (요리)하다, (활동)하다, (운전)하다, (증가)하다, (감시)하다, (선택)하다 등 동사/명사를 구별하지 않고 같은 sign으로 본다면 위의 비교 비율은 훨씬 더 높아짐.
- (2) 자-자동차, 정간-건정연 등 유사어를 같은 sign으로 본
- (3) 앞으로 = 앞, 후에 = 후, 때문에 = 때문 등 후치사가 없는 명사와 같은 sign으로 본
- (4) 더하다 = 많하다, ㅇㅋ = OK = 오케이 등 의미상으로 같은 단어로 쓰이는 sign으로 본

46

4. Results and Discussion (cont'd)

조선대학국어수어코퍼스 vs. 국립국어원 한국수어사전과의 Top 200 signs 비교 (182/200 =91%)

The First 100 signs	Chosen KSL corpus	KSL online dictionary compiled by the National Institute of Korean Language	Rate (%)
명사	47	43 (저기, 저기, 앞줄어플, (인)형인, 인성님, 거기, (승호)기)	91.5
동사	36	28 (생각)하다, 안)다, 할)다, (수)어)다, (대)회)하다, (사)물)하다, (의)행)하다, (경)험)하다	77.8
형용사	10	9 (말)다)하다, 졸)다, (가)능)하다	90.0
부사/나머지	7	7	100
The Second 100 signs	Chosen KSL corpus	KSL online dictionary compiled by the National Institute of Korean Language	Rate (%)
명사	47	45 (교)사, 어)름)	91.5
동사	36	33 (결)혼)하다, (사)람)하다, 타)다 (= 할)다 73, 마)무)리)하다, (보)안)하다, (공)부)하다, 불)나)다, (계)속)하다, (미)고)하다, (문)물)하다, (질)문)하다, (구)경)하다, (표)지)하다, (발)동)하다, (운)전)하다, (종)가)하다, (감)시)하다	91.7
형용사	10	10 (피)리)하다, (물)권)하다	100
부사/나머지	7	7 (때)문)에	100

5. Conclusion and More issues to be considered

5.1 Conclusion: A summary of answers to the two research questions

(1) How many words should comprise general vocabulary in Korean Sign Language?

→ 전체 19,183개의 단어가 273,086번의 빈도수를 가지지만 **목표 단어**는 18,515개 단어로 268,421번 빈도수를 보여주어 **96.5%** 분포하고 있는 것으로 보이나, 갑탄사 1.7%, ‘-면’, ‘(-)때문(에)’ 등 접속사로 사용되는 단어가 1.1%, POS에서 내용어로 분류된 인정대명사(3.9%)와 비인칭대명사와 지시대부사(3.8%)와 의문사 (1.3%)를 기능어로 재분류하면 실질적으로 기능어는 **10.1% 이상**이고 **89.9% 정도가 내용어**임.

(2) Which words are proper to create spoken language textbooks of Korean and English for the first grade to the ninth grade deaf students ?

→ 초중등 농학생들에게 자신의 모국어인 한국수어로 한국어와 영어 등 외국어를 배우는 경우에는 학년에 맞는 한국수어어휘로 외국어언어교재를 만드는 것이 필요함 그래서 누적 약 **84%를 차지하는 1,800여개의 단어**를 9개 학년으로 나누면 학기당 100개의 어휘를 학습하는 것이 적절함. 여기서 학년별 200개의 어휘로 **균등하게 배분**하였으나 **용사**별로 학습해야 하는 단어는 빈도수의 **비율**로 **용사**별로 **차등**하여 나눌 필요가 있음.

→ The first 100 signs과 The second 100 signs은 초등1학년용 교재에 사용되어야 하는 어휘이고, The third 100 signs와 the fourth 100 signs은 초등2학년용 교재에, the 17th 100 signs와 the 18th 100 signs은 중학교 3학년용 교재에 적합한 어휘로 보이나 **용사**별 학습단어 빈도수는 차등있게 분배할 필요가 있음.

5.2 More issues to be considered

- (1) Do we need to define words in a sign language corpus as a collection of glosses as type, lemma, or word-family?
- (2) How to count words in the sign language corpus: **glosses** or **signs** in sentences?
- (3) We need more studies about parts of speech in KSL.
 - Are there any function words in sign language? If there are, which parts of speech exist.
 - For example, conjunctions ('-면', '-때'), postpositions ('후에'), pronouns, auxiliary verbs '가능하다'), modals('꼭 완료 상'), qualifiers, question words, etc.
 - In case of content words, how to identify nouns, verbs (plain verb 끝나다 vs. auxiliary verb 끝), adjectives, or adverbs ('예전', '옛날'(과거 시제), '앞으로'(미래 시제))?
- (4) Is the children KSL corpus necessary to add their preferable vocabulary for elementary and middle school deaf students?
- (5) To compare Chosun KSL corpus with the KSL online dictionary compiled by the National Institute of Korean Language, we need some criteria to collect the data and register words in the KSL dictionary.
- (6) To compare Chosun KSL corpus with the spoken corpus out of the contemporary Korean corpus called 'Modun Mal Mungchi,' and explore some similarities and differences between KSL as a sign language and Korean as a spoken language, we need further studies.

45

(2) **동향**, 만드신 1~9학년용 어휘목록을 보시고 초등 한국어학습용 1학년~6학년별 1개 총 6개 문장과 초등 영어학습용 3학년~6학년 6개 문장 (총 12개 문장, 주의: 영어는 초등 3학년부터 실시함), 중등 한국어/영어학습용 학년별 7학년~9학년 3개씩 총 6개를 만들어서 모두 공유한 후에 빠른 시일내에 확정짓고 난 후에 확정된 문장들을 조선대 공대 학교수님팀에서 겨울방학동안 중강현실(아바타 AR)을 만드시면 되겠습니다. 참고로 아래는 첨부한 연구소재로 제출한 연구계획서에 나와 있는 연구요약입니다.

연구목적: 코퍼스와 중강현실 기반 한국수어어휘 수준별 농인 초·중등학교 한국어 와 영어 등의 언어 교재 개발 (한국수어 교재는 2단계에서 개발) • 연구내용 : (1) 1~3년차: 한국연구재단 지원 (a) 한국수화언어법에 의거 문화체육관광부의 제2차 5개년 한국수어 발전기본계획(2022년~2027년)에 본 과제와 연관된 한국수어 교육 및 연구를 위한 계속 연구비 지원에 대한 정책제안 (b) 1~2년차: 한국수어 담화 동영상 100개 자료를 새로 수집하고 기 수집된 담화 동영상 30개를 재분석하여 총 130개 자료(약 200,000~300,000 단어)로 구성된 한국수어코퍼스 구축을 목표로 1년차는 기 수집된 30개의 재분석 및 50개 신 자료 수집과 2년차에도 50개 신 자료 수집·정밀전사·분석 (c) 1~2년차: 동영상자료 분석을 기초로 품사와 배수지신호 태깅으로 주석된 한국수어코퍼스를 구축 (d) 2년차: 중강현실(AR) 교재 샘플링을 위한 초·중등 과목별 아바타 선정과 시험 (초등 한국어 6 문장과 초등 영어 6 문장 샘플링) (e) 3년차: 한국수어의 초·중등 학년별 기본어휘(개별어휘와 관용표현 포함 단어) 목록 추출 (f) 3년차: 초·중등 학년별 수준별 한국수어 어휘기반으로 한국어와 영어 두 개 언어 교과의 학년별 한 단원의 대화 및 비대화 텍스트 및 문법과 어휘학습, 독해력 측정 등 연습문제 원고 집필 (g) 3년차: 초·중등 한국어와 영어의 교과 학년별 한 단원의 한국수어로 번역 (h) 3년차: 중강현실(AR) 교재 샘플링 (중등 한국어 3 문장과 초등 영어 3 문장)

50

References

- Baser, L., & Nation, P. (1993). Word families. *International Journal of Lexicography*, 6, 253-279.
- Brezina, V., & Gablanova, B. (2015). Is there a core general vocabulary? Introducing the new general service list. *Applied Linguistics*, 36(1), 1-22.
- Carroll, J. B., Davies, P., & Richmond, B. (1971). *The American heritage word frequency book*. New York: Houghton Mifflin, Boston American Heritage.
- Coxhead, A. (2000). A new academic word list. *JESOL Quarterly*, 34, 213-238.
- Dang, T. N. Y., Coxhead, A., & Webb, S. (2017). The academic spoken word list. *Language Learning*, 4, 959-997.
- Gansner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics*, 35, 305-327.
- Hsu, W. (2011). The vocabulary thresholds of business textbooks and business research articles for EFL learners. *English for Specific Purposes*, 30(4), 247-257.
- Hsu, W. (2014). Measuring the vocabulary load of engineering textbooks for EFL undergraduates. *English for Specific Purposes*, 33, 53-64.
- Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, 22(1), 15-30.
- Liu, N., & Nation, P. (1985). Factors affecting guessing vocabulary in context. *RELC Journal*, 16(1), 33-42. McLean (2018)
- Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly*, 47, 91-108.
- Nation, P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, P. (2016). *Making and Using Word Lists for Language Learning and Testing*. Amsterdam, The Netherlands: John Benjamins.
- Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, 47(4), 484-503.
- Schmitt, N., & Zimmerman, C. B. (2002). Derivative word forms: What do learners know? *JESOL Quarterly*, 36, 145-171.
- Webb, S., & Rodgers, M. (2009a). Vocabulary demands of television programs. *Language Learning*, 59(2), 335-366.
- Webb, S., & Rodgers, M. (2009b). The lexical coverage of movies. *Applied Linguistics*, 30(3), 407-427.
- West, M. (1953). *A general service list of English words*. London, UK: Longman.
- Zeno, S. M., Ivens, S. H., Millard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide*. Touchstone Applied Science Associates.

55

THANK YOU very much.



QUESTION

2[yes] GIVE 1[me]



Any Questions to me?

56



0 증강현실 기술의 이해



1. 증강현실(Augmented Reality, AR) 기술

- **증강현실이란?**
 - 현실의 정보에 가상의 정보를 합성해 사물이나 이미지의 정보를 ‘증강’시키는 것
 - 현실 세계와 가상 세계라는 두 차원이 겹치는 위치에 증강현실이 자리한다(Paul Milgram, 1994).
 - “**현실 세계 위에 3차원의 정보를 입혀**” 보여주는 기술
 - 옮긴된 Augmented **실재**현실 Reality
- **현실과 가상의 조합 / 실시간 상호 작용 / 3D에 기반**





출처: TED <http://www.ted.com> / Matt Mills: Image recognition that triggers augmented reality / 역시 인용함

2. 가상현실



- **가상현실(Virtual Reality, VR)이란?**
 - 존재하는 현실의 사물과 공간을 ‘가상’으로 재구성
 - 최초의 가상현실: 센서라마(sensorama) 기기(Morton Heilig, 1961) - proto-type
 - HMD형 가상현실 디스플레이(Huan Sutherland, 1968) 개발: 하버드 대학(대모클레스의 할)
 - “**현실 세계를 차단한 완벽한 디지털 환경**” 을 구축하여 가상의 세계에서 활동하는 것





Virtual Reality & Augmented Reality

출처: 안보니가우디움 VR다이트 플리만에서 인용함

3. 혼합현실

- 혼합현실(Mixed Reality, MR)이란?
 - 가상의 영상을 현실에 '겹쳐 더욱 발전된' 가상세계를 구현하는 기술
 - 가상의 이미지로 구현된 공간에 현실 정보를 덧붙이거나 현실 정보와는 무관하게 가상 이미지를 삽입
 - "가상현실 + 증강현실"의 장점을 결합한 형태



www.BANDICAM.com

혼합현실...가상 공간 + 실제 공간

• 출처: 구글이 만든 3D 홀로그람 '구글에서 만들었다'

4. 증강현실 기술의 이해와 실례

혼합현실
Mixed Reality(MR)

← 현실세계 Real Environment 증강현실 Augmented Reality (AR) 증강가상 Augmented Virtuality (AV) 가상현실 Virtual Environment →

현실과 가상세계 환경의 연속체 Milgram, Kerkin, 1994

 현실 세계	 증강현실 세계	 가상현실 세계
 실제 현실	 증강현실	 가상현실

1 증강현실 적용 방법

1.1 증강현실(Augmented Reality, AR)

- 2015 개정 교육과정
 - 교육부(2015): **창의융합형 인재 양성 목표**
 - 컴퓨팅사고력 향상을 위해 도입
 - 교육부(2018): **"증강현실 기술의 실감형 콘텐츠를 제공하는 디지털 교과서 제작"**
 - 사회, 과학, 영어 교과목 디지털 교과서 보급
 - 초등학교 3-4학년, 중학교 1학년
 - 교육부(2019): 실시간 예뮬테크 콘텐츠 제공하는 **교실환경 구축**
 - 학교공간혁신사업
 - "미래 교실"은 어떤 모습? www.etrn.go.kr



* 출처: 교육과 기술의 융합, 에듀테크가 꿈꾸는 미래(다산출판사) / YTN 사이언스캐서 인용됨

1.2 연구배경 및 목적

2016년 한국수화언어법 제정

- 한국수어가 한국어와 똑같은 지위를 인정받게 되고 우리나라 농언들은 그들 모국어인 한국수어를 통해 학습할 수 있도록 하는 **농인 언어권 및 학습권을 보장하는 법적 지위를 얻게 됨(Law, 2016)**

첫째 모국어를 사용하여 학습할 수 있는 농인의 언어권과 학습권을 보장하기 위해서는 농인 초·중등 언어 교재를 개발하는 일이 교육과 연구인 차원에서 상당한 의미가 있는 일뿐만 아니라 농인의 언어권과 학습권 및 **둘째** 직업창출이라는 생존과 관련된 안건의 차원에서 반드시 이루어져야 할 정책(agenda)

1.2 연구배경 및 목적

- 증강현실이라는 새로운 컴퓨터 기반 멀티미디어 기법을 적용시켜 농인들의 능동적인 반응과 교육적 관심과 흥미를 유도할 수 있는 증강현실 기반의 콘텐츠를 개발하여 농인들을 위한 교재개발을 통해 교육적 학습 성취도에 도움을 주고자 함
- 본 연구에서는 이러한 현실적 문제를 해결하기 위해서 우리나라 최초로 농인을 위한 코퍼스와 증강현실 기반으로 수준별 초중 한국어와 영어 등 필수 언어 교과 교재 개발을 연구 목표로 구축된 수어 코퍼스를 통해 증강현실 적용방법에 대해 알아보하고자 함

2

연구 방법 및 내용

2.1 한국수어 동영상 자료수집 및 주석 코퍼스 구축과 코퍼스언어학 방법론을 통해 자료 수집

- 첫째, 영상수집 방법
 - 식생활 및 수어 통역의 다양한 장르에서 수어 동영상 자료를 수집함
- C (video recording of right person), A (overall), B (video recording of left person)

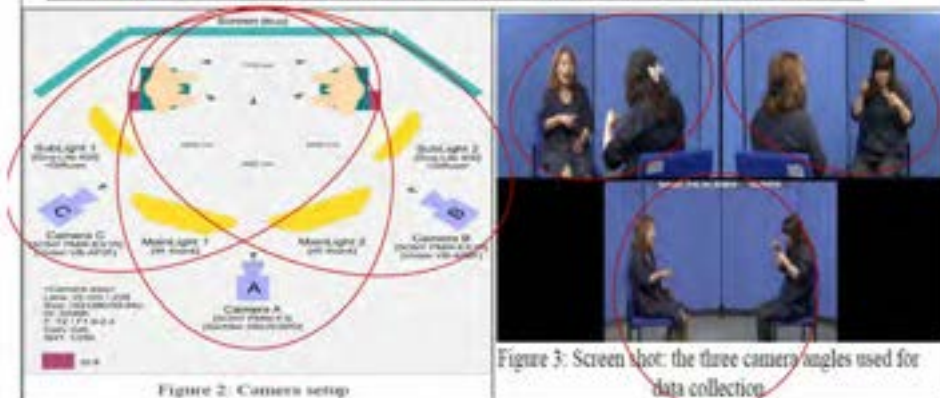



Figure 3: Screen shot: the three camera angles used for data collection.

idea

- **소재, 전사와 분석방법**
 - 동영상 자료를 ELAN(EUDICO Linguistic Annotator) 프로그램을 사용하여 청밀전사와 어구주석(gloss) 분석으로 개별어휘와 관용어휘 포함한 데이터에 사용되는 어휘 비수지선호, 태깅으로 주석된 한국수어코퍼스를 구축

A compilation stage of Gwangju KSE Corpus



idea

- **소재, 코퍼스 구축방법**
 - 소리언어의 코퍼스언어학에서 채택하고 있는 BNC-COCA 25000Range/COCA_MWU20 ColloGram 프로그램(Shin et al., 2018)과 같은 방법으로 코퍼스언어학 방법론을 활용하여 한국어의 초·중등 학년별 한국수어의 기본 개별어휘와 데이터(관용표현 포함) 목록을 제시

• Naming of Video + Annotation Files
 Current staying region_informants number_Birth place_Age group_Sex_Video topics

WJGS_017.018_ILJL_40.40_FM_3-3.3-4.4-1.5-1
 ① ② ③ ④ ⑤ ⑥

구분	설명	예시
지역	1. 원거주 지역	1. 원거주 지역
지역	2. 원거주 지역	2. 원거주 지역
지역	3. 원거주 지역	3. 원거주 지역
지역	4. 원거주 지역	4. 원거주 지역
지역	5. 원거주 지역	5. 원거주 지역
지역	6. 원거주 지역	6. 원거주 지역
지역	7. 원거주 지역	7. 원거주 지역
지역	8. 원거주 지역	8. 원거주 지역
지역	9. 원거주 지역	9. 원거주 지역
지역	10. 원거주 지역	10. 원거주 지역
지역	11. 원거주 지역	11. 원거주 지역
지역	12. 원거주 지역	12. 원거주 지역
지역	13. 원거주 지역	13. 원거주 지역
지역	14. 원거주 지역	14. 원거주 지역
지역	15. 원거주 지역	15. 원거주 지역
지역	16. 원거주 지역	16. 원거주 지역

① WJ : Wan-Ju(완주), GS : Gun-San(군산)
 ② 017, 018: Informants
 ③ J : Birth place (Jeon-Ju, 전주)
 ④ 40: forbes year old
 ⑤ Female, Male: Sex
 ⑥ Topics(주제)
 3-3 (Food Life: most favorite food that you recently ate),
 3-4 (Food Life: any food that you want to eat in future),
 4-1 (Clothing Life: Tell me how to purchase clothes and shoes),
 5-1 (Housing Life: your preferred place and time to meet with you're a few)

*Currently 16 major categories , 88 subcategories, 199 video topics

idea

Top 10 and bottom 1 ranked words

	Noun	freq	Verb	freq	Adjective	freq	Adverb	freq
1	농인 deaf	120	맞 be correct	240	같 same	223	또 again	81
2	때 when	118	생각하 think	149	없 not exist/dot have	178	자연히 naturally	65
3	사람 person	110	하 do	133	있 exist/have	168	가장 most	60
4	끝 end	100	가 go	97	좋 good	156	지금 now	59
5	면 if	92	알 know	88	많 many	117	보다 than	52
6	곳 where	81	말하 say	84	괜찮 fine	63	안 not	47
7	영화 movie	70	일하 do	79	다르 different	62	왜 why	33
8	집 house	68	되 become	67	힘들 hard	55	오두 all	30
9	옛날 old times	68	모이 gather	64	즐겁 happy	52	함께 together	27
10	결국 finally	68	오 come	61	가능하 possible	48	그냥 just/as it is	26
136							정자표 silently	1
215					저렴하 cheap	1		
732			시합하 play a game	1				
1144	천리 Cheonli	1						

idea

• 넷째, 증강현실 적용

- 공간과 시각언어인 수어의 특성을 초중등 한국어와 영어, 한국수어 (교과과정 개편이 이루어진 2단계에서 교재 개발) 등 필수 언어 교과 교재에 반영하고 전 제4차 산업혁명시대의 신교재의 개념에 가장 적합한 증강현실(AR) 기술을 시험적으로 도입



2.2 연구대상(주제별, 연령별, 지역별 통계)

1. 촬영 기간 및 전사 기간

가) 촬영 기간 : 2020.12 - 2022.04(5개월)

나) 전사 기간 : 2021.01 - 2022.03(5개월)

가. 주제별 참여 인원

주제	참여인원
연간	18
살	24
식생활	18
피생황	14
주생활	14
사회생활	18
갈계생활	14
교육	12
종교	12
문화	18
정치와 행정	14
자연	18
통식물	18
계명	12
국가명 및 지역	16
기타	18

나. 연령층: 평균 49.6세(남 16명, 여 60명)

나이	참여인원
20대	1
30대	2
40대	12
50대	6
60대	3
70대	1
80대	1
총 인원	26

다. 지역

지역	성별	인원	지역별 참여 인원
전남	강진	1	23
	함양	2	
	장흥	7	
	나주	1	
	남악	3	
	목포	3	
	무안	2	
	보성	1	
	순천	1	
	화순	2	
전북	완주	1	2
	진주	1	
서울(경기)	인천	1	1

3 증강현실 적용 : AR이모지 카메라 활용

3.1 교육적 효과

· 장점

- 학습 몰입감, 성취도, 흥미, 주의집중, 학습태도에 긍정적 영향
- 콘텐츠 유형: 소용형(91.5%), 실습형(4.3%), 협동형과 체험형(2.1%)
- 콘텐츠 내용:
 - 소용형: 여러 기법 콘텐츠를 카메라 화면에 교과서나 도서를 비추면 가상의 물체가 나타나는 형식
 - 실습형: 자동차 매뉴얼 학습을 실습
 - 협동형: 사물놀이 합주의 음악 교과 콘텐츠
 - 체험형: 특정지역에 카메라 화면을 비추고 건물 상호영을 터치하면 관련 영어 문장이 제공되는 학습 콘텐츠

구분	교육효과										총계
	몰입도	성취도	흥미	주의집중	학습태도	소용형	실습형	협동형	체험형	기타	
2020	1	1	1	1	1	1	1	1	1	1	11
2021	1	1	1	1	1	1	1	1	1	1	11
2022	1	1	1	1	1	1	1	1	1	1	11
합계	3	3	3	3	3	3	3	3	3	3	33

*출처: 박유빈, 김성민(2023), 「AR 증강현실 활용 교육 효과 분석」, 한국언어과학회 2023년 겨울학술대회



4 증강현실 콘텐츠 내용

4.1 농인 교재 개발 절차

데이터 구축

동영상 수집과 전사 **코퍼스 구축**

>>이절 추출>>

3D아바타 **증강현실**

4.2 농인 교재 개발 : 국어

(1) 3차원 영상 정보를 보여줌으로써 청각장애학생의 수어(hand language)교육에 효율적 적용

- 초등 교과서 국어 문형의 선정 기준(초등 3학년)
 - 한국수어 코퍼스 기준(조선대 언어활용연구소) 상위 80%의 빈도수(20만 어휘)를 가지고 있는 어휘 중에서 세술어를 선택함
 - 수어의 일반동사, 공간동사, 일치동사(한국어 세술어를 3개의 단원에 배치)

<초3 국어 단원>

- 1단원: 일반동사('먹다', '만들다')**
 - 문형 1: (-를 먹다) 지호가 **편지**를 먹어요.
 - 문형 2: (-를 만들다) 지호가 **떡갈이**를 만들어요.

80%를 벗어나는 어휘거나 한국수어코퍼스가 현재 20만 어휘로 매우 제한적인 정음 고려하여 한국어에서 문장에 자연스럽게 사용할 수 있는 빈도수가 높은 명사를 선정
- 2단원: 공간동사('오다', '가다')**
 - 문형 1: (-에 오다) 수지가 **통계** 왔어요.
 - 문형 2: (-에 가다) 수지가 **학교**에 가요.

상위 80%의 빈도수에 해당되는 어휘
- 3단원: 일치동사('주다')**
 - 문형 1: (-에게 주다) 지호가 **안타까**게 **편지**를 줬어요.
 - 문형 2: (-를 주다) 지호가 친구에게 **상여지**를 줬어요.

4.3 농인 교재 개발 : 영어

(a) 3차원 영상 정보를 보여줌으로써 청각장애학생의 수어(Deaf Sign Language)교육에 효율적 적용

- 초등 교과서 영어 문형의 선정 기준(초등 3학년)

- 국어와 동일한 기준으로 선정했으며, 국어와 동일한 문형을 영어로 학습하도록 선정

<초3 영어 단원>

- 1단원: 일반동사(‘먹다’, ‘만들다’)

- Jiho eats fish. We eat fish everyday.
- Jiho makes a necklace. Jiho and Suji make a necklace.

- 2단원: 공간동사(‘오다’, ‘가다’)

- Suji comes home. Come here!
- Suji goes to school. Don't go!

- 3단원: 일치동사(‘주다’)

- Jiho gives his mother a book. I will give my friend a book tomorrow.
- Jiho gives his friend a puppy. I will give Suji flowers tomorrow.

5

수어코퍼스 기반 증강현실 적용 제시 방법

5.1 코퍼스 전사 데이터 수집 후 빈도수를 고려한 문장 선정

• 제시 방법

- 코퍼스 데이터를 바탕으로 초등학생의 각 단원별 문장 선정

- 초등 3학년 코퍼스 전사 내용을 기준으로 문장 선정

1. 글자(문장)

- 국어 문장(6개)
 - 단원 1, 단원 2, 단원 3
- 영어 문장(6개)
 - 단원 1, 단원 2, 단원 3

2. 음성



3. AR 이모지



4. 농인 수어 영상

초3단원1	(-를 먹다)	지호가 생선을 먹어요.
초3단원2	(-를 만들다)	수지가 목걸이를 만들어요.
초3단원1	(-에게 오다)	지호가 집에 와요.
초3단원2	(-에게 가다)	수지가 학교에 가요.
초3단원1	(-에게 주다)	지호가 엄마에게 책을 줘요.
초3단원2	(-를 주다)	수지가 친구에게 장미꽃을 줘요.

We eat fish everyday.

Jiho and Suji make a necklace.

Come here!

Don't go!

I will give my friend a book tomorrow.

I will give Suji flowers tomorrow.

5.2 초등학교 교재에 증강현실 적용(국어, 영어문장)

• 증강현실 방법적 모색 <글자(1), 음성(2), AR이미지(3), 농인 수어(4)>

• 3학년 1학기(초등학교)


• 국어 문장

• 단원 1 : (문형 1) : (-을 먹다) 지호가 생선을 먹어요.

• 문장 (글자) ⇒ 지호가 생선을 먹어요.

• QR1 (음성) ⇒ 

• QR2 (AR 이미지) ⇒ 

• QR3 (농인) ⇒ 

• 영어 문장

• 단원 1 : (문형 1) : (-을 먹다) 지호가 생선을 먹어요.

• 문장 (글자) ⇒ We eat fish everyday.

• QR1 (음성) ⇒ 

• QR2 (AR 이미지) ⇒ 

• QR3 (농인) ⇒ 

5.3 교재에 증강현실 적용: 교재 개발 형식

• 3학년 1학기(국어 문장)

• (단원 1)

- 학습 목적: 학습 몰입감, 심취도, 흥미, 주의집중, 학습태도에 긍정적인 영향을 준다.
- 콘텐츠 내용: 여러 기법 콘텐츠로 커리어 화면에 QR코드를 비추면 기존의 문장의 글자, 음성, 사람이 나타나는 형식

• 문형 1 : (-을 먹다) 지호가 생선을 먹어요.

• 문형 2 : (-을 만들다) 지호가 목걸이를 만들어요.

• 문장 (글자)

• QR1 (음성)

• QR2 (AR 이미지)

• QR3 (농인)

지호가 생선을 먹어요.



지호가 목걸이를 만들어요.



5.3 교재에 증강현실 적용: 교재 개발 형식

3학년 1학기(영어 문장)

(단원 1)

- 기초 어휘: 기초 회화형, 단원교, 꽃이, 거장님요, 새우요요와 교양에 도움을 받아
- 문장의 내용: 자연 관찰 보고서, 자연의 아름다움, QR코드로 새우와 거장님 문장의 글자, 음성, AR이모지, AR이모지

- 문장 1(1회 직사): We eat fish everyday.
- 문장 2(1회 직사): Jiho and Suji make a necklace.

- 문장 (글자)
- QR1 (음성)
- QR2 (AR 이모지)
- QR3 (농인)

We eat fish everyday.	QR1	QR2	QR3
Jiho and Suji make a necklace.			

5.4 교재에 증강현실 적용: 학습방법 형식(3개 단원) 전체: 3학년 국어(글자, 음성, AR이모지, 농인)

초3단1문1	(-를 먹다)	지호가 생선을 먹어요
초3단1문2	(-를 만들다)	수지가 목걸이를 만들어요
초3단2문1	(-에 오다)	지호가 집에 와요
초3단2문2	(-에 가다)	수지가 학교에 가요
초3단3문1	(-에게 주다)	지호가 엄마에게 책을 줘요
초3단3문2	(-를 주다)	수지가 친구에게 감아지를 줘요



5.4 교재에 등장현실 적용: 학습방법 형식(3개 단위) 전체: 3학년 영어(교차, 음성, ... , 동영상)

초3단1문1	(-를 먹다)	We eat fish everyday.	
초3단1문2	(-를 만들다)	Jiho and Suji make a necklace.	
초3단1문1	(-에 오다)	Come here!	
초3단1문2	(-에 가다)	Don't go!	
초3단1문1	(-에게 주다)	I will give my friend a book tomorrow.	
초3단1문2	(-를 주다)	I will give Suji flowers tomorrow.	

감사합니다

Leap Forward. PNU



Gesture and/to Sign in Language

2023. 02. 16

장세은
(한국해양대학교)

한국언어과학회 2023 겨울학술대회
부산대학교 ZOOM 온라인

목차

서론 Q: 우리나라의 공용어?

국어기본법 vs. 한국수화언어법
연도별 한국수화언어와 관련된 주요 역사
국립국어원 한국수어사전 vs. 조선대한국수어사전

Sign에 대한 이론적 배경
소리언어와 수어언어와의 차이점과 유사점

본론: Gesture vs. Sign
How to distinguish **gesture** from **sign** in Linguistics ?
A history of sign language and gesture studies over the last 50 years

Case study: 수어언어에서 gesture와 sign은 어떻게 구별하나?

결론, 문제점, 제언
참고문헌

Q: 우리나라의 공용어?



한글전용에 관한 법률

[시행 1948. 10. 9.] [법률 제6호, 1948. 10. 9., 제정]
대한민국의 공용문서는 한글로 쓰되, 일정기간 한자를 병용할 수 있도록 함

국어기본법

[시행 2005. 7. 28.] [법률 제7368호, 2005. 1. 27., 제정]
5년마다 국어발전기본계획을 수립해서 시행해야 하고, 2년마다 국회에 국어의 발전과 보급에 관한 시책과 시행 결과를 보고해야 한다. 또한 국가기관과 지방자치단체에 국어의 보전과 발전 업무를 총괄하는 국어책임관을 두도록 함

제3조(정의) 이 법에서 사용하는 용어의 뜻은 다음과 같다.
1. "국어"란 대한민국의 공용어로서 한국어를 말한다.
2. "한글"이란 국어를 표기하는 우리의 고유문자를 말한다.

한국수화언어법

[시행 2016. 8. 4.] [법률 제13978호, 2016. 2. 3., 제정]

<http://www.law.go.kr/lsInfoP.do?lsiSeq=180821&efYd=20160804#0000>

- [제정이유]
- 주요 내용

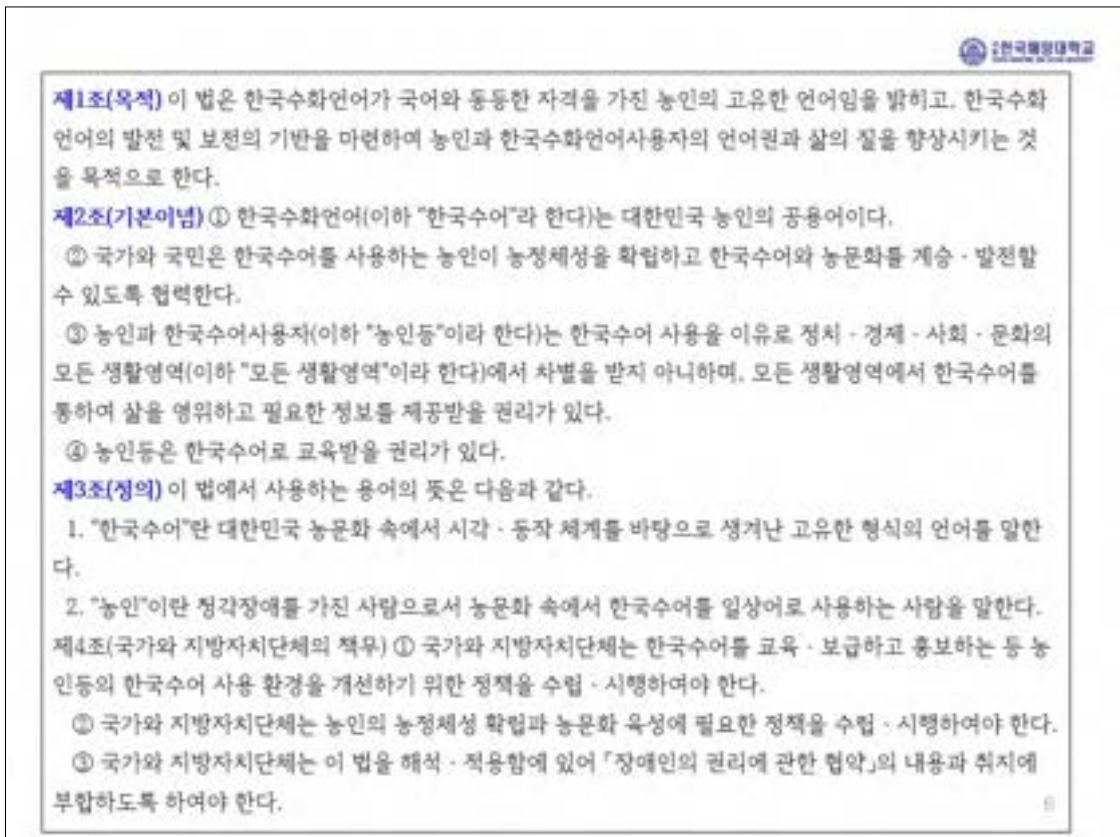


[제정이유]


농인은 음성언어인 한국어를 대신하여 **한국수화언어를 제1언어**로 사용하여 왔으나, 현행 우리 사회의 수화언어 사용 환경이 미비하여 수화언어를 이용한 의사소통·정보이용·학습 등에 제약이 있음. 이는 결과적으로 농인이 일상 생활뿐만 아니라 교육, 취업 등 삶의 여러 분야에서 열악한 처지에 놓이게 되고 **사회적 소외계층**으로 머물러 있는 요인이 되고 있음.

현행 「장애인복지법」 등에서 농인이 정보에 원활히 접근하도록 하고 수화언어통역 등에 대한 제도를 확충하는 등 농인의 권리를 높이기 위하여 노력하고 있으나, **농인의 근본적인 언어권 신장과 삶의 질 향상에는 여전히 한계가 있는 실정임.**

이에 **한국수어가 한국어와 구별되는 고유한 자격의 공용어**임을 선언하고, 한국수어의 보급·발전과 농인의 교육·사회·문화 등 모든 생활 영역의 기반을 마련하는 **법률을 제정함**으로써 농인의 삶의 질을 향상시키고 사회활동의 참여를 증진하려는 것임.



연도별 한국수화언어 주요 역사와 사전 및 법제정



연도	주요 내용	
1909	사립중앙영어학교 대한제국 인가	
1921	조선영어학회 창설	
1947	국립영어학교(현 서울 농어학교) 운영권 교장 한글지문자 장안	
1982	표준수화사전 (5,494단어, 서울농어학교 표준수화사전편찬위원회)	
1994	한국수화연구회 (한국수화학회, 한국수어학회 전신) 설립	
1996	학술지 '수화 연구' 창간호 발간	
1997	제1회 수화통역사자격 인정시험 실시	
2000	한국표준수화규범 제정 추진위원회 구성 (문화관광부, 한국농어인협회)	
2001	나사렛대학교 재향복지대학원 국제수화통역학과 개설	→ 현재 4년제 후안재활학부 수어통역교육전공
2002	국립한국복지대학교 수화통역과 개설	→ 현재 3년제 한국수어교육원과
2005		
2007	한국표준수화사전 초판 및 개정판 (17,600 단어)	
2011	문화체육관광부 국립국어원 한국수어사전(2005년부터 수화표준화사업 이어 발음)	
2013	한국수화언어법 국회 발의	
2014	국립국어원 한국수어 발전을 위한 기본계획수립연구과 한국수어의 알맞은 표현을 위한 기초연구 (참고: 한국어 국가 발음서 21세기 세종계획 1998-2007 약 2억 어절)	→ 국립국어원 https://www.korean.go.kr/
2015	국립국어원 한국수어 발음서 시범구축 5월-12월 표본 68,000개 타입 2,400개	
2015	한국수화언어법 국회 통과 (12.31)	
2016	한국수화언어법 [법률 제13978호 2016. 2.3 제정] 대한민국 농인의 공용어 (??)	
2016	국립국어원 한국수화언어 정책 토론회 (4월 8일 여의도 이룸센터 누리홀)	
2016	한국수화언어법 [시행 2016. 8. 4.]	
2018	국립국어원 한국수어사전 국제학술대회 (11.8 이화여대)	
2018	국립국어원 한국수어사전 정비 및 구축 (예: "한국수어사전" 수어용량상 12,703개 한국어 표제어 중 7,053개 포함 55.5%)	
2019	7, 28, 현재: 국립국어원 한국수어의 언어적 특성에 맞는 사전은 현재 준비 중 일상생활 수어 12,594 + 전문용어 수어 10,984 + 문화정보 수어 1,627 = 25,205건. 영상자료	

국립국어원 한국수어사전




-2022년 현재:
-한국어 대응표현
한국수어 lemma (표제어)
웹사전제공
→ 3,676개

-한국수어 표제어 등록수
→ 24,705개(??)

- 일련번호 1~ 20 예 (우측표)

번호	수어 표제어 번호	한국어 대응표현
1	24705	복직 계약
2	24704	이혼, 쓰다
3	24701	(곳)이/자다
4	24699	할거야
5	24647	(마음)이/는다, 흥물하다, 흥하다
6	24641	(마)지/을/었다
7	24637	인도(문화)국
8	24626	안니다, 앓다
9	24634	여지, 여경, 여
10	24633	살거지
11	24632	배부르다
12	24630	(여)지/을/유록, 유물하다
13	24629	살, 셋
14	24628	시간, 할 시간
15	24626	놓여우다
16	24624	둘, 여의, 여러가지
17	24622	시작, 개사, 개회, 거할, 시발, 착수, 출발, 많다, 이할, 하수
18	24620	틀 (바닥)에 닿기는(문)질
19	24619	참다, 건다, 만나, 자제, 죽게
20	24615	회리

조선대한국수어코퍼스 (Dec. 2020~May 2022) Top 20 and Bottom 10 words 

Top 20 in The Study Corpus						Bottom 10 in The Study Corpus			
1	Gloss	POS	types	types(%)	types(CUM%)				
2	나	NP	7,452	2.9665	2.9665				
3	저기	NP	4,571	1.8196	4.7861				
4	같다	VA	3,816	1.5191	6.3052				
5	있다	VA	3,244	1.2914	7.5966				
6	맛다	VV	3,182	1.2667	8.8633				
7	없다	VA	2,820	1.1226	9.9859				
8	꿈	NNG	2,610	1.0390	11.0249				
9	너	NP	2,559	1.0187	12.0436				
10	좋다	VA	2,459	0.9789	13.0224	17644	전부/전체	MAG	1 0.0004 99.9964
11	보다	MAG	2,399	0.9550	13.9774	17645	가웃똥	MAG	1 0.0004 99.9968
12	많다	VA	2,004	0.7978	14.7752	17646	전.영.수	MAG	1 0.0004 99.9972
13	-면	NNG	1,754	0.6982	15.4734	17647	김.영.란	MAG	1 0.0004 99.9976
14	집	NNG	1,532	0.6099	16.0833	17648	-면/-문/S	MAG	1 0.0004 99.9980
15	생각하다	VV	1,470	0.5852	16.6685	17649	전.영	MAG	1 0.0004 99.9984
16	말하다	VV	1,425	0.5673	17.2357	17650	디.	MAG	1 0.0004 99.9988
17	때	NNG	1,376	0.5478	17.7835	17651	역치로/강?	MAG	1 0.0004 99.9992
18	살하다	VV	1,374	0.5470	18.3304	17652	영.휘.	MAG	1 0.0004 99.9996
19	농인	NNG	1,370	0.5454	18.8758	17653	-중.후.	MAG	1 0.0004 100.0000
20	안하다	VV	1,314	0.5231	19.3989				
21	무엇	NP	1,311	0.5219	19.9208				

251,205 tokens and 17,652 types

수어/점자

수어 소개

수어	
수어 소개	한국수어의 의미 청각장애인은 소리로 말을 배울 수 없어서 '보이는 언어'를 사용한다. 이 '보이는 언어'가 바로 '수어 (手語, Sign language)'다. 이차음 수어를 일상어로 사용하는 사람을 '농인'이라고 한다. '한국수화언어법'에 따르면 '한국수어'는 '한국수화언어'를 줄인 말로, 한국어나 영어와 같은 독립된 언어라는 의미를 갖고 있다. 한국수어는 한국어와는 문법 체계가 다른, 대한민국 농인의 고유한 언어이다.
한국수어 관련 자료	
한국수어 용어 명칭	
새수어	한국수어, 농인, 농문화의 의미
한국수어	우리나라 농인들이 사용하는, 보이는 언어 (한국어와는 다른 고유한 형식이 있음)
농인	청각장애인을 달리 이르는 말로서 수어를 일상어로 사용하는 사람
농문화	농인으로서의 정체성과 가치관을 바탕으로 형성된 모든 생활양식

출처: 국립국어원 홈페이지
https://www.korean.go.kr/front/page/pageView.do?page_id=P000300&mn_id=202

수어와 몸짓(제스처)의 차이

수어를 두고 한국어로 말할 때 쓰는 몸짓(제스처) 같은 것이 여자가 모해할 수 있다. 하지만 수어는 단순한 몸짓과는 전혀 다르다. 수어는 손과 손가락이 모양수형(손바닥의 방향수형), 손의 위치수형, 손의 움직임수형(형태 따라 의미가 달라진다. 또한 같은 목적을 위해서도 어떤 표현을 정교하게 나타내거나 다른 의미가 있다).





오전



오후

오전(왼쪽으로) / 오후(오른쪽으로) 손의 모양은 같으나 손의 움직이는 방향에 따라 의미가 달라짐.




맛있다



맛없다

맛있다 / 맛없다(표정함으로) 손의 모양은 같으나 표정(제스처)에 따라 의미가 달라짐.

국립국어원 홈페이지
https://www.korean.go.kr/front/page/pageView.do?page_id=P0003008&mn_id=202



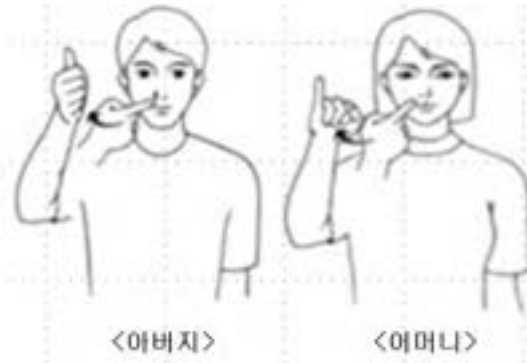
Theoretical Background

Modern Approaches to Language

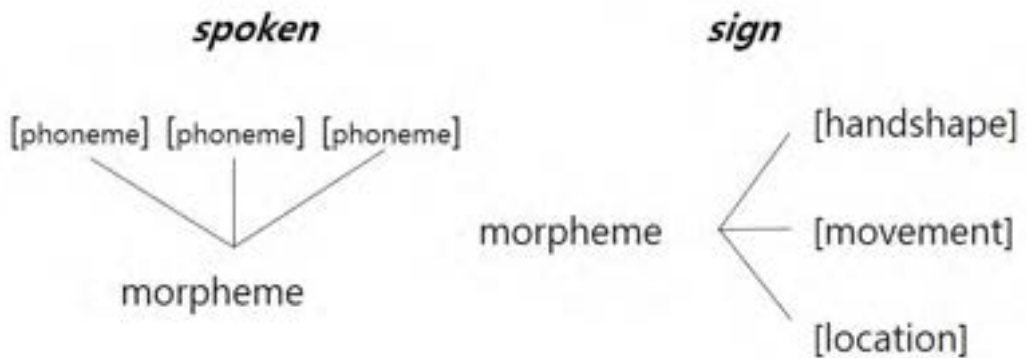
- “All human beings were born with **an innate capacity for language**” (Chomsky)
- Grammar: the unconscious mental rules that govern linguistic behaviors in speech.
- Sign languages are completely linguistic & rule-governed.

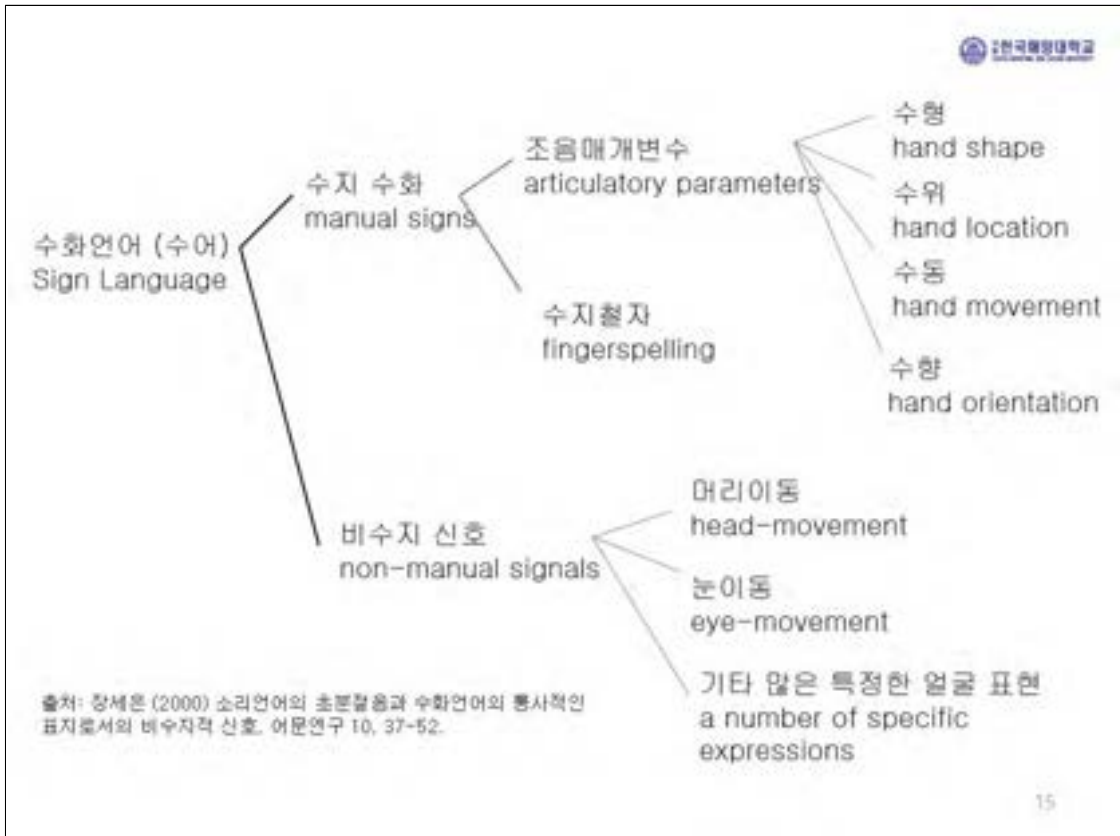
Sings: linguistically organized??

Signs have **PARTS**.



Phonology & Morphology of Languages





Differences & Similarities in Spoken vs. Sign

			Auditory-vocal	Visual-gestural
Phonetics	Physical modality	Production	lip	hands
		perception	ears	eyes
Phonology	Sublexical units	segments	(non)-linear sequences	(non)-linear sequences

제스처(gesture) vs. 수화(sign)

제스처와 수화의 사전적 정의
Oxford English Dictionary (1989)

- **Gesture**: "a movement of the body or any part of it that is expressive of thought or feeling."
[생각이나 느낌을 나타내는 몸의 움직임]
- **Sign**: "a **gesture** or motion of the hand, etc. serving to convey an intimation or to communicate some idea"
[어떤 아이디어를 전달하거나 어떤 암시를 전하려는 제스처 또는 손짓]



이러한 의미에서 gesture와 sign은 상호 호환성있는 개념으로
19세기에는 gesture languages는 language of signs와 혼용

17

How to distinguish **gesture** from **sign** in Linguistics ?

- **Nonarbitrariness vs. Arbitrariness** in relation to form and meaning
- Distinction between imagistic **gestures** and categorical **sign (or speech)**
- Problem of combining **two antinomical types of meaning**: lexical and gestural → objective meaning bearers (words and signs) vs. idiosyncratic and contextually dependent meaning (gestures)

(ex. **Co-speech gestures** modify or complement verbal expression.)

Although both **sign** and **gestures** involve the use of the hands (with other parts of the body), they are rather different. **Sign** is like speech and is used instead of speaking, whereas **gestures** are mostly used while speaking.

A history of sign language and gesture studies over the last 50 years

Sign language has, at times, been viewed

- (1) as a language of gestures and therefore very different from spoken language, and at other times,
(→ Sign language is not a language.)
- (2) as a language characterized by structures just like those found in spoken language; more recently,
(→ Sign language is just like spoken language and therefore a language)
- (3) researchers have recognized that sign language has gestural components just as spoken language does.
(→ Sign language is not like spoken language in all respects – could the differences be gestural?)

(Goldin-Meadow and Brentari (G-M& B), 2015/2017)

Gesture, sign, and language: The coming of age of sign language and gesture studies. *BEHAVIORAL AND BRAIN SCIENCES* (2017), 1-60.

도상성 (iconity) vs. 자의성(arbitrariness)

언어 기호의 형태가 의미와 체계적으로 닮아 있는지의 여부 → ???

(수어의 형태와 의미와의 관계) - 임지룡 (2018) 한국수어의 도상성 양상과 의미 특성

- 비유
- 대립어
- 동기화 양상
- 신체어
- 은유와 환유



a.	머리 옆: 지적, 정신적 활동
b.	눈 부위: 보는 것
c.	코 부위: 가치 평가, 냄새의 판단
d.	입 부위: 말하는 것과 먹는 것
e.	귀 부위: 소리 관련 의미 표현
f.	뺨 부위: '진가'의 표현
g.	턱 부위: '인생'의 의미 표현
h.	목 부위: '경험'과 '욕망'의 의미, '권위'의 상징
i.	배 부위: '마음'과 관련된 의미
j.	어깨 부위: '의무'와 '책임'의 의미
k.	위팔 부위 관련 기호: '힘'과 '권세'의 의미
l.	아래팔 부위 관련 기호: '노동'의 의미 (장진권 1996)

장진권(1996/1998) 한국수화의 어휘적 의미, 단국대 석사학위논문.
한국수화학회지(1996) 1, 51-62.

기호의 형식과 내용의 관계



Charles Sanders Peirce (퍼스, 1839-1914)

- (1) 도상(icon) → 기호는 닮은 모양 (사물의 형태나 행동을 모방)
- (2) 지표(index) → 기호는 자연적 관계 (인접성)
- (3) 상징(symbol) → 기호는 관습적 관계 (자의성)

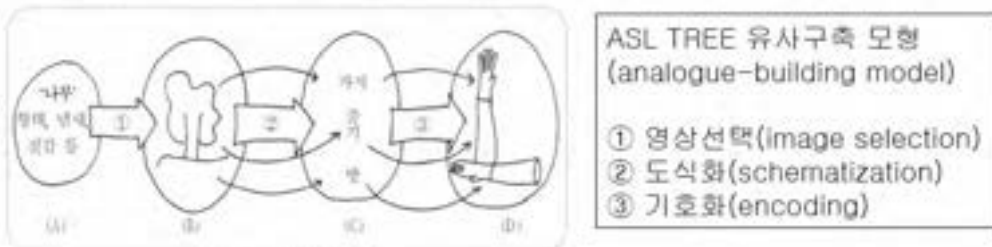
※ 남기현(2012) KSL 2,674개 단입어 중 [도상+지표] vs. [상징]의 비율은 각각 50% 차지

- (1) 도상(icon)
- (2) 지표(index)
- (3) 상징(symbol)

				
<p>{집} 두 손의 손끝을 맞대어 좌우로 비스듬히 세운다</p>	<p>{가볍다} 약간 구부린 오른손 가락 끝으로 왼손 등을 긁는다</p>	<p>{눈} 오른 주먹의 1지를 펴서 끝으로 오른쪽 눈을 가리킨다</p>	<p>{귀다} 오른 주먹의 1지를 펴서 끝 바닥으로 목을 스쳐 내린다</p>	<p>{간단하다} 1 5지 끝을 맞대어 동그라미를 만든 오른손을, 주먹 끝이 위로 향한 왼 팔뚝에 내려 놓으며 1 5지를 떼다</p>

21

형태와 의미 간의 도상적 수어 생산되는 과정 (Taub 2001: 44; 임지홍 2018: 70)



(그림 2) 유사구축 모형

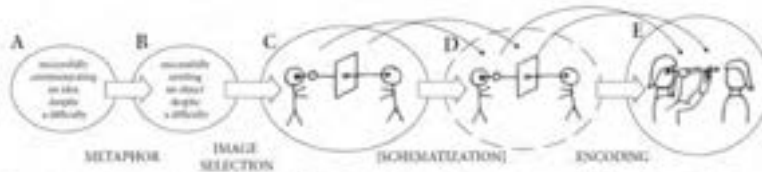


Figure 2. Analogue-Building Model of Metaphorical Iconicity applied to the ASL sign THINK-PENETRATE (Taub 2001: 111)

Q: Why should iconic mapping be a part of the creation of metaphorical signs?

A: Metaphors are not inherently linguistic, but a part of cognition: metaphorical mapping happens between concepts, not between linguistic entities.

22

한국국립대학교

환유1: 집 전체의 구체적인 영상에서 영상을 선택
환유2: 집의 여러 구성 요소 가운데 가장 현저한 지붕을 도식화

〈그림 3〉 유사구축 모형에 의한 '집'의 수어 과정

한국 수어와 중국 수어의 차이점: 문화 간 수어 차이의 인지 과정

〈그림 15〉 '집'의 한국 수어(KSL)와 중국 수어(CSL)의 차별적 영상

이미지 선택 (image selection) → 도식화 (schematization) → 무호화 (encoding)

23

한국국립대학교

A case of (3): Karen Emmorey (2017)

How to distinguish gesture from sign: New technology is not the answer. *BEHAVIORAL AND BRAIN SCIENCES* (2017), 40.

Better models are needed to account for differences between co-speech and co-sign **gesture** (e.g., different degrees of optionality, existence of beat gestures).

Beat gestures are small movements of the hand (or head) that contain little semantic information (unlike deictic or iconic gestures)

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7147075/>



A case of (3): Emmorey (2017) (cont'd)

As noted by G-M&B, most linguistic analyses of sign language assume that directional (“agreeing”) verbs and pronouns comprise both a linguistic and a gestural component, although whether the gestural component (indicating the location of a referent) is always expressed is an empirical question.

The primary problem is that **sign and gesture** cannot be separated by articulatory systems, **unlike speech and gesture**. Although motion capture technology was able to characterize small changes in head movements, it is hard to find noticeable differences between head movements that seemed more gestural (e.g., nods indicating affirmation) and those that were likely to be more linguistic (e.g., nods occurring at syntactic boundaries).



A comparison of different models proposed for **speech** and **gesture** production

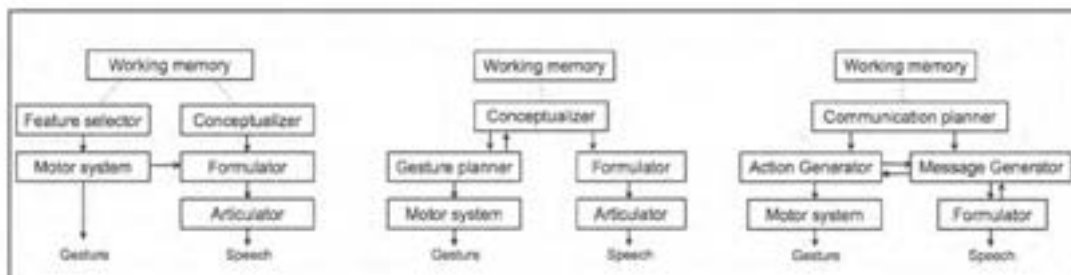


Figure 25.1 Adapted schematic overview of different models in relation to speech and gesture production (from left to right: Morrel-Samuels & Krauss (1992); de Ruiter (2000), Kita & Özyürek (2003) (taken from Wagner et al. (2014) with permission))

Source: Woll, Bencie and David Vinson (2011) Chapter 25 Gesture and Sign: Theoretical and experimental perspectives. In Josep Quer, Roland Pfau, and Annika Herrmann (eds.) *The Routledge Handbook of Theoretical and Experimental Sign Language Research*, pp.566-590. New York: Routledge.

A case of (2): William Stokoe (1960)

Sign language structure: An outline of the visual communication system of the American deaf. *Studies in Linguistics Occasional Papers* 8 & 7-79, SUNY at Buffalo.

수화언어의 어휘를 구성하는 기본단위 (2)는 아래의 (3)과 같이 소리언어에서 분절음의 음소를 규명하고 음소의 변별적인 자질을 분류할 때 사용하는 방법과 유사함

(3) a. The choice of articulator
 b. The place of articulation
 c. The manner of articulation

(4) a. 소리언어(spoken language)
 [] [] [] []
 ↓
 μ

b. 수화언어(sign language)
 [] (수형 Handshape)
 [] (이동 Movement)
 [] (위치 Location)

Hurst (1993: 210), μ = 형태소 (morpheme). Battison (1978)부터 손바닥 방향(orientation of the palm)도 독자적인 단위로 봄.

(4a) : 소리는 선형적 결합이지만 vs. (4b): 수화는 공간적 동시성 결합으로
 다중조음(multiple articulation)의 특성 vs. 다중조음(multiple articulation)의 특성
 → multi-linear phonological analysis in which different features may be placed on separate tiers → 대립 (opposition)과 대조(contrast)의 체계를 갖는 소리언어의 음운론에서의 변별적 자질과 같은 기능을 함
 → Autosegmental Phonology (Goldsmith 1979)
 CV phonology (Kahn 1980, McCarthy 1981)

Place	Labial		[bilabial]	p, b, m
		[labiodental]		f, v
Coronal	[dental]		θ, ð	
	[alveolar]		t, d, n, l, s, z	
	[post-alveolar]		r	
	[palato-alveolar]		ʃ, ʒ	
Dorsal		[velar]	k, g, ŋ	
Aperture	Fricative:		[stop]	p, t, k, h, d, g, m, n
			[vibrant]	r, l, r, ʒ
			[nonvibrant]	f, θ, v, ð
	Approximant Vowel	Height	[high]	i, u
			[mid-high]	e, o, eɪ, oʊ
			[mid]	a, ɔ
			[mid-low]	ɛ, ɔ
			[low]	æ, ɒ
			Rackness	[front]
	[back]	u, ɔ, oʊ, o, ɒ		
	[rounded]	ɔ, oʊ, u, ɒ		
	[unrounded]	i, e, ɛ, ɪ, æ, ɒ		
Nasality		[nasal]	m, n, ŋ	
		[oral]	(all others)	
Laterality		[lateral]	l	
		[central]	(all others)	
Laryngeal	Glottal Structure		[voiceless]	p, t, k, f, θ, s, ʃ
			[modal] voice	(all others)

음소 성운상태 조음위치 조음방법
Phoneme Voicing Place Manner

/p/ voiceless bilabial stop
 /b/ voiced bilabial stop
 /m/ voiced bilabial nasal
 /n/ voiced alveolar nasal

높이 전후위치 원순성
Height Backness Rounding

/i/ high front unrounded
 /æ/ low front unrounded
 /u/ high back rounded

최소대립쌍 (Minimal pairs)
 pet, bet, met, net
 pail, bail, mail, nail

/i/ vs. /æ/: beach, batch
 /i/ vs. /æ/: cheat, chat
 /i/ vs. /u/: cheese, choose

From Ladefoged and Johnson (2011: 278)

미국수어와 한국수어의 단일형태소 sign의 최소대립쌍

종소단위	미국수어(American Sign Language)	한국수어(Korean Sign Language)	
a. 수형 handshape			<p><오늘> 두 손을 펴서 손끝이 위로 손끝이 밖으로 향하게 하여 가슴 앞에서 약간 아래로 내린다.</p> <p><지금> 1 손 5지를 펴서 끝이 위로 끝이 밖으로 향하게 한 두 주먹을 아래로 약간 내린다.</p>
	CANDY	APPLE	
b. 위치 place or location			<p><추억> 오른 주먹의 1지를 펴서 끝을 오른쪽 관자놀이에 댄다가 두 번 구부렸다 펴며 밖으로 펼친다.</p> <p><옛> 오른 주먹의 1지를 펴서 끝을 코에 댄다가 천천히 구부렸다 폄다 하면서 밖으로 내린다.</p>
	SUMMER	UGLY	
c. 이동 movement			<p><교양어> 오른 주먹의 5지를 펴서 들어 입으로 향하게 하여 입 왼쪽으로 두 바퀴 돌린다.</p> <p><종역> 오른 주먹의 5지를 펴서 세워 입 앞에서 좌우로 흔든다.</p>
	TAPE	CHAIR	
d. 방향 orientation			<p><연습> 오른손을 구부려 손바닥이 밖으로 향하게 하여 입 앞에서 왼쪽으로 두 바퀴 돌린다.</p> <p><철안> 오른손의 손가락을 벌려 한글 구부려 입 가까이 대고 두 번 돌린 다음, (주먹의 4 5지를 펴서 끝이 위로 향하게 맞대고 세워 입앞으로 펼친다.)</p>
	CHILD	THING	

출처: Klima and Bellugi (1979) *The Signs of Language*, Harvard Univ. Press, Cambridge, MA. 출처: 강세민 (1997) 수화언어와 음성음성현상, 언어학 21, 263-267, 한국언어학회.

WORK_v / JOB TALK_v ON TRUE
 MIND_v HEALTH / WELL BROTHER PAPER

Figure 25.2 Minimal pairs in BSL (British Sign Language)

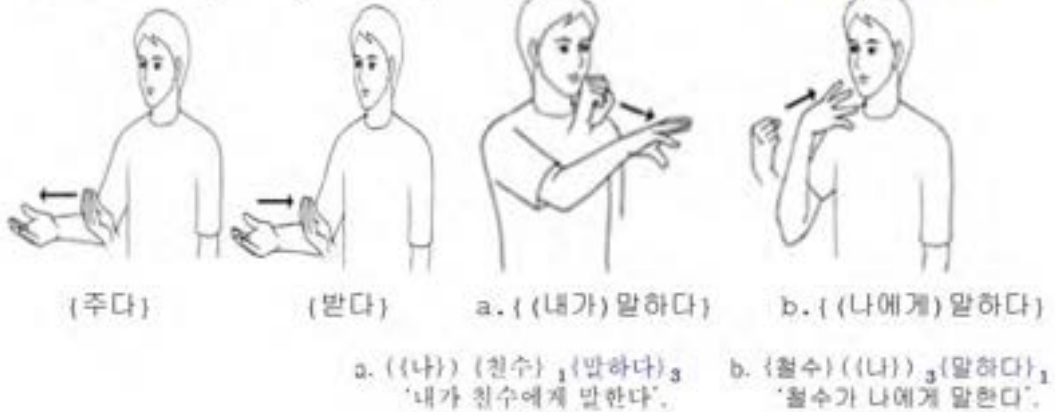
1. 굴절형태소 (Inflectional morphemes)



- 1.1. 동사 일치(verb agreement):** 주어와 목적어 논항의 문법적 공간 위치에 따라 이동방향이나 손바닥이나 손끝의 수향을 바꾸는 '굴절동사'가 갖는 인칭 일치를 말함
 ((내가) 주다). ((내가) 받다), ((내가)말하다. (나에게) 말하다)
- 1.2. 수의 일치(number agreement):** 동사가 주어나 목적어의 수(복수)에 일치하는 것으로, 주어 또는 목적어의 수(하나 또는 여럿)에 따라 반복 이동을 보이는 형태-동사적인 현상

1.1. 동사 일치(verb agreement):

주의: {못생기다}




31

1.2. 수의 일치




32



2. 어형성 조어법

2.1 품사가 변하지 않는 접사 (접두사) Derivants I: class-maintaining affixation


2.1.1 파생접사 (접두사) a. 비우세손 접두사: {#말-} / b. 한글/한자 지문자형 접두사: {#외}, {#공(공)-}



{말} + {아들} {말말}

(말어플) 손바닥의 방향 중요(아플은 2지(대구방안 경우 5지)를 피서 손바닥이 위로 향하게 하여 2지(대구방안 경우 5지) 끝이 비우세손의 5지에 연다)

Q1: 파생접두사? 비우세손 5지 끝 + 접촉
Q2: 왜 {아들} 5지 수형이 아니고 2지를 사용하나? 손구조상 접촉이 불가능하므로 대체할 수 있는{행}, {동성} 수형에 2지를 사용




한글 지문자 접두사 {의} + {말아버지}

한자 지문자 접두사 {공} + {사무} → {공무원}

접두사 {공} 비스듬히 세운 왼손 밑에 오른손으로 '스' 모양을 그린다.

{사무} 주먹을 쥔 왼 팔뚝에 1 5지 끝을 맞춘 오른 주먹을 올려 놓고 전후로 움직인다

33

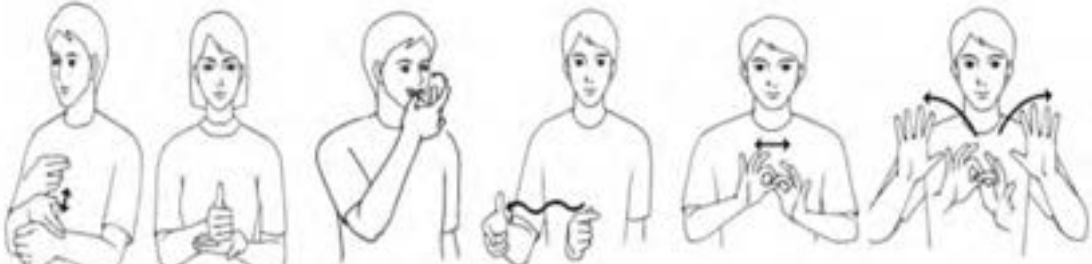


2.1.2 품사가 변하지 않는 접사: 파생접사 (접미사)

a. 존칭 접미사: {#-님}

b. 범주 접미사: {#-여러 가지}

c. 부정 접미사: {#-없다}



{선생} + {님} {관} + {계} → {무정} + {관계}

{관} + {계} → {무정} + {관계}

{관계} 두 손의 1 5지로 고리를 만들어 끼고 좌우로 두 번 움직인다.

{관계없다} 두 손의 1 5지 끝을 맞대어 고리를 만들어 끼었다가 고리를 풀며 두 손을 손바닥이 밖으로 향하게 양옆으로 벌린다.

34

한국해양대학교

범주 접미사의 다른 예 : (#-여러 가지)

{사과}+{ 여러 가지 } 오른 손바닥으로 왼 주먹을 한 번 감싸 돌린 다음, 양손(또는 한손) 주먹의 1·5지를 펴서 모로 세워 좌우로 약간 흔들며 바깥쪽으로 이동시킨다.

35

한국해양대학교

2.2 품사가 변하는 접사 (Derivants II: class-changing affixation)

2.2.1 반복이동: 파생접미사 (동작만 반복한 것(동사) → 명사(연제) 아니면 모든 수화가 반복되는지?)

명사: {의자} :
repeated movement
왼 주먹의 1·2지를 펴서 끝이 오른쪽으로 움직이 위로 향하게 하고, 오른 주먹의 1·2지를 펴서 약간 구부려 바닥을 왼 주먹의 1·2지 등에 두 번 낸다.

{침대}
repeated movement
손바닥이 위로 향하게 한 왼 손바닥에 오른 주먹의 1·2지를 펴서 끝을 대고 두 손을 동시에 상하로 두 번 움직인다.

{먹다}
[동사 + repeated movement 접미사] v → [명사]N

{밥}

동사: {앉다} :
single movement

{눅다}
single movement
손바닥이 위로 향하게 한 왼 손바닥에 오른 주먹의 1·2지를 펴서 들을 낸다.

36

2.2.2 **완전중첩 1(Total Reduplication)** → 품사 변할 수도 있음 (부사 → 부사/명사/동사)

{반복, 거둬, 자꾸, 자주, 잦다}

Q: 접두사로 쓰이는 구속형태소?



{또}+{또} 손등이 밖으로 향하게 세운 오른 주먹을 왼쪽으로 이동시키며 1·2지를 펴는 동작을 두 번 반복한다.



{반복} 오른손을 구부려 손바닥이 밖으로 향하게 하여 오른쪽으로 두 번 반원을 그리며 내린다.

[같은 모양이 반복되는 것을 나타내는 동작]

완전중첩 2: 3번이상 중첩현상이 원형 이동으로 복수를 단순화



{결석}

{결석} 손끝이 밖으로 손등이 위로 향하게 편 두 손을 양옆에서 중앙으로 이동시켜 맞댄다



+ {결석} + {결석} ...



= {방학}


{방학} 손등이 위로 손끝이 밖으로 향하게 편 두 손을 마주 대고 왼쪽에서 오른쪽으로 (원형모양으로) 움직인다.

2.3 단어어와 합성어 한국국립대학교


+

➔

{(붉은)사과}
'apple'

[빨강] 오른 주먹의 1지를 펴서 옆면으로 입술 밑을 오른쪽으로 스쳐 낸다. [사과] 모로 세운 왼 주먹 등을 오른 손바닥으로 닦아 내린다


+

➔

{풋사과}

[파랑] 오른 손바닥으로 오른쪽을 스쳐 올린 다음, 모로 세운 왼 주먹 등을 오른 손바닥으로 닦아 내린다.

39

한국국립대학교

•2.3.1 병렬(대등)합성어 : 두 단어가 원래의 의미를 유지하면서 동등한 자격으로 연결되는 것. 예) 마소(말+소), 흑백, 높푸르다(높고푸르다), 오가다(오고가다) 등

• 2.3.2 종속합성어 : 두 단어가 서로 주종관계로 연결되는 것. 예) 소나무 (나무는 나무인데, 숲나무 - 나무가 주, 숲이 종) 국밥 (밥은 밥인데, 국에 알은 밥 - 밥이 주, 국이 종) 봄비, 꽃게, 똥다리 등



주의: 입에 대면 (흑스)

[검은색] 오른 손바닥으로 머리 오른쪽을 스쳐 내린 다음, 오른손의 손가락을 구부려 턱에 대고 오른쪽에서 왼쪽으로 좌우로 두 번 움직인다.



[비밀법] 모로 세운 오른 주먹을 왼 손바닥에 올려 놓고 좌우로 움직인 다음, 손바닥이 위로 손끝이 왼쪽으로 향하게 면 오른손을 입으로 두 번 올린다.

40

•2.3.3 융합합성어 : 두 단어의 본래의 뜻은 사라지고, 새로운 하나의 뜻이 생겨나는 합성어. 예) 밤낮 → 항상, 늙 / 춘추 → 나이(연세), / 세월 → 시간, / 넘어가다 → 속다, / 보릿고개 → 매우 어려운 시절



{파랗다}

+



{밝다/넓게 펼치다}



{하늘}

오른 손바닥으로 오른쪽을 스쳐 내린 다음, 두 손을 펴서 손바닥이 밖으로 손끝이 위로 향하게 하여 얼굴 앞에서 교차하게 하였다가 양옆으로 벌린다.



{검다}

+



{희다}



{두다} 바둑을 두는 동작



{바둑}

오른손으로 머리 오른쪽을 스쳐내리고, 오른 주먹의 1지를 펴서 끝으로 치아를 가리킨 다음, 오른 주먹의 1-2지를 펴서 끝바닥을 왼 손바닥에 두 번 낸다.

2.3.4. 융합합성어를 사용한 관용적 표현



{불편} 오른 주먹의 5지를 펴서 세워 두 번 구부린다

+



{귀신} 손등이 위로 향하게 벌려 편 왼 손등에 오른손의 2지를 대고 두 손을 동시에 위로 빠르게 약간 올린다. (비수지동반: 입을 벌려 눈을 크게 뜨고 눈썹을 치켜올림)



{감쪽같이 사라지다}



{상각} 오른 주먹의 1지를 펴서 오른쪽 관자놀이에 대고 좌우로 약간 돌린다


+

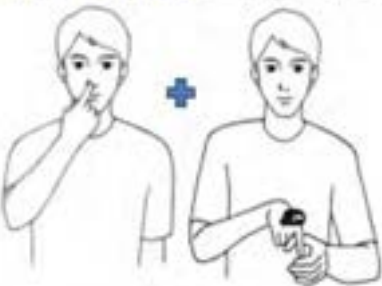


{박수} 두 손바닥이 X자로 맞닿게 박수를 친다



{너가 알아서 해}

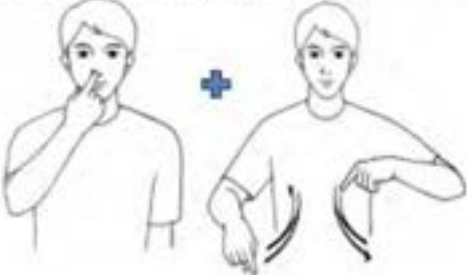
3. 기타 조어방법 (Coinage: 새로운 단어가 만들어지는 방법): 음차 (#코-) 



{코} + {차}{물잔에 젖는 동작} = {커피}

오른 주먹의 1지를 펴서 끝 바닥을 코에 닿 다음, 그 끝을 둥글게 구부려 모로 세운 왼손 안에 넣어 돌린다.

[코(숨을 쉬는 기관 '코'를 coffee의 'CO'로 빌려 씀)·차]

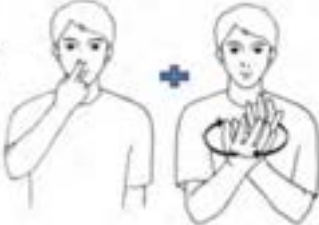


{코} + {다스리다} = {코치}

{다스리다} 두 주먹의 1지를 펴서 끝이 약간 아래로 향하게 하여 좌우로 엇갈리게 움직인다.


[코(숨을 쉬는 기관 '코'를 '코치'의 '코'로 빌려 씀)·다스리다]

참고: {코스모스} = {코} + {꽃}



{꽃} 약간 구부린 두 손의 손목을 'X'자로 맞댔다가 손목을 천천히 돌려 두 손 바닥이 마주 보게 한다.

43

동시적으로 결합된 합성어 



←



{산} 오른 주먹의 2지를 펴서 등이 밖으로 향하게 세운다.

+



{폭발} 두 손의 손가락을 오므려 맞댔다가 위로 빠르게 울리며 활짝 편다.

44



Figure 25.3 Movement contrast between the derivationally related BSL signs KEY and LOCK.

Source: Woll, Bencie and David Vinson (2011) Chapter 25 Gesture and Sign: Theoretical and experimental perspectives, In Josep Quer, Roland Pfau, and Annika Herrmann (eds.) *The Routledge Handbook of Theoretical and Experimental Sign Language Research*, pp.566-590. New York: Routledge.

KSL 어순

출처: 이영재(2010a: 40-45)

<나>



<빵>



<먹다>



(i) SOV 어순

- a. 나는 빵을 먹었다. (황도순 1994)
'I ate the bread.'
- b. 나의 빵을 먹었다. (이영재 2010a)
'(pro) ate my bread.'

a'. <빵> <먹다> <누구> <나>

a''. <빵> <먹다> <나>



(ii) OVS 어순:

pronoun copy의 경우



한국국립대학교

(iii) OSV 어순

(22) ㄱ. 개가 고양이를 쫓다.
 ㄴ. *<개> <고양이> <쫓다>
 ㄷ. <개a><IXa> [indexing by left to right]
 S
 <고양이b><IXb> [indexing by right to left]
 O
 <쫓다>
 V

ㄹ. <고양이a><IXa> [indexing by right to left]
 O
 <개b><IXb> [indexing by left to right]
 S
 <쫓다>
 V

ㅁ. <고양이><쫓다><무엇><개>

<쫓다>

한국국립대학교

Q: 수화언어에서 제스처(gesture)와 수화(sign)를 어떻게 구별하는가?

기본 질문에 대한 해답을 찾고자 KSL 인칭대명사를 중심으로 논의

연구 질문

Q1: KSL 인칭대명사에서 사용되는 수형은 한국어 인칭대명사 제스처와는 어떤 관계?
 Q2: KSL 인칭대명사의 수화소들이 인칭, 수, 성의 문법형태소들과는 어떤 연관성?
 Q3: KSL 일인칭대명사의 변이형은 무엇이고 왜 이런 변이형을 갖고 있는가?

형식언어학과 인지심리학에서는 제스처와 수화 구별을 위해 어떤 서로 다른 견해를 제시하고 있고 어떠한 논쟁을 하고 있으며 KSL 인칭대명사의 특성에서 어떤 견해가 우위에 있다는 증거는 무엇인지에 대하여 예를 들어 논의

※ 한국수어학회 발표 (2013. 8. 15)
 ※ 서울대학교 언어학과 수어언어학 10주차 수업자료 (2021. 5/2022. 5.)

40

초록

수화언어에서 *gesture*와 *sign*은 어떻게 구별되나? 한국수어의 인칭대명사를 중심으로 다음과 같은 세 개의 연구 질문을 통하여 논의하고 하고자 한다. (1) 한국수어의 인칭대명사에서 사용되는 수형은 한국어 인칭대명사 제스처와는 어떤 관계인가? (2) 한국수어의 인칭대명사의 수화소들이 인칭, 수, 성의 문법형태소들과는 어떤 연관성을 갖는가? (3) KSL 일인칭대명사의 변이형은 무엇이고 왜 이런 변이형을 갖고 있는가? 특히 형식언어학과 인지심리학에서는 제스처와 수화 구별을 위해 어떤 서로 다른 견해를 제시하고 있고 어떠한 논쟁을 하고 있으며 한국수어의 인칭대명사의 특성에서 어떤 견해가 우위에 있다는 증거는 무엇인지에 대하여 예를 들어 논의한다. 본 연구의 결과 우리는 음성 동반 제스처 (*co-speech gesture*)의 문법화에 대한 체계적인 연구가 필요하고 한국수어의 인칭대명사의 기능과 형식 및 체계에 대한 심도 있는 논의가 필요하다고 주장한다.

키워드: 수화언어, 한국수어, 인칭대명사, 수화, 음성동반 제스처, 문법화

49

1. 연구질문
2. 음성(Speech), 제스처(Gesture), 수화(Sign)
 - 2.1 사전적인 정의
 - 2.2 음성 동반 제스처(Co-speech Gesture)의 네가지 유형
 - 2.3 다섯 번째 제스처의 유형
 - 2.3.1 제스처의 연속체: 표상의 위치
 - 2.3.2 표상제스처의 네 가지 예: Thumbs-up, ASL V-수형, OK, Bull horns
 - 2.4 제스처(Gesture)와 수화(Sign)의 구별
 - 2.4.1 형식언어학의 관점 대 인지심리학적 관점
 - 2.4.2 제스처와 수화(Gesture and Sign) 대 제스처를 수화로(Gesture to Sign)
 - 2.5 학자별 음성, 제스처, 수화의 주요 키워드 비교
 - 2.6 언어이론에서 도상성과 지시성
 - 2.6.1 도상제스처 (음성제스처 포함) 대 도상수화
 - 2.6.2 지시제스처 대 지시수화
 - 2.7 문법화
 - 2.7.1 단일 경로 대 두 개의 경로
 - 2.7.2 문법화 연쇄과정
3. 수화언어 인칭대명사의 특징
4. KSL 인칭대명사의 제스처와 수화
 - 4.1 KSL 인칭대명사 수형과 Korean 인칭대명사의 음성 동반 제스처와의 비교
 - 4.2 KSL 인칭대명사의 지시(pointing)는 제스처인가 수화인가?
 - 4.3 KSL 인칭대명사의 수형 변이형
 - 4.4 KSL 인칭대명사의 위치 변이형
5. 결론과 제언
- References

50

1. 연구질문

수어언어학10주차

- (1) 건청인은 말을 할 때 gesture를 사용한다. 그러면 농인은 수어를 할 때 gesture를 사용할까? “예”라고 답하면, 수어를 할 때 동반된 gesture (“co-sign” gesture)는 말을 할 때 동반된 gesture(co-speech gesture)와는 구별될까? 즉, 수화언어에서 gesture와 sign은 어떻게 구별되나?
- (2) 한국수화언어(KSL)의 인칭대명사에서 사용되는 수형 1-형()과 9-형()은 한국어(Korean) 인칭대명사 수형 gesture(제스처)와는 어떤 관계를 가질까?
- (3) KSL의 인칭대명사를 나타내는 sign(형식)은 어떻게 될까? 즉, 수형(handshape), 수위(location), 수동(movement), 수향(orientation) 등의 음운적 요소인 수화소들이 인칭(person), 수(number), 성(gender) 등의 문법형태소들과 어떤 연관을 갖고 있을까?
- (4) KSL 일인칭대명사는 얼마나 많은 변이형을 갖고 있고 왜 그런 현상이 일어날까?

91

2.2 음성동반제스처(co-speech gesture)의 네 가지 유형


수어언어학10주차

- 말을 할 때 동반되는 gesture의 네 가지 주요 유형 (McNeill, 1992)



- (1) **Iconic gestures**, which **imagistically** represent object attributes, actions and spatial relationships.
[대상, 속성, 행위, 공간적인 관계를 형상화하여 나타내는 도상 제스처]
- (2) **Deictic, or pointing, gestures**, which index, or connect, some aspect of speech to some other **idea**, object, location or action.
[말의 어떤 양상을 어떤 다른 생각, 대상, 위치 또는 행위에 색인을 하거나 연결하는 지시 제스처]
- (3) **Metaphoric gestures**, which convey an **abstract idea** in a concrete form.
[추상적인 생각을 구체적인 형태로 전달하는 은유 제스처]
- (4) **Beat gestures**, which are hand movements that keep the **rhythm** of speech.
[말의 리듬을 유지하는 손 동작인 박자 제스처]

92

2.3 다섯 번째 주요 제스처

 수어언어학10주차

(5) **Emblems**, such as the 'thumbs up' or 'peace sign', are hand configurations that have a culturally specified meaning, and unlike co-speech gesture, they convey information independently of speech. (Kelly, Manning, and Podak, 2008)

[ 이니  와 같은 표상은 문화적으로 명시된 의미를 갖는 수형이고 음성이 동반되는 제스처와는 달리 음성과 관계없이 정보를 전달해준다]

2.3.1 제스처의 연속체: 표상의 위치

언어의 특성 존재 변명 (Kendon Continuum: McNeill, 1992: 3)

출처: McCleary & Viotti (2010: 181-182)

- linguistic properties (분절적, 범주적, 결합적, 선형적, 규약적, 계층적 구성)
discrete, categorical, combinatorial, linear, conventional, hierarchically organized
- gestural properties (연속적, 비범주적, 비결합적, 공간적, 특이적, 비계층적)
continuous, analogical, noncombinatorial, spatial, idiosyncratic, nonhierarchical

Gesticulation — **Pantomime** — **Emblems** — **Sign Language**

언어적 특성 결여

어느 정도 있음

비제스처로 풍부한 언어적 특성

음성 동반 제스처(co-speech gesture)

- ★ 음성이 선택적인 경우: Gesticulation — Emblems 의 순서
- ★ 음성이 없는 경우: Gesticulation — (Pantomime, Sign Language)의 순서

53

2.3.2 표상 제스처의 네 가지 예

 수어언어학10주차

(Four Examples of Emblematic gestures)

Emblematic gestures easily misunderstood abroad

[문화적으로 잘못 이해하기 쉬운 표상 제스처]

(<http://matadomnetwork.com/abroad/10-common-gestures-easily-misunderstood-abroad/>)

(1) Thumbs-up in Thailand



그림1 agreement or approval (동의, 만족, 승인, 좋아, 잘했어라는 표상)

It's typically a child's gesture, the Thai equivalent of sticking out your tongue(혀를 내밀), which means a sign of condemnation (비난).

54

(2) V in England/North America and Australia

With the palm facing **outward** (손바닥이 밖으로 보여도쪽), this V means victory in England, or peace in North America. (승리, 평화)



그림2 Jodie Foster flashes V for victory after winning the 61st Annual Academy Awards, 1989. Photo by Alan Light

In 1992, George Bush visited Australia and, from the window of his limousine, he gestured with his palm facing **inward** (손바닥이 안쪽으로 보여도쪽), the nonverbal equivalent of "up yours."



그림3 미국수화 수지알파벳 V

up yours (속어) 「항오-한항 따위를 나타내어」 그 따위 얄게 뛰아, 엉망할, 밧어먹을, 찌기할.

Other emblematic gestures easily misunderstood abroad

(3) OK in France

- ◆ Making a circle with your thumb and forefinger means "great" or "fine" in North America. It's also used by scuba divers to communicate that there are no problems.
- ◆ In **France**, however, this gesture means "zero." A bad idea when trying to compliment a chef on your meal.



그림3 OK 표상 제스처

(4) Bull Horns in Italy and Norway

- ◆ North Americans raise their index and pinkie fingers like bull horns when they want to rock and roll all night... or cheer on a sports team with a name like "Cowboys" or "Longhorns."
- ◆ In **Italy**, this "cuckold" gesture (부정한 아내의 남편) means that a man's wife is being unfaithful, and he is a fool because of it.
- ◆ In **Norway**, satanic associations (사탄과 같은 악마를 찬양함) In 2005, George W. Bush visited the University of Texas at Austin.



그림4 소뿔 표상 제스처

2.4 제스처와 수화 구별 선행연구 1

2.4.1 형식언어학 대 인지심리학 관점

(1) 형식언어학의 관점: 서로 다른 범주로 본 최근 연구

- Emmorey (1999) Do signers gesture? ➡ 논문개요와 최근 세미나 소개: 다음 슬라이드 참조
- Duncan (2005) Gesture in signing : A case study from Taiwan Sign Language

▷ Sign: 언어적인(linguistic) 어떤 것

- 음운구조 (Sandler, 2012)
- 형태구조 (Padden et al, 2010)
- 운율과 통사구조 (Sandler et al, 2011)

▷ Gesture: 언어 밖에 있는 비언어적(nonlinguistic)인 것으로 준언어적(paralinguistic) 또는 비언어적 의사소통(nonverbal communication)의 일부분으로 봄

질문: gesture를 형태소와 단어의 형식 언어적 단위와 대립하는 준언어적요소로 본다면 gesture와 sign이라는 두 개의 의사소통 유형이 동일한 조음자들을 갖고 있는 수화언어에서는 무엇이 gesture이고 무엇이 언어적(또는 형태소)인가를 어떻게 정의할 것인가? (Quinto-Pozos, 2002)

[참고] Karen Emmorey 논문과 2013년 세미나 요약

- Emmorey (1999) Do signers gesture?

Prof. **Karen Emmorey**, School of Speech, Language, and Hearing Sciences, San Diego State University, USA.

- Seminar at UBC, Canada on **January 17, 2013**

- Karen Emmorey explores the relation between gesture and sign language from linguistic, psycholinguistic, and cognitive neuroscience perspectives. **"Co-sign" gesture** differs from **co-speech gesture** with respect to the primary articulators (the face/body vs. the hands), the optionality of gesture, and types of gestures that are possible. She also presents the results of studies with bilinguals who are fluent in a signed and a spoken language, which indicate knowledge of American Sign Language (ASL) can impact the form and content of co-speech gesture. Finally, she presents neuroimaging results that reveal **differences between** the neural systems that support gesture (pantomime) **and** sign production, as well as **differences between** the neural regions engaged when deaf signers vs. hearing nonsigners view meaningless **and** meaningful gestures.

제스처와 수화의 구별 선행연구 2

(2) 인지심리학적 관점: 명확하게 구별하지 않음

- Kendon (2004, 2008) : Visible action as utterance

형식언어학의 관점에서 수화언어를 분석한 결과 **gesture**와 **sign**은 서로 대립적인 관계에 있는 것처럼 서로 다른 범주로 구분한 것은 이들의 차이점을 과장한 것이고 서로의 중복성의 영역을 모호하게 하는 경향이 있다. 그래서 **gesture**의 인지심리학적 관점에서 **시각신체행위 (visible bodily action)**가 발화의 구성에 사용되는 서로 다른 방법들에 대한 비교연구가 필요하다고 보고 시각신체행위가 기호학적 관점에서 기능하는 다양한 방법을 모색해야 한다고 주장함.

- Wilcox and Xavier (2013) : A framework for unifying spoken language, signed language, and gesture

gesture와 **sign** 사이의 경계가 흐려있기 때문에 제스처는 다양한 방법으로 수화언어에 병합되어진다.

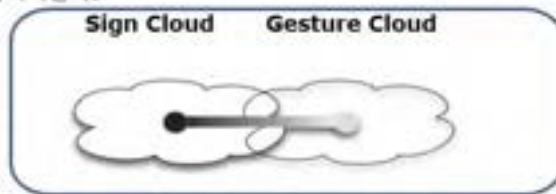


그림4 Sign and cloud continua (sign과 구름 연속체)

59

두 가지 견해 요약

2.4.2 제스처와 수화(gesture and sign) vs. 제스처를 수화로(gesture to sign)

<Gesture와 Sign이 동시에 또는 공시적으로 상호작용되는 두 가지 견해 요약>

gesture and sign

- 서로 다른 범주로 봄
- gesture와 sign은 한꺼번에 또는 교체적으로 같이 발생할 수 있음

gesture to sign

- 제스처가 수화언어체계에 병합
- 전형적인 예는 통시적인 현상인 어휘화(lexicalization)와 문법화(grammaticalization)
- gesture와 sign은 상호관계에 있고 끊임 없는 변화를 겪고 있어 오늘날 gesture로 알고 있는 것이 미래에는 수화언어의 언어적인 요소가 될 수 있음

60

2.5 학자별 음성, 제스처, 수화의 비교

표1. 학자별 음성, 제스처, 수화의 주요 키워드 비교

	Speech	Gesture	Sign
Goldin-Meadow (1999)	analytic, segmented, combinatorial	analogous, mimetic	
Cicourel(2002)	categorical, morphemic, conventionalized	gradient, analogous, unconventionlized	Gesture and Sign
	combinational, categorical	noncombinational, gradient	
Little(2003)	discrete, countable	analogous, uncountable	
Kelly, Mirzang & Podak(2008)	conventionalized, arbitrary	idiosyncratic, analogic	
McCleary & Votts(2010)	discrete, categorical, combinatorial, linear, conventional, hierarchical, organized	continuous, analogical, noncombinatorial, spatial, idiosyncratic, nonhierarchical	Gesture to Sign
	gradiently patterned phenomena		
Duncan(2003)	intonation	analogous gesture	
	tempo	analogic conceptualization	
	loudness	expressions of affect and arousal	
	voice quality	discourse prosody	

61

2.6 언어이론에서 도상성(iconicity)과 지시성(pointing)

2.6.1 도상제스처(iconic gestures) 대 도상수화(iconic signs) (Pfau et al., 2012: 394-396)

	iconic gestures	iconic signs
Similarity	Both involve structure-preserving mappings between form and referent.	
Differences	<ul style="list-style-type: none"> They are not bound by linguistic constraints on form. They tend to represent a specific action at a specific time and place. They are interpreted as meaningful via an on-line conceptual blending process. The iconic gesture is embedded in a particular discourse event; it could not be interpreted if removed from its context. (그림 5) 	<ul style="list-style-type: none"> They obey the phonotactic constraints of the respective sign language. They denote a concept rather than a specific event. They have a directly accessible, memorized meaning. Though the iconic sign is very similar to the gesture, it is fully conventional and comprehensible in the absence of context. (그림 7)

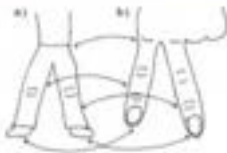


그림5 Structure-preserving correspondences between a) human legs and b) extended index and middle fingers (V-handshape) to mean 'two-legged entity'.



그림6 바나나 껍질을 벗기는 도상제스처

그림7 ASL BANANA 수화 (바나나를 벗기는 특정사람의 image 또는 특정행위를 나타내는 것이 아니라 과일의 한 종류인 바나나의 개념을 나타낸다.)

62

음성제스처(spoken gesture)와 도상 제스처(iconic gestures)

◇ 음성제스처(spoken gestures) = 도상제스처(iconic gestures) (Okrent 2002: 187-188)

Ex1) It was a **loooooong** time. [장모음화(vowel lengthening)]

- 위의 장모음화는 단어 long의 음소자질도 음소배열제약(phonotactic constraint)의 결과도 아닐.
- 한정된 음성자질의 집합에서 [+long]이라는 자질을 선택한 결과가 아닐.
- 위의 장음화는 모음의 규범화된 형식과 동시조음으로 새롭게 만들어진 형태를 통한 형상화(imagery)의 표현임.

Ex2) The bird flew up [high pitch] and down [low pitch]. [높낮이(pitch)]

- 위의 성조는 단어 up, down의 음소자질이 아닐.
- up에 높은 성조는 높은 공간 위치서 고주파수와 연관된 음유를 통하여 높음의 형상화를 표현하고, down에 낮은 성조는 위의 음유의 다른 면을 통한 낮음의 형상화를 표현한다.
- 위의 성조의 높낮이는 고정된 어휘항목과 동시에 조음되는 새롭게 만들어진 형태를 통한 형상화(imagery)의 표현임.



- 화자들은 동일한 modality에서, 동일한 조음자료, 언어적인 것과 제스처적인 것 중시 표현
- 소리언어단위는 음성제스처(spoken gestures)와 동시에 조음할 수 있다. (이와 같은 원리로)
 - 수화도 수제스처(manual gestures)와 동시에 조음할 수 있다.

- Liddell (2003): 일치동사(agreement verbs)는 (위치자질 또는 동사가 지시하는 곳과 같은) **지시 제스처**와 동시에 조음된 (수형, 장형 등의 자질과 같은) 언어단위라고 주장.

지시 제스처(pointing gesture)와 지시수화(pointing sign)

2.6.2 지시제스처(Pointing gesture) vs. 지시수화(pointing sign)

Q1: 수화언어의 인정대행사는 지시제스처(pointing gesture)인가? 지시수화(pointing sign)인가?

Q2: 지시 때 발생하는 이동(movement)의 방향과 목표가 있는 공간장소는 제스처인가? 또는 언어적 단위가 되는가?

(Spatial loci are gestural (non-linguistic) or linguistic?)

지시수화(pointing signs)에 대한 논쟁 (Pflau and Steinbach 2006: 63)

- Liddell (2000, 2003): 공간장소는 제스처 성분이다. ➡ pointing gesture
- Meier (1990, 2002, 2011): 공간장소는 언어적 단위이다. ➡ Pointing sign
- Pflau and Steinbach (2006): 음성동반 지시제스처(co-speech pointing gestures)는 위치지시를 위해 사용. 지시제스처는 건청어동과 농어동 모두 초기 습득(대명사 사용의 오류 발생은 도상제스처로 설명하지 못할)하는 것이어서 지시제스처가 위치표지(locative marker)로서의 수화언어 문법체계에서 행해지는 문법화과정으로 설명

◇ 지시수화(pointing sign)의 사용과 문법은 언어적인 원리(문법적 체계의 일부분)가 작용한다. 그래서 대명사와 일치동사가 보여주는 일정한 언어적 패턴은 공간장소가 비언어적인 요소라고 가정하면 설명할 수 없다.

- 서로 다른 지시수화의 분포와 관련하여 언어개별맥락이 작용되는 예:

주어생략, 주어대행사 복사 (Pro-drop, subject pronoun copy: Padden (1988), 이영재 (2009: 51), 장세은 (2011: 92))



"나는 뺨을 먹었다"

2.7 문법화(grammaticalization)

수어언어학10주차

- 문법화(grammaticalization): 다음의 두 과정에 의한 통시적 과정
 - 명사나 동사와 같은 어휘형태소가 시간이 경과함에 따라 문법형태소로 발전
 - 조동사와 같이 일 문법적인 형태소는 시제나 양상 표시와 같은 더 문법적인 형태소로 발전 (Bybee et al. 1994)

2.7.1 단일경로와 두 개의 경로

(1) 일반화된 단일 경로 (Shafer, 2000; Jansen & Shafer, 2002)



예) Jansen & Shafer (2002)

ASL		고대불안사수화
FUTURE	제스처 'to go'	→ 동사 'to go' → 문법형태소 'future'
CAN	제스처 'strong'	→ 어휘 'strong' → 문법형태소 'can'
MUST	제스처 'owe'	→ 동사 'owe' → 불안사수화/ASL 문법형태소 'must', 'should'

(2) Wilcox(2009)가 제시한 문법화와 두 개의 경로:

- 음성이 동반되는 시각제스처(co-speech visible gesture)는 수화언어체에 병합(incorporation), 법제화(codification), 마침내 문법화(grammaticalization)되는 sign의 자연적인 근원이다.

- 첫 번째 경로: 제스처 → 어휘형태소 → 문법형태소 (수지적 제스처가 근원인 경우)



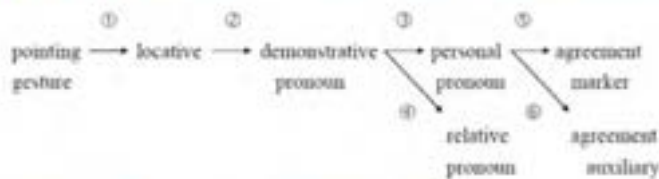
- 두 번째 경로: 제스처 → 문물/역량 → 문법표지 (이동법법, 얼굴, 입, 눈의 제스처 등의 수지적 제스처가 만들어지는 방법이 근원인 경우)

65

문법화 연쇄과정

수어언어학10주차

2.7.2 문법화 연쇄과정 [Grammaticalization chain: Pfau and Steinbach (2006), Pfau, et al. (2012)]



- 소리언어에서 흔히 발견되는 유형 (③ 단계)

예: Turkish의 3인칭대명사의 발달과정

- INDEX가 목표를 하는 수화공간에 있는 위치가 이동의 시작점과 끝지점을 정의함으로써 수화언어 동사일치에 중요한 역할을 한다면 수화언어의 일치표지는 대명사에서 발달되었다고 볼 수 있다. (⑤ 단계)

예: Use of space in DGS (독일수화언어): 대명사 겹사로서 동사 일치자질 (Keller 1999)

66

3. 수화언어 인칭대명사의 특징 1

- (1) 수화언어의 인칭대명사는 지시수화(pointing sign)이다. (Sandler & Lillo-Martin 2006)
- (2) 소리언어와는 달리 대명사를 사용하면 모호성(ambiguity)이 만들어지지 않는다.
 - ◇ English
 - The governor, told the senator, that they_{1,2,3,4,5,6,7,8,9,10,11,12} should leave
 - ◇ ASL
 - (a) a-GOVERNOR TELL-b b-SENATOR c-IX SHOULD c-LEAVE
 - The governor, told the senator, that they_c should leave'
 - (b) a-GOVERNOR TELL-b b-SENATOR b:c-IX SHOULD b:c-LEAVE
 - The governor, told the senator, that they_{b,c} should leave'
 - (c) a-GOVERNOR TELL-b b-SENATOR a:b-IX SHOULD a:b-LEAVE
 - The governor, told the senator, that they_{a,b} should leave'
- (3) 소리언어와는 달리 지시적 색인(referential indices)은 수화공간에서 서로 다른 위치를 나타내는 형태로 명시적으로 실현된다. (Sandler & Lillo-Martin 2006: 378)
- (4) 수(number, 단수, 이중수, 삼중수, ... 다중수 등)에 대한 표지를 보여준다.
- (5) 수화언어에서 인칭은 논쟁(three-person vs. two-person system)의 대상 NEXT

3. 수화언어 인칭대명사 특징2: 인칭체계

- ◇ 소리언어에서 대화역할(speech roles)에 따른 인칭(person) 정의 (Shat 2004)
 - being the speaker 1st person : 수화자에게 지시
 - being the addressee 2nd person : 듣는이에게 지시
 - non-speaker/non-addressee 3rd person : 실질적인 대상자가 없는 대화참여자에게 지시
- ◇ 수화언어에서 인칭은 논쟁의 대상
 - 인칭을 나타내는 방향성(directionality): 제스처 설명 vs. 문법적 설명 (일치agreement)
 - 세 개 인칭체계(three-person system)
 - 전통적으로 수화언어의 대명사체계는 소리언어의 대명사체계와 유사한 세 개의 인칭체계로 보는 견해 (Friedman 1975, Klima & Bellugi 1979, Berenz 2002)
 - 두 개 인칭체계(two-person system: first vs. non-first person pronoun)
 - 일인칭대명사는 **변어형이 없어 어휘부에 기재되는(istability)** 반면에, 비일인칭대명사는 **듣는어 또는 실질적인 대상자가 없는 대화참여자들에게 지시할 수 있는 수화공간은 열거하지 못할 정도로 공간 값이 열거 때문** (Padden 1990; Engberg-Pedersen 1993; Lillo-Martin 2002)
 - Liddele (2003): 수화언어 대명사는 제스처로(gesturally) 지시물(referents)을 지시(point)한다고 주장.
 - 수화언어대명사는 언어적 요소(수형과 수동과 같은 음운적으로 명시된 매개변수)와 제스처 요소(=수화의 방향성)의 융합(fusion)의 결과라고 봄.
 - Meier (1990): 그러나 방향성에 대한 제스처 설명은 1인칭 행위, 특히 **지시물에 대한 지시가 필수적이지 않은 1인칭 복수행위를** 설명하지 못한다. 이러한 점이 ASL에는 독립적인 1인칭이 존재한다는 주장.
 - 듣는어와 실질적인 대상자가 없는 대화참여자들을 언급하는 공간의 사용은 범주(categorical)가 아닌 **gradient**라고 주장.

3. 수화언어 인칭대명사 특징3: Pfau et al (2012)

- ◆ **위치:** 말인칭대명사는 수화자의 가슴 쪽으로 밖에서 안쪽으로 방향하여 지시하지만 수화자의 코 쪽으로 지시하는 수화언어도 있음(Japanese Sign Language, Plains Indian Sign Language, McBurney 2002).
- ◆ **공간:**
 - 수화자 주위의 공간은 담화상에서 대명사 지시(pronominal reference)를 위하여 사용된다.
 - 담화상에서 이러한 공간장소의 또 다른 용도는 일치동사/지시동사(agreement/indicating verbs)를 포함한다. 이러한 장치는 대명사화 또는 공간장소의 문법화된 형태(grammaticized forms of pronominalization or spatial loci)로 본다. (Pfau and Steinbach 2006)
- ◆ **지시대상물의 존재여부:**
 - 지시물이 존재하면 수화자는 대명사 또는 그 지시물의 위치로 지시하는 일치/지시 장치를 사용한다.
 - 그러나 지시물이 대화상에 존재하지 않으면 수화자는 그 지시물을 위한 어떤 공간에 어떤 지점에 어떤 사람을 임의적으로 정할 수 있다.
 - 어떤 지시물에 대한 어떤 위치가 정해지면 이문적 의미를 갖지 않게 계속적으로 언급될 수 있다.

예: 영국수화언어에서의 수화공간의 사용 (Pfau et al 2012: 229)

SISTER INDEX3a UPSET, INDEX1 IASK3a WHAT, INDEX3a LOSE BAG.
 sister there upset I I-ask-her what She lost bag
 'My sister was upset. I asked her what was wrong. She had lost her bag.'

4. KSL 인칭대명사의 제스처와 수화

4.1 KSL 인칭대명사 수형과 Korean 인칭대명사의 음성 동반 제스처와의 비교

- ◆ 최상배(2012)의 고빈도 수형(분석대상 단어수 9,557, 전체수형 27,675) 대 ASL의 무표수형(Battison 1978)

	1	2	3	4	5	6	7
KSL 수형어휘	9형	1형	주먹형	지성형	편손형	5형	부리형
빈도수	6,438	2,561	2,511	1,879	1,539	1,538	920
%	23.3	9.3	9.1	6.8	5.6	5.6	3.3
ASL 수형어휘	A	S	O	C	G	B	5
KSL 수형어휘	9형	주먹형	대화형	집형	1형	9형	편손형

- ◆ 국립국어원 한국어수화사전 홈페이지 (<http://222.122.196.110>)에서 제공하는 수화 단어를 대상

- ◆ Korean 인칭대명사의 음성 동반 제스처

수형어휘	(9형)	(1형)	(5형)	(여자형)
1,2,3인칭 자신이나 자신보다 신분이 높은 사람지칭	1,2,3인칭 자신이나 자신보다 낮은 신분의 사람 또는 사물 지칭	1,2,3인칭 지시하는 양쪽의 공통형을 나타냄	3인칭의 여자친구나 애인을 지칭할때	

KSL 인칭대명사의 지시(pointing)

4.2 KSL 인칭대명사의 지시(pointing)는 제스처인가 수화인가?

- 세 개 인칭체계 (Lee, 2009: 55-56)



- 지시수화(pointing signs)라는 증거

(Petersmutter 1991; Meier & Lillo-Martin 2010; 음성언어의 대명사체계 특성을 수화언어도 갖고 있다는 논의)

- (1) 대명사 어휘 항목의 약정성 (conventionality):


- 어느 특정 집단에 사용되는 모든 사람들의 의사소통의 행위의 한 약정된 부분의 예:



(여우형) '여우, 땡계, 교활, 왕궁' 단어에 사용되는 수형
(소불 표상) '텍사스대학 상징적 표상'
'악마 표상'

지시수화(pointing signs)라는 증거

- 음성언어에서 단어와 그 의미간의 관계는 약정적(conventional)이고 임의적(arbitrary)인 것과 같이 수화언어의 지시수화가 형식과 의미에서 어느정도까지 약정적일까?

- (a) ASL는 인칭대명사의 지시수화가  으로 비언어적 제스처와 동일하고 한 개의 수형을 가지고 있어 지시수화에 대한 약정성에 대한 증거가 약하지만, KSL에서 1인칭대명사 (1)는 세 개의 서로 다른 형태를, 2인칭대명사 (2)는 두 개의 서로 다른 형태를 갖는 점에서 KSL 지시수화는 약정성이 있다고 봄.



지시수화(pointing signs)라는 증거

수어언어학10주차

(b) 숫자 수형과 병합된 대명사의 집합은 약정적이다.

ASL TWO-OF-US THREE-OF-US FOUR-OF-US
 KSL



(c) 앞선 선행연구에서 언급한 ASL 방향성에 대한 제스처 설명은 1인칭 행위, 특히 **지시물에 대한 지시가 필수적이지 않은 1인칭 복수행위**를 설명하지 못하고, 이러한 점이 ASL에는 독립적인 1인칭이 존재한다는 것은 지시는 약정성이 있으므로 제스처가 아닌 수화라는 증거.



ASL sign WE (Meier & Lillo-Martin 2010: 351)



KSL sign 1인칭 복수 <우리> WE (이영재 2010: 109)



지시수화(pointing signs)라는 증거

수어언어학10주차

❖ (2) 음운단위에서의 대명사의 구성분석(compositionality)

• 음성언어든 수화언어이든 의미 있는 언어적 표시(단어, 구, 문장 등)는 의미없는 음운단위(음소, 음절 등)로 구성되어 있듯이 인칭대명사도 음운단위로 구성되어 있다.

KSL 1인칭 대명사의 음운단위:

- (a) 수형 (index-finger handshape, 1형 ) or 9형 ()
- (b) 이동 (path movement)
- (c) 장소 (location, 가슴 또는 코)

➡ ASL과 같이 KSL 비1인칭 대명사(2인칭과 3인칭)의 수형이 이동을 하여 지시하는 장소는 셀 수 없을 정도로 많기 때문에 음운론 음운단위로 열거하지 못하고 이러한 장소는 일반 어휘에서 대조를 보여주는 음소적인 역할을 못하고 어떤 형태소적 기능을 한다고 보기 어렵기 때문에 언어적이라기 보다는 제스처라고 보는 것이 타당할 수 있음.

지시수화(pointing signs)라는 증거

(3) 문법적인 제약을 갖는 대명사의 분포(grammatically constrained distribution)

지시수화라는 문법적인 증거는 대명사가 갖는 통사적인 분포에서 찾을 수 있음.

- 주어 대명사 복사 (subject pronoun copy, Padden 1988): 문장 주어의 대명사 복사는 문장 끝 위치에서 일어난다. 이러한 문장 끝에는 완전한 명사가 오지 않고 대명사가 놓인다는 것은 ASL과 같이 KSL의 대명사도 실제 대명사와 같은 행동을 한다는 것이다.
- 일치동사의 지시수화는 문법적인 제약을 받는다. 즉, 일치동사에 주어 또는 목적어가 일치하는 일치동사에만 해당되고 일치를 보여주지 않는 plain과 spatial 동사에는 적용이 되지 않는다.

(4) 대명사에 표시하는 문법적 구분(grammatical distinction marked by pronouns)

- 수(number)표시:
2중수, 3중수, 4중수 (left-right short movement) or 복수(arc movement)

- 성(gender)표시



KSL 인칭대명사의 수형 변이형

4.3 KSL 인칭대명사의 수형 변이형

- ◆ From gesture to sign 접근
- ◆ 음성 동반 제스처에서 다양한 수형변이형으로의 문법화 과정
(The grammaticalization of co-speech hand gestures)

- 1형  , 9형  ➡ 1인칭, 2인칭 대명사(단수, 복수)의 수형



- 5형  과 여성형  수지 제스처 ➡ 성구별 3인칭 대명사 수형으로 문법화됨

Pfau & Steinbach (2011); Janzen (2012); Pfau et al. (2012); Wilcox and Xavier (2013)

수어언어학10주차

KSL 인칭대명사의 수형 변이형

4.4 KSL 1인칭대명사의 위치 변이형

- 타수화언어의 1인칭대명사의 위치변이형: 1인칭 불규칙 목적어 일치동사
 Meier(2002:120): ASL and Danish Sign Language(DSL)



a. CONVINCE (citation form)



b. CONVINCE-1 'convince me'

From Meier (2011: 115)

77


수어언어학10주차

KSL 1인칭 불규칙 목적어 일치동사

- ❖ KSL 1인칭대명사의 위치 변이형
 - 장세은(2011: 96-97) KSL도 1인칭 불규칙 목적어 일치동사를 가짐.
 - 1인칭 불규칙 목적어 일치동사
 (KSL agreement verbs with irregular first-person object forms)

a. 가슴(chest) > 얼굴(face)
 <말하다> 'say' <불러서 가다> 'be summoned'
 <미워하다> 'hate' <묻다> 'ask'

b. 얼굴(face) > 가슴(chest)
 <가르치다> 'teach' <나무라다> 'scold'
 <노려보다> 'scowl at' <쳐다보다>, 'stare at'
 <겨냥하다> 'aim at' <쏘다> 'shoot'



ASL, DSL과 같이 KSL에서도 1인칭 불규칙 목적어 일치동사의 존재는 1인칭 대명사의 문법적인 범주에 대한 증거가 됨.


78

5. 결론, 문제점과 제언


- (1) 음성 동반 제스처(co-speech gesture)의 문법화에 대한 체계적인 연구가 요구됨.
- (2) KSL의 인칭대명사의 기능과 형식에 대한 더 심도 있는 연구가 필요함.
- (3) KSL의 인칭대명사가 two-person 또는 three-person 체계인지에 대한 언어학적 연구가 필요함.
- (4) 농아아동의 KSL 대명사 습득에 대한 연구가 필요함.
- (5) 웹기반 KSL의 코퍼스기반 사전구축을 통한 건청인 언어학연구자들에게 자연수화언어 자료제공과 자료중심의 언어학적 연구가 필요함.

References

1. 국립국어원 한국표준수화규범제정 추진위원회 (2007). KSL Dictionary, 한국수화사전(개정판). 도서출판 사단법인 한국농아인협회.
2. 이영희 (2010). 한국수화언어와 인칭대명사체계 연구. 전북대학교 석사학위논문.
3. 장세원 (2009). 한국수화언어와 동사유형과 수분류사. 언어과학 13(3), 101-135.
4. **장세원 (2011)**. Personal Pronouns and Reflexives in Korean Sign Language. *Journal of Language Sciences* 18(3), 87-111.
5. 최상범 (2012). 한국수화언어의 수할소 문법. 특수교육저널: 이론과 실천 13(1), 233-256.
6. Battison, R. (1978). Phonological deletion in American Sign Language. *Sign language Studies* 5, 1-19.
7. Bhat, D.N.S. (2004). *Pronouns*. Oxford: Oxford University Press.
8. Bybee, Joan, Revere Perkins, and William Pagluga. (1994). *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago, IL: University of Chicago Press.
9. Duncan, Susan (2005). Gesture in signing: a case study from Taiwan Sign Language. *Language and Linguistics*, 6(2), 279-318.
10. Emmoney, Karen (1999). Do signers gesture? In Lynn H. Messing and Ruth Campbell (Eds.), *Gesture, Speech, and Sign* (pp. 132-159). Oxford: Oxford University Press.
11. Goldin-Meadow, Susan (1999). The role of gesture in communication and thinking. *Trends in Cognitive Sciences*, 3(11), 419-429.
12. Janzen, T. (2012). Lexicalization and grammaticalization. In R. Pfau, M. Stenbach, and B. Wolf (Eds.) *Sign Language: An International Handbook*. 816-841. Berlin: Mouton De Gruyter
13. Janzen T. & B. Shaffer (2002). Gesture as the substrate in the process of ASL grammaticalization. In: MEIER, R.; QUINTO, D.; CORMIER, K. (Ed.). *Modality and structure in signed and spoken languages*. Cambridge: Cambridge University Press, 199-223.
14. Keller, Jörg. (1999). Using space in German Sign Language. *Sign Language & Linguistics* 2, 249-57.
15. Kelly, Spencer D., Sarah M. Manning, and Sabrina Rodak. (2008). Gesture gives a hand to language and learning: Perspectives from cognitive neuroscience. *Developmental Psychology and Education. Language and Linguistics Compass* 2 pp. 1-20.
16. Liddell, Scott K. 2003a. *Grammar, gesture, and meaning in American Sign Language*. Cambridge: Cambridge University Press.



수어언어학10주차



Any Questions?

81

‘쉬운-한국어(Plain-Korean)’를 통한 사회적 소통의 인문사회과학적 연구

이상수

(부경대학교)

1. 들어가기

본 연구는 ‘쉬운-한국어’¹⁾를 통하여 한국사회의 소통문제를 파악하고 언어 민주주의 실현을 목표로 한 인문사회과학적 연구이다.

한국사회는 이미 다문화 사회²⁾가 되면서 급변하는 언어현실에 직면해 있다. 외국인들에게는 각각의 다언어로 정보를 제공하는 것이 이상적이지만 번역할 언어가 증가할수록 행정 비용을 늘려야 하고 증가하는 다언어를 완전하게 대응하는 것은 불가능하다. 한국사회는 다수의 언어 약자³⁾들을 위한 언어적 배려가 부족하고 평등한 사회적 기회를 보장받지 못하여 사회적 분열이 야기되고 있다.

다양화된 현대 사회에서는 복잡한 어휘를 피하고 명확하고 간결한 언어표현이 요구되고 있다. 이러한 발상은 1960년대 미국에서 법률 및 정부문서를 쉬운 언어로 작성할 것을 요구하는 소비자 단체의 쉬운 영어 운동(plain English movement)에서 시작되었다. 이 운동의 목적은 쉬운 언어 사용을 통해 정부와 국민 간 의사소통을 활성화하고 관료주의적인 정부문서의 비효율을 줄이기 위한 것이다(Joanne Locke 2003:p.5). 1970년대 유럽(영국, 스웨덴 등)에서도 공문을 간결하고 쉽게 작성하는 공문서 개혁 운동을 추진해 왔다. 일본에서는 1995년 한신·아와지 대지진⁴⁾을 계기로 재난 시 필요한 정보를 입수할 수 없었던 외국인을 위해 사회언어학자와 NHK아나운서들이 쉬운 일본어(Yasashii Nihongo)에 관해 연구하기 시작했다. 이후, 평상시 정착외국인의 정보제공을 위해 경어나 한자표현을 없애고 간결한 어휘로 정리하는 등 일정한 원칙을 기반으로 한 쉬운 일본어가 제시되었다(庵功雄 2016:p.36).

한국사회는 한국어와 의사소통에 대한 각성 없이 언어로 인한 갈등, 오해, 분열 등으로 심각한 사회문제가 발생되고 있다. 정치적 올바름(*political correctness*), 성 중립(*gender*

1) 쉬운-한국어’란 쉬운 영어 운동(Clive, M. & Russo, F. (1981)), 쉬운 일본어(庵功雄(2016))와 같이 본 연구가 처음으로 제시하는 용어이다. 이하에서는 ‘쉬운-한국어’를 쉬운-한국어로 사용한다.

2) e-나라지표에 의하면 2021년 기준으로 체류외국인은 1,956,781명이며 장기체류자는 1,569,836명, 단기체류자는 386,945명이다. 단, 본 연구에서는 불법체류자는 제외한다.
(<http://www.index.go.kr/potal/main> 검색일:2022.12.13)

3) 본 연구의 언어약자는 언어소외 계층으로서 한국어 모어의 인식과 사용에 있어 일반적인 능력이 부족한 사람을 말하며, 언어약자에는 다문화, 장애인, 외국인, 새터민, 중국동포 등이 속한다.

4) 1995년 1월 17일 효고현의 아와지 섬 북쪽이 진원으로 발생한 규모 6.9로 6천명 이상의 사상자와 4만 명 이상의 부상자를 낸 대지진이다.

neutral)과 같은 존중과 소통의 언어생활은 시대적 요청이다.

본 발표에서는 언어사용에 관한 국내외 앞선 연구를 개관하고, 한국사회의 언어풍경을 진단하여 인문학적 성찰을 통해 쉬운-한국어 개념을 세우고자 한다.

2. 선행연구

쉬운 언어 사용에 관한 국내외 연구동향은 다음과 같다.

국내에서는 한자폐지(이석린 1949, 허웅 1969), 국어순화⁵⁾(김민수 1988, 남기심 2003, 민현식 2003, 김선철 2009, 김석득 2013, 조태린 2019)등 한자의 비근대적, 비능률적 측면에서 폐지론 주장, 국어를 다듬는 일, 비속어를 고운 말로, 맞춤법에 맞게 바르게 쓰는 법을 제시하였다. 고길섭(2000), 조태린(2019)등은 국어순화의 핵심 대상이 외래어가 아니라 어려운 말이 되어야 함을 강조하였다. 남영신(2009), 조태린(2010) 등은 공공문서의 비효율적 문제 지적, 공공언어 문제별로 사회적 합의를 통한 정부의 개입방식을 제안하였다.

국외에서는 1960년대 미국에서 법률 및 정부문서를 쉬운 언어로 작성할 것을 요구하는 소비자 단체의 쉬운 영어 운동(Plain English Movement)을 시작으로 공공기관을 중심으로 쉬운 언어쓰기 정책이 추진되어 왔다. 쉬운 언어를 통해 정부와 시민 간 의사소통을 활성화하고 관료주의적인 정부 문서의 비효율성을 줄이기 위한 것이다(Clive, M. & Russo, F. 1981:p.209, Locke, J. 2003:p.5).

이외에도 법률 문서를 쉽게 작성하는 원칙을 제시(Kimble 1994-1995), 법률 관계자들을 대상으로 법률 문서의 쉬운 영어 버전에 대한 선호도를 증명하여 쉬운 영어 사용의 필요성을 제시(Dubose 1991)하였다.

1970년대부터 유럽(영국, 스웨덴 등)에서도 공문을 간결하고 쉽게 작성하는 공문서 개혁 운동을 추진해 왔다. 독일에서는 2006년부터 언어 약자를 위하여 쉬운 언어(Leichte Sprache)를 국가 정책으로 도입하고, 교육적 관점에서 쉬운 독일어를 분석하여 이를 활용한 소통 방안을 제안하였다(Blum 2011, Holtz 2014, Alkhalil 2015).

일본에서는 1995년 한신·아와지 대지진⁶⁾을 계기로 재난 시 필요한 정보를 입수할 수 없었던 외국인을 위해 사회언어학자와 NHK아나운서들이 쉬운 일본어(*Yasashii Nihongo*)의 필요성을 논하기 시작했다. 이후, 평상시 정착 외국인의 정보제공을 위해 경어나 한자표현을 없애는 등의 간결한 문법, 어휘 등을 통한 쉬운 일본어가 제시(庵功雄 2016)되면서 사회구성원 간 소통의 메커니즘으로 부상하였다.

5) 국어순화의 용어에 대해서 민현식(2003)은 ‘國語純化’는 이상 개념, ‘國語醇化’는 방법 개념으로 구별하고, 김선철(2009)는 외래적인 요소의 축출로만 그 개념을 한정하여 ‘우리말 다듬기’ 또는 ‘국어 정화’라는 용어를 사용하였다.

6) 1995년 1월 17일 효고현의 아와지 섬 북쪽이 진원으로 발생한 규모 6.9로 6천명 이상의 사상자와 4만 명 이상의 부상자를 낸 대지진이다.

3. 쉬운-한국어는 왜 필요한가?

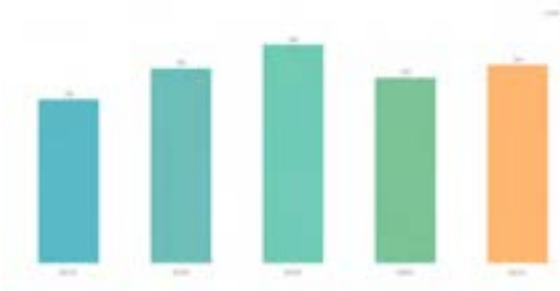
3.1. 다문화 사회

한국사회는 저출산 고령화 및 생산연령인구의 감소와 정착외국인의 증가로 다문화 사회⁷⁾가 되었다. 현재 체류 외국인의 현황은 다음과 같다.



<그림 1>⁸⁾

<표 1>과 같이 연도별 인구대비 체류외국인 현황('16~'21년) : 전체 인구 대비 체류외국인 비율은 2017년 4.21%에서 2019년 4.87%로 매년 증가하다 코로나19의 영향으로 2021년에는 3.79%로 감소하였다.



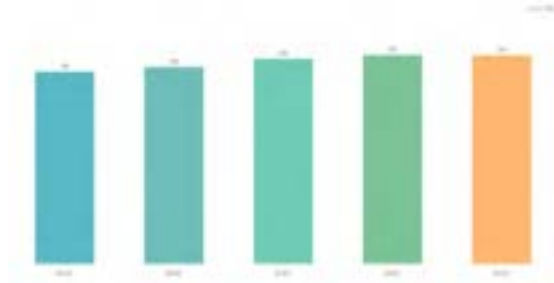
<그림 2> 연도별 유학생 현황⁹⁾

<그림 2>와 같이 연도별 취업자격 체류외국인 현황('16~'21년) : 2021년 말 기준 취업자격 체류외국인은 406,669명으로 전년 대비 10.1% 감소하였다.

7) e-나라지표에 의하면 2021년 기준으로 체류외국인은 1,956,781명이며 장기체류자는 1,569,836명, 단기체류자는 386,945명이다. 단, 본 연구에서는 불법체류자는 제외한다. (<http://www.index.go.kr/potal/main> 검색일:2022.12.13.) 단, 다문화사회는 정치, 사회, 경제 등 여러 문제와 관련되지만, 본 연구에서는 ‘언어’에 초점을 두어 고찰한다.

8) 법무부 (<https://www.moj.go.kr/moj/index.do> 검색일:2022.11.30)

9) 법무부 (<https://www.moj.go.kr/moj/index.do> 검색일:2022.11.30)



<그림 3> 연도별 결혼이민자 현황¹⁰⁾

<그림 3>과 같이 연도별 결혼이민자 현황('16~'21년) : 2021년 말 기준 결혼이민자는 168,611명으로 전년 대비 0.01% 증가하였다.

이처럼 한국사회는 구성원으로서 외국인을 수용해야 하며 외국인들과의 공존 방안을 모색해야 할 시점에 이르렀다.

3.2. 한국어 사용의 문제

한자어, 비문법적 표현, 차별어, 폭력어 등으로 인한 언어사용의 문제는 한국사회의 심각한 불통과 갈등을 야기하고 있다. 최근 몇 년 간 일상에 수시로 제공되는 “안전안내문자”나 “긴급재난문자”는 한자어나 비문법적 표현으로 인하여 불통을 낳는 경우가 적지 않다.

- (1) · AI관련, 5월23일 이후 전통시장에서 **가급류**를 구입하신 분은~
- 여러 가지 대응노력을 하고 있습니다만, **적수가** 나오고 있어~
- 코로나19환자 3,844명. 60세 이상 및 만 12세 이상 **기저질환자** 확진시~
- 태풍 **내습시 하천, 해안가 등 위험지역 접근금지**
- **별도로** 신청하지 않아도 신청한 것으로 **간주**하여 최초 긴급고용안정 지원금 신청서에 **기재된** 계좌로 지급됩니다.
- 계좌이체오류 발생시 **지급이 상당기간** 지연될 수 있어 **가급적** 계좌 확인 필요
- **금일** 한파주의보 **발령** 및 눈 **예보**가 있으니 안전에 유의
- **승강장 연단간격**이 넓은 곳이니 **승하차** 시 주의
- 수도권량기 **동파동결** 예방조치, **전열기** 사용시 화재에 주의
- 교통 **정체가 예상**되니 ~ 보행시 **낙상사고**에 주의

한국사회는 차별과 불통의 언어로 사회적 악을 양성하고 이로 인한 사회경제적 불평등이 심화되어 심각한 계층 갈등을 양산해 왔다.

(2)

전래/일차별어	세대간 차별어	계층간 차별어	지역간 차별어
저출산	노년노년	김삼면	삼골
미숙어	한구/할양구	황금쟁이	공노
미혼모	노산재	장사꾼	콘스럼다
지매공연	나다래	남사꾼	지병대
효자성동	애채들	노기다	백강남공
장사/방오	장쟁이	간오래	택수도공
방머리 장길	젊은것들	청소노동자	
도연님	유즘것들	똥만	
머기회		기름부	
김서방/안서방		신용불량자	
채내직		부하직원	
엥스태머션		허급직원	

10) 법무부 (<https://www.moj.go.kr/moj/index.do> 검색일:2022.11.30)

- (3) · 민원인과 콜센터 직원간의 언어폭력: 지난 2월 콜센터 직원에게 민원인은 “빨리 좀 해” “이 XX가”와 같은 폭언을 퍼부었다. 직원의 소관이 아닌 업무를 요구하며 30분 넘게 폭언을 이어 갔다.¹¹⁾
 - ‘oo충’ 온라인이나 우리 실생활에서 공공연하게 쓰이고 있는 단어이다. 별레라는 의미의 충을 붙여서 ‘급식충, 진지충, 맘충, 한남충’ 등 특정한 사람을 비하하는 어.
- (4) · 술먹고 춤추는 이태원, 가면쓰고 귀신놀이, 외국서 온 헬러원 금지하라, 급조유족 거부, 자식 못 막은 부모탓, 이태원 놀러 간 건 그들의 선택, 사람 많은 곳, 서로 민 그들 책임, 국위선양 했냐, 국가책임 물을 생각말라¹²⁾ 등

3.3. 사회적 불통 문제

현재 한국사회는 언어적 불통으로 인하여 세대, 계층 등 심각한 갈등을 양산하고 있다. 최근 SNS상에 사용되는 언어로서 새로 만들어진 단어 및 용어 가운데 표준어로 등재되지 않은 말들은 흔히 접할 수 있다.

- (5) · 점메추(점심메뉴추천)
 - 주불(주소 불러)
 - 켈박(캘린더에박제)
 - 완내스(완전 내 스타일이라는 뜻이다. 사람, 옷, 물건 등에 사용)
 - 많관부(많은 관심 부탁드립니다)
 - 스불재(스스로 불러온 재앙의 준말로 자신이 원해서 한 일이 능력 부족이나 환경의 변화로 예기치 않게 일을 그르쳤을 때 사용)
 - 너 뭐 돼?(너 혹시 뭐라도 돼? 라는 준말로 너가 얼마나 대단하길래 그렇게 하나 라는 의미)
 - 다꾸(다이어리 꾸미기의 준말)
 - 켈박(캘린더(달력)에 박제, 중요한 것을 잊지 않겠다는 뜻)
 - 좋댓구알(좋아요, 댓글, 구독, 알림설정. 유튜브 증가로 인한 신조어)
 - 믿거(믿고 거르는)
 - 취존(취향 존중)
 - 솔까말(솔직히 까고 말해서)
 - 자냥괴(자본주의가 낳은 괴물)
 - 자만추(자연스러운 만남 추구)
- (6) · 킹리적갓심(합리적의심)
 - 갓생(성실하고부지런한 삶)
 - 조삼모사(사자성어가 아닌 학생들이 시험에서 찍을 때 쓰는 말. 조금 모르면 3번, 완전 모르면 4번)
 - 식집사(식물을 기르는 사람. 코로나 이후 집 생활이 늘며 생김)

11) <https://www.khan.co.kr/national/national-general/article/202212141010001> 콜센터직원언어폭력

12) 국민일보·카리스트 이원제교수팀의 빅데이터 분석결과, 댓글문장 분석을 통해 뽑힌 혐오 사례 (<https://m.kmib.co.kr/view.asp?arcid=0017753310> 검색일: 2022.12.22.)

- (7) · 개이득(개+이득)
 - 주린이(주식+어린이)
 - 복돌이(복사+돌이: 불법 복제품을 사용하는 사람들을 일컫는 말)
 - 정돌이(정품+돌이: 정품을 사용하는 사람들을 일컫는 말)
 - 친목질(친목+질)
- (8) · 옷프다(옷기다+슬프다)
 - 덕통사고(덕후+교통사고: 아이돌 입덕은 사고처럼 우연히 일어나는 것)
- (9) · 사회적 거리두기(감염 차단을 위해 사람들 간 거리를 유지하는 캠페인)
 - 언택트(서로 만나지 않고 비대면으로 소통하는 방식)
 - 확진자(코로나 사태 속에서 활동량이 줄면서 단기간에 살이 확 썩 사람을 일컫는 말)
 - 텍스크/코스크(마스크를 내려 턱 밑에 걸쳐두거나 코는 내놓고 입만 가리는 경우를 일컫는 말)
 - 코로나 블루(코로나에 우울을 뜻하는 ‘blue’가 합쳐진 말로 일상 변화로 인해 우울, 불안 등을 느끼는 증상)
 - 상상코로나(‘상상 임신’에서 파생된 말로 기침, 미열 등의 증상이 보이면 바로 코로나로 의심하는 것을 의미)

4. 쉬운-한국어란

4.1. 정확한 정보제공을 위한 쉬운-한국어

현재 한국사회에서는 외국인은 물론 언어약자에 대한 배려가 부족한 언어현실에서는 이들이 평등한 사회적 기회를 보장받기 어렵다. “긴급재난문자”와 “안전안내문자”와 같이 생존과 연관된 정보와 더불어 상시 정보제공도 중요하다. 정부와 국민 간 의사소통의 상시 정보 제공으로서 공문서(공공언어(각종 공문서 포함)가 중요한 의미를 가지지만, 여전히 이해하기 어려운 표현으로 제공되고 있다.

(10)

참고	편본을 통해 본 「군중 압착 사고」 발생 시 행동요령
<p>1. 상황 인지 방법</p> <p>가. (밀집도 상승 신호) 군중들이 전전히 움직이기 시작할 때</p> <p>나. (밀집도 위험 수준) 움직이면 인파의 흐름이 갑자기 멎을 때</p> <p>다. (상황특제 불능 판단) 주변 사람들이 불관함이나 고통을 호소하는 소리가 들릴 때</p> <p>2. 위험한 인파 속에 헤엄칠 때 대처 방법</p> <p>가. (상황 파악) 인파의 이동이 멎을 경우 주의, 스스로 의지대로 움직일 수 있는지 확인 후 가능하면 빨리 탈출</p> <p>나. (신호 확보) 고관 비명 자백, 최대한 질작함을 유지하여 머리를 들어 신호 확보</p> <p>다. (공간 확보)</p> <ul style="list-style-type: none"> - 팔로 땅과 단둘이 손 뺄 공간 확보, 주로 쓰인 손으로 반대쪽 손목 위를 잡거나 두 팔로 땅잡기 - 두 발로 땅을 굳게 딛고 서기, 핵심 선수처럼 다리를 앞뒤로 벌리고 무릎 살짝 굽히기 라. (여동) 인파의 흐름을 거스르지 말고, 외도록 흐름에 따르면서 천천히 대각선으로 가장자리를 향해 움직이며 탈출로 찾기 <p>3. 기타 유의 사항</p> <p>가. (본인이 넘어졌을 때) 곧바로 일어나기 불가능하다면, 왼쪽 옆으로 쏠려서 머리를 보호하기, 등, 배를 대고 눕거나 옆으로만 매우 위험</p> <p>나. (물건을 떨어뜨렸을 때) 주우려고 허리를 숙이지 말고 바로 포기하기</p> <p>다. 보호자(부모, 교사 등)의 경우</p> <ul style="list-style-type: none"> - (어린이 넘어졌을 때) 옆 사람이 두 발로 서 있어야 내 생존 확률도 올라가므로 최선을 다해 일으켜 세워주기 - (어려 동반 시) 아이를 어깨 위로 올리거나 언어 불려 아이가 어른의 허리를 감쌀 것, 아이의 팔 끌어당기지 않기 	

□ 연파 밀집 장소 기기 전 유의 사항

※ (출처) 연합뉴스(2022.11.17), 국민일보(2022.11.27), (자료) 워싱턴포스트(WP)



(10)의 일상에 제공되는 공공언어에는 밀집도 상승, 상황통제 불능, 자제, 확보, 귀가조치, 동선 조정, 수시 환기 등 어려운 한자가 여전히 제공됨으로써 외국인, 언어약자는 물론 모어화자도 이해하기 어렵다.

4.2. 원활한 커뮤니케이션을 위한 쉬운-한국어

실제 모어화자와 외국인의 접촉장면에서 외국인만이 아니라 모어화자도 오해하는 경우가 있다. 다음은 외국인 학습자와 한국인 화자 사이에 나타나는 질의응답 결과를 제시한 예13) 중에 하나이다.

(11)가. 한국에 어떻게 오셨어요? (외:78.9%, 한:32.7%)

나. 비행기 타고 왔어요. (외:26.9%, 한:70.9%)

나'. 한국어를 공부하러 왔어요. (외:9.6%, 한:54.5%)

나". 한국어를 공부하고 싶어서 왔어요. (외:9.6%, 한:70.9%)

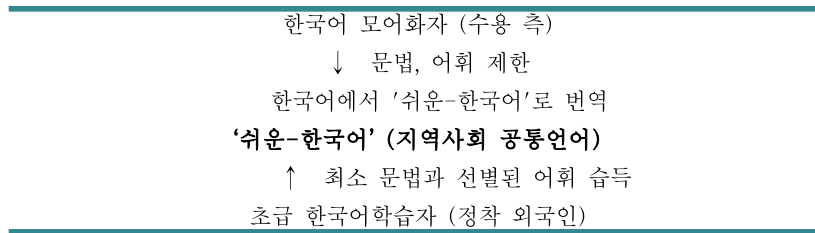
우선, (11가)은 일차적으로 이동 방법에 대한 것은 직설적, 일차적 응답으로 (11나)이다. 그러나 이동이나 경로와 관련되지 않은 (11나'), (11나")와 같은 목적이나 이유 표현 등도 (11가)에 대한 응답이 될 수 있다. 이러한 장면은 경로를 묻는 것이 아니라 한국에 온 목적이나 이유를 묻는 것으로도 사용할 수 있기 때문이다. 어떠한 응답이든 질의 맥락에 따라 응답은 달라 질 수 있으나 모어화자가 질의에 적절한 응답을 하는 것은 아니라는 것을 알 수 있다.

13) 서진숙(2015) 「외국인 학습자와 한국인 화자의 질의응답 양상 분석 연구」, 『국어국문학』제171호, p.209

4.3. 지역사회 공통언어로서의 쉬운-한국어

쉬운-한국어는 실질적으로 모어화자에게 중요한 의미가 있다. 한국사회가 추구해야 할 배려와 존중의 언어를 구축하는 데는 모어화자가 무의식적으로 사용하는 한국어를 대화 상대의 눈높이에 맞추어 조정하는 의식화가 필요하다. 외국인이 사회구성원으로서 최소한의 한국어를 습득해야 한다면, 한국인도 외국인의 한국어 표현을 이해하기 위해 조정하는 훈련도 필요하다.

<표 1> 지역사회 공통언어로서 쉬운-한국어



<표 1>은 지역사회 공통언어는 모어화자가 일상적으로 사용하는 한국어가 아니라 공통언어로 성립되는 것이 쉬운-한국어이며, 평등 사회를 실현하는 데 필수적이다. 본 연구의 핵심은 누구나 소통할 수 있는 쉬운-한국어를 정립하여 소통을 원활하게 하는 것이다. 또한 외국인에 대한 초기 한국어교육으로서 공적으로 보장되어야 하며 구체적이고 체계적인 교육 내용이 제시되어야 한다.

5. 나오기

본 연구는 한국사회의 소통 문제를 해결하기 위해 지난 연구와 아울러 언어, 문화, 사회간의 유기적 상관성에 천착한 심도 깊은 연구가 절실히 필요하다는 취지 아래, 쉬운-한국어를 개발하여 사회 전반에 적용하기 위한 시도이다.

한국사회가 공존, 평등, 소통의 언어민주화가 정착되어야만 건강한 민주주의를 완성할 수 있다. 그러기 위해 언어사용자, 문서생산자, 발화자 등은 바르고 편하고 쉬운 언어를 쓰고 언중들은 본의를 쉽고 빠르게 전달할 수 있는 기술적인 언어기법을 익혀야 한다. 최소한의 문법규칙과 다양한 어휘를 통해 커뮤니케이션능력을 향상시키는 한국어 사용에 관한 발상의 전환이 필요하다. 쉬운-한국어를 통해 한국어 모어화자가 쉽게 한국어를 표현하고 이를 외국인이 자연스럽게 받아들였을 때 한국은 성숙한 다문화 공존사회가 될 수 있다.

참고문헌

- 남기심. 2003. 국어 순화와 관련하여. 「국어 순화 실천 방안 마련을 위한 학술대회」, 국립국어연구원, 9-11.
- 민현식. 2003. 국어 순화의 국어학적 연구. 「국어 순화 정책 연구 보고서」, 국립국어연구원, 33-212.
- 이석린. 1949. 한자를 폐지하자. 「한글」 107, 한글학회, 416-418.
- 한국독일어교육학회. 2010. 언어학습, 교수, 평가를 위한 유럽공동참조기준. 「한국문화사」, 1-122.
- 허웅. 1969. 한자는 폐지되어야 한다. 「한글」 143, 한글학회, 23-49.
- 고길섭. 2000. 국어 순화 정책, 무엇이 문제인가? -국어 순화 정책의 역사적, 문화적 비판-. 「국어 순화 실천 방안 마련을 위한 학술대회 자료집」, 국립국어연구원, 35-44.
- 민현식. 2015. 언어 불통의 진단과 소통의 방법. 「산청어문」 43, 서울대학교국어교육과, 113-162.
- 유재원. 2005. 국어 순화, 왜 그리고 어떻게 해야 하나?. 「새국어생활」 15.1, 국립국어원, 7-17.
- 조태린. 2010. 공공언어 문제에 대한 정책적 개입 방식. 「한말연구」 27, 한말연구학회, 379-405
- 庵功雄. 2013. やさしい日本語研究の現状と今後の課題. 「一橋日本語教育研究会」, 一橋日本語教育研究会, 1-12.
- 庵功雄. 2016. やさしい日本語—多文化共生社会へ—. 「岩波書店」, 2-87.
- 庵功雄・イヨンスク・森篤嗣. 2017. やさしい日本語は何を目指すか—多文化共生社会を実現するために—. 「ココ出版」, 3-48.
- 岩田一成. 2010. 言語サービスにおける英語志向—『生活のための日本語：全国調査』結果と広島の事例から—. 「社会言語科学」 13.1, 社会言語科学会, 81-94. 72-86.
- 森篤嗣. 2013. 語彙から見た「やさしい日本語」, 「やさしい日本語」は何を目指すか. 「ココ出版」, 99-115.
- 山田 泉. 2002. 地域社会と日本語教育, ことばと文化を結ぶ日本語教育, 凡人社, 118-135.
- Council of Europe. 2001. Common European Framework of Reference for Languages: Learning, teaching, assessment. Cambridge: Cambridge University Press.
- Council of Europe. 2002. Common European Framework of Reference for Languages: Learning, Teaching, Assessment Case Studies. Strasburg, Council of Europe.
- Kevin Dubose. 1991. *The Court Has Ruled, The Second Draft*, Legal Writing Institute, 36-48.
- Kimble, Joseph. 1994-1995. "Answering the Critics of Plain Language", *The Scribes Journal of Legal Writing* 51, Vol.5, 53-60.
- Michael Clive and Francine Russo. 1981. "The Plain English Movement in America: A View from the Front", *Information Design Journal*, Vol.2, No.3-4, 208-214.

한국어 교육을 위한 한국어 피동 표현 연구

-일본어 피동 표현과의 대조 분석을 중심으로-

김미루

(외국어로서의 한국어교육전공)

1. 서론
 - 가. 연구 목적
 - 나. 선행 연구
2. 한국어의 피동 표현과 일본어의 피동 표현
 - 2.1. 한국어 피동 표현과 일본어 수동 표현의 특징
 - 2.1.1. 한국어 피동 표현
 - 2.1.2. 일본어 수동 표현
 - 2.2. 한국어 피동 표현과 일본어 수동 표현의 대조 및 학습 전략
 - 2.2.1. 한국어 피동 표현과 일본어 자동사 대조 분석
 - 2.2.2. 일본어의 수동 표현이 한국어에서는 피동 표현으로 사용되지 않는 경우
 - 2.2.3. 일본어 피해 수동 표현의 한국어적 표현으로의 전이
 - 2.2.4. 일본어 사역 수동 표현의 한국어적 표현으로의 전이
3. 결론

I. 서론

한국어와 일본어는 그 유사성에서 기인하여 초기 학습에 용이한 점이 있다는 것은 부정할 수 없는 사실이나, 언어 학습이 심화될수록 학습에 어려움을 겪는 요소들이 나타나게 된다. 그 대표적인 예 중 하나가 바로 피동 표현이다. 특히나 한국어의 피동 표현의 경우 일본어 뿐 아니라 영어나 기타 언어의 피동 표현에 비해서 그 생산성이나 의미에 있어서 차이점이 있으며, 그에 따르는 의문점이 여럿 있다. 본고에서는 한국어 피동 표현과 일본어 피동 표현을 대조 분석하여, 한국어 피동 표현이 가지는 의미와 피동성에 대해서 살펴보고, 한국어 피동 표현 학습에 대해서도 논의해 보고자 한다.

1.1. 연구 목적 및 방법

한국어에서는 피동법이라는 개념은 일본어의 수동(이는 한국어에서 대응된 개념이며, 정확히는 受身ukemi, 수신)에 해당하는 것인데, 한국어의 피동 표현의 의미 범주와 일본어 피동 표현의 의미역에 차이가 있는 것을 알 수 있고, 한국어와 일본어의 피동 표현에 대한 개념적, 인식적 차이가 있다는 것을 알 수 있다.

- (1) 가. 私はバスである男に足を踏まれた
 watashiwa basude aru otokoni ashio humareta.
 나. 나는 버스에서 어떤 남자에게 발을 밟혔다.
- (2) 가. 私は学生たちに見られた
 watashiwa gakuseitachini mirareta
 나. 나는 학생들에게 보였다?/봐졌다?/보게 되었다?.
- (3) 가. 私は去年母に死なれた
 watashiwa kyoneN hahani shinareta
 나. 나는 작년 어머니에게 죽었다?/죽어졌다?/죽게 되었다?.
- (4) 가. 居眠りしすぎて `私は先生に立たされたまま `授業を受けた
 inemuri shisugite, watashiwa senseini tatasaretamama,jyugyouuketa
 나. 너무 졸아서, 나는 선생님에게 서게 하는 것을 당한? 채로, 수업을 들었다.

(1)의 경우는 일본어의 피동 표현이 한국어에서도 피동 표현으로 동일하게 나타난 경우이다. 그러나 (2)의 경우는 일본어에서는 피동 표현으로 사용하나 한국어에서는 피동 표현으로 사용하지 않는 경우, (3)의 경우는 일본어에서 '간접 피동 표현'으로 대표되는, 한국어에서는 사용하지 않는 표현방식인 경우로, (3 가)의 의미는 여격으로 나타난 어머니가 죽은 동작으로 인해 자신이 어떠한 방식으로든 피해(감정적 우울함, 경제적 어려움 - 일반적으로 다른 문장에서 그 원인이 추가적으로 설명되는 경우이다)를 받았다고 생각하는 경우에 사용하는 표현으로, 한국어의 피동 표현의 성격과 큰 차이를 보인다. (4)의 경우는 일본어에서는 소위 사동 피동 표현이라 하는 것이다. 의미는 사동 하는 행위를 '당하다'라는 의미로, 화자=주어는 그 행위를 하기 싫지만, 여격으로 표현되는 행위주가 나에게 어떠한 행위를 시켰고, 나는 그것을 당했다(억지로 했다)는 의미가 된다. 따라서 (4)의 의미는, 내가 너무 졸았기 때문에 나는 선생님이 나를 (억지로) 자리에서 서 있게 하였고, 나는 그 행위를 당했다. 라는 해석이 된다.

- (5) 가. 部屋のドアが開いた
 heyano doaga aita
 나. 방 문이 열렸다.
- 다. 私は部屋のドアを開けた
 watashiwa heyano doao aketa
 라. 나는 방 문을 열었다.

(5 가)의 경우 일본어에서는 피동 표현을 사용하지 않으며, 다만 동사 자동사의 성격을 가진다. 그러나 한국어의 경우는 '-리'의 피동 접미사가 사용되어 피동 표현으로 나타난 것을 알 수 있다. (5 다)의 경우 (5 가)에서 사용된 '開く(열리다)'의 타동사 '開ける(열다)'를 사용한 문장이다. 이처럼 일본어에서 자동사로 사용되는 것이 한국어에서는 피동 표현으로 나타나는 경우들이 많고, 이러한 점에서 살펴보다라도 한국어의 피동 표현과 일본어의 피동 표현의 영역이 다소 어긋나 있는 것을 확인할 수 있다.

위 내용과 같이 한국어 피동 표현과 일본어 피동 표현은 의미 영역이 일치하는 부분도 있

으나, 본질적인 차이점이 존재하는 것으로 보인다. 또한 일본어에서만 사용하는 특수한 피동 표현 영역도 존재하기 때문에, 이를 한국어로 적절히 옮기는 것에 대한 고민 또한 필요할 것이다. 본고에서는 이 때문에 발생할 수 있는 피동 표현 학습상의 어려움에 대해 고찰하고, 이를 한국어 학습에 적용할 때 어떠한 전략을 취해야 하는지를 연구하는 것을 목적으로 한다.

이를 위해 먼저 한국어 피동 표현의 유형과 다루고자 하는 범주에 대해서 먼저 제시하고, 다음으로 일본어 피동 표현의 특징과 특수한 유형에 대해서 살펴보기로 한다. 이후 한국어 피동 표현과 일본어 피동 표현을 유형별로 대조하여 분석한 이후, 그 공통점과 차이점을 기반으로 피동 표현 학습의 방향을 제시하고, 일본어권 학습자들의 한국어 피동 표현 수용성에 대한 문제와 일본어의 특수한 피동 표현에 대응하여 한국어 표현으로 옮기는 방법에 대해서도 살펴보려고 한다.

1.2. 선행 연구

먼저 이정택 · 기타무라다다시 · 전나영(2005)과 이효숙(2008)에서는 한 · 일어 피동 범주에 대한 고찰을 바탕으로 한국어 피동 범주와 교수 · 학습법에 대하여 논하였는데, 일본어 피동문은 생산적인 문법 장치를 통해 이루어지나, 한국어 피동 표현은 여러 가지 유표적 의미와 제약을 지닌 다양한 어휘요소에 의존한다는 것을 명시하면서 한국어 피동문 생산에 따르는 제약 요소를 분명히 하였다. 이정택 · 기타무라다다시 · 전나영(2005)은 이러한 차이로 인해 일본어 화자의 일본식 피동문 양산에 대한 우려를 제기하였다.

허경행(2016)에서는 '-이, -히, -리, -기' 접미사에 의한 피동의 경우 의미에 따라 구분해서 학습해야 한다는 점을 강조하기도 하였다. 다만, 일본어 '-られる' 형태를 논함에 있어, 다만 형태만 같을 뿐인 존경어 사용과 가능형 사용까지 언급한 것은 취지에도 맞지 않으며, 다만 형태가 같다고 해서 피동 표현의 다른 의미로 다룬 것은 적절치 못한 것이 아닌가 하는 의문이 있다.

양경모(2017)에서는 주로 [V-られる]와 한국어 {이} 피동사 및 {V-어지다}를 중점적으로 비교 분석하였는데, {V-어지다}가 [V-られる]와 대응할 수 있는 경우와 아닌 경우를 들어 설명하면서, 이것이 통어 · 의미적 기능에 있어 등가의 것이 아님을 시사하였다. 다만 제한적으로 통어적 접점이 있는 경우에 제한적으로 대응할 수 있음을 밝혔다.

일본어 자동사와 한국어 피동 표현의 대조에 관한 선행 연구에는 앞서 언급한 허경행(2016)이 있는데, 이 논문에서는 일본어 타동사의 피동형이 한국어 피동 표현에는 없다는 것을 이야기하며, 일본어 자동사의 일부가 한국어의 피동 표현에 해당하는 것을 언급하였다. 다만 정리가 잘못된 것인지, 일본어의 자 · 타동사 구별이 반대로 되어있는 경우가 있어, 앞서 언급한 오류와 함께 아쉬움을 남기기도 하였다.

양경모(2017)에서는 [V-られる]동사가 한국어 {이} 피동사에 대응되지 않는 경우들을 언급하는 가운데 일본어 자동사와 한국어 피동 표현이 대응되는 경우에 대해서 언급했고, 한국어가 피동적으로 사태를 표현하기보다는 능동적으로 표현하는 경향이 높다는 점을 시사했다. 정해권(2016)에서는 한국어 피동 표지의 다의성에 대해 언급하며, 이것이 피동 > 자발 > 가능 > 사동의 순서로 습득이 되며 피동과 가능을 구별하여 단계별로 학습해야 한다 덧붙였다. 이는 피동 표현이 온전히 피동의 표현만을 가지는 것이 아님을 시사하기도 한다.

2. 한국어의 피동 표현과 일본어의 피동 표현

2.1. 한국어 피동 표현과 일본어 피동 표현의 특징

2.1.1. 한국어 피동 표현

한국어 피동 표현의 유형은 동사에 '-이, -히, -리, -기'와 같은 접미사에 의해 이루어지는 피동법이 있고, '되다, 받다, 당하다'와 같이 어휘가 결합한 표현이 있으며, '-아/어지다'와 같이 통사적 결합에 의한 피동 표현도 있다.

이를 각각 파생적 피동, 어휘적 피동, 통사적 피동이라고 했을 때, 파생적 피동의 경우에는 '깎이다, 묶이다, 막히다, 먹히다, 걸리다, 깔리다, 안기다, 뜯기다'와 같이 타동사에 접미사가 결합하여 생산되는데, 각각의 동사가 어떤 접미사와 결합하는지에 대한 명확한 근거나 기준이 없으며, 피동 접미사가 아예 결합하지 못하는 경우도 존재하기 때문에(하다, 자다, 가다, 끄다 등) 생산성의 제약이 다소 존재하며, L2 언어로서 한국어를 학습하는 경우, 이것이 접미사를 통한 생산이라기보다는 피동 접미사가 결합한 형태의 피동사를 개별적으로 암기해야 하며, 이러한 방법 외에는 특별히 효율적 학습 방법의 여지가 없다고 보는 것이 현실이다.

'-아/어지다'의 피동 표현을 어떻게 취급할 것인가에 대한 문제를 살펴보면, 고영근·구본관(2018)에서는 [형용사+-{아/어지다}]가 피동보다는 상태 변화에 가까운 형태라고 보았다.

- (6) 가. 안 보는 사이 예뻐졌다.
 나. 그새 키가 많이 커졌다.

이 문장을 각각 일본어 표현으로 표현하였을 때는 다음과 같은 문장이 된다.

- (7) 가. 見ないうちにきれいになった
 minai utsini kireini natta
 나. そのうち `背が高くなった
 sono utsi, sega takaku natta

일본어 표현으로 바꾸어 보아도, 피동 표현이 아니라, 한국어의 '되다'에 대응되는 표현인 '-くなる'의 형태가 된 것을 볼 수 있다. 위와 같은 논의에 따라, [형용사+-{아/어지다}]의 형태 대조에 대한 논의는 본고에서는 제외하기로 한다.

또한 '되다'와 같은 파생적 피동 표현을 어떻게 다룰 것인가에 대해서도 동일한 논의가 필요하다. 아래의 예문을 보도록 하자.

- (8) 가. 그가 먼저 집에 가게 되었다.
 나. 나 혼자 밥을 먹게 되었다.
 다. 색깔이 밝게 되었다.

(8 가)의 경우 자동사에 '-게 되다'가 결합한 경우 (7 나, 다)는 각각 타동사와 형용사에 결합한 경우인데, 이 경우도 마찬가지로 의미적으로도 상태의 변화를 나타내는 것으로 보이며, 이를 일본어로 옮기더라도 마찬가지다.

- (9) 가. 彼が先に家に帰るようになった
karega sakini ieni kaeruyouninatta
나. 私一人でご飯を食べるようになった
watashi hitoride gohaNo taberuyouninatta
다. 色が明るくなった
iroga akarukunatta

동사의 경우 '-ようになった'의 형태로 나타나고, 형용사의 경우 '-くなった'의 형태로, 피동 표현으로 나타나지 않고, 그 의미도 피동 표현의 그것과는 거리가 있다. 따라서, 본고에서는 '-게 되다'의 피동 표현 또한 논의에서 제외하기로 한다.

2.1.2. 일본어 피동 표현

일본어의 경우, 피동 표현이 능동 동사에 조동사 '-れる, られる'가 결합하여 생산된다는 특징이 있고, 이는 거의 모든 어휘에 제약 없이 결합하여 사용할 수 있으므로, 한국어 피동 표현에 비해 규칙성이 있고, 생산성이 있다고 볼 수 있다.

- (10) 가. 行く `脱ぐ` 読む `呼ぶ` 死ぬ `問う` 待つ `折る` 押す
iku, nugu, yomu, yobu, sinu, tou, matsu, oru, osu
가'. 行かれる `脱がれる` 読まれる `呼ばれる` 問われる `待たれる` 折られる `押される`
ikareru, nugareru, yomareru, yobareru, yowareru, matareru, orareru, osareru
나. 見る `降りる` 食べる `植える`
miru, oriru, taberu, ueru
나' 見られる `降りられる` 食べられる `植えられる`
mirareru, orirareru, taberareru, uerareru
다. する `来る`
suru, kuru
다'. される `来られる`
sareru, korareru

일본어의 경우 동사 어미의 활용 형태에 따라서 동사의 종류가 총 세 가지로 나뉘는데, (10 가)와 같은 형태를 5단 동사라고 하며, 5단 동사의 경우 어미의 모음이 /a/로 바뀌고, 조동사 '-れる'가 결합된다. (10 나)와 같은 동사들은 (상, 하)1단 동사라고 하고 이 동사들은 어간에 '-られる'가 결합되는 형태로 피동 표현을 만든다. (10 다)의 경우는 활용 방식이 불규칙적으로 고정된 경우이다.¹⁾ 그러나 동사의 활용 종류와는 관계없이, 조동사 '-れ

1) 근대 이전의 일본어 활용 형태가 굳어져서 활용된 것이므로, 엄밀히 말하면 활용 규칙이 존재하지

る, -られる'가 결합된 것을 볼 수 있으며, 동사의 성격이나 어휘 의미와 상관없이 거의 모든 동사에 결합되어 사용하고, 대부분은 실사용 문장에서 활용할 수 있으므로, 한국어의 피동 표현 생산과는 확연한 차이가 있다.

2.2. 한국어 피동 표현과 일본어 피동 표현의 대조

2.2.1 한국어 피동 표현과 일본어 자동사 대조 분석

한국어 피동 표현과 일본어 피동 표현에는 본질적으로 다른 측면이 존재한다. 허경행 (2016)의 논지에서 본다면, 한국어 피동 표현의 경우 능동사의 행동이 이루어진 상황이나 상태를 서술하는 데 주로 사용되고, 일본어의 피동 표현은 그 행위의 주체가 명확한 경우가 많고, 문장 주어는 그 행위를 받거나 당한다는 측면의 의식이 강하다고 한다. 다음 예문을 통해 살펴보자.

(11) 가. 部屋のドアが開いた

heyano doaga aita

방 문이 열렸다.

나. 私は部屋のドアを開けた

watashiwa heyano doao aketa

나는 방 문을 열었다.

다. 私は部屋のドアを誰かに開けられた

watashiwa heyano doao darekani akerareta

?나는 방문이 누군가에게 열렸다.

(11 가)의 문장을 보면, 일본어의 자동사 '開く'는 한국어 '열다'의 피동 표현 '열리다'가 되고, (11 나)의 문장을 보면 일본어의 자동사 '開ける'는 한국어의 능동 동사 '열다'가 된다. 그러나 (11 다)의 경우, 그 의미는 '나는 누군가에게 방문을 여는 행위를 당했다'라는 의미가 된다. 이것을 표로 정리하면 다음과 같다.

<표 1> 일본어 자동사와 한국어 피동 표현의 관계

일본어 타동사 開ける	일본어 자동사 開く	일본어 피동 표현 開けられる
한국어 타동사 열다	한국어 피동 표현 열리다	x

모든 어휘에서 그러한 것은 아니지만, 일본어에서 자·타동사의 대응이 있는 경우는 대부분 <표 1>에서 볼 수 있듯이, 일본어 타동사는 한국어의 타동사와 동일한 대응이 일어나지만, 일본어 자동사의 경우 한국어에서는 피동 표현에 대응하는 경우가 많고, 일본어의 피동 표현의 의미를 생각했을 때, 온전한 의미로 번역할 수 있는 표현이 없다. 사실 이러한 측면

만, 현대 일본어 관점에서 설명하려면 고정된 불규칙 형태라고 설명하는 것이 본고에서는 유용하다고 판단하였다.

이 일본어 피동 표현의 의미적인 부분을 여실히 보여주는 것이라 생각한다. 한국어의 피동 표현과 비교했을 때는 그것이 더욱 대비적으로 이해된다고 생각한다. 사실 일본어의 본 장의 가장 첫 부분에서 언급했듯, 일본어의 피동 표현은 주어가 어떠한 행위를 '받는다' 혹은 '당하다'의 측면이 더욱 강하고, 한국어의 피동 표현의 경우는 '받는다' 혹은 '당하다'의 의미와 더불어 어떠한 행동이 '이루어지는' 측면을 표현하고자 할 때도 피동문을 사용하게 되는 것이다. 이러한 '이루어지는' 측면의 성격이 일본어에서는 자동사로 나타나고, 행위를 '받는다', '당하다'라는 성격의 동작을 피동으로 표현하고 있으므로, 일본어 피동 표현이 보다 일관적인 측면을 가짐과 동시에 피동적 성격과 '이루어지는' 성격의 동사 활용이 더욱 분화되어 사용되고 있다고도 이해할 수 있다. 또한 일본어의 자·타동사의 대응이 한국어와 비교하여 더욱 뚜렷하게 드러나고 있다는 것 또한 알 수 있다.

또한 자·타동사 대응이 있는 일본어 자동사가 한국어 피동 표현에서 '-이, -히', '-리, -기'로만 나타나는 것이 아니라, '-되다'의 형태로도 나타나는 것을 알 수 있다.

<표 2> 한국어 피동 표현과 일치하는 일본어 자동사

'-이, -히', -리, -기'	줄이다, 바뀌다, 닫히 다, 접히다, 열리다, 걸 리다, 말리다, 찢기다	일본어 자동사	減らす `変わる `閉まる `折れる `開く `かかる `乾 かす `破れる
-되다	시작되다, 계속되다,		始まる `続く
-어지다	정해지다, 넘어지다		決まる `倒れる

일본어 자동사와 일치하는 한국어 피동 표현 중, '-당하다, -받다' 류가 없는 것은, 일본어 어휘 중 의미상으로 비교적 확실하게 '-당하다' 혹은 '-받다'의 의미를 가지는 경우는 대부분 피동 표현으로 나타나기 때문이라고 생각하며, 바꿔 말하면 일본어의 자동사의 경우 행동이 '이루어지는' 경우를 나타내고 있기 때문이라고 생각한다. '-어지다' 류의 피동 표현도 마찬가지로, 그 행동이 '당하다' 혹은 어떤 대상으로부터 행위를 '받다'의 의미인 경우는 포함되지 않는 것으로 보인다. 또한 'V+되다'의 구성으로 이루어지는 한국어 동사의 경우에는, 일본어에서는 '-당하다', '-받다'의 의미를 가지는 피동 표현인 'V+される'(される는 する의 피동 표현)으로 나타나거나, 'V+する'의 형태가 의미상으로 '-되다'의 의미를 가지기 때문에 대응이 나타나지 않는다.

한 가지 눈에 띄는 점은, 한국어의 능동 동사가 '한자어+하다'의 구성인 경우, 일본어에서는 자/타의 대응이 일어나지 않음을 알 수 있다. 이것은 한국어에서의 '한자어+하다'의 구성이 일본어에서 동일하게 나타나는 경우, 일본어에서도 동일하게 '한자어+する'의 형태로 나타나고, 한국어에서 이 구성의 어휘가 '한자어+되다'의 형태로 나타나는 경우에는 일본어에서는 '한자어+される'의 형태로 나타나기 때문에, 한국어에서의 '한자어+하다'의 동사가 '-이, -히, -리, -기' 접사를 결합한 형태의 피동 표현이 불가능한 것과 마찬가지로, 일본어에서의 '한자어+する' 형태를 한 동사는 한국어로 옮길 때 파생 피동 형태는 불가능한 것을 의미한다. <표 2>의 시작되다, 계속되다 같은 경우는, 대응되는 일본어가 한자어 구성이 아닌 고유어이므로, 그 의미만을 대응한 것이므로 이에 해당되지 않는다.

이상의 대조 분석을 바탕으로 일본어 자동사와 한국어 피동 표현의 관계에 대해서 정리해 보자면,

- a. 한국어의 피동 표현은, 일본어 피동 표현과는 성격이 다르다.
- b. 한국어 피동 표현은 '-당하다', '-받다'의 의미에 더해 행동이 '이루어진' 상황이나 상태를 나타내는 경우가 포함된다.
- c. 한국어 피동 표현 중 '이루어진' 상황이나 상태를 나타내는 경우는 보통 일본어의 '자동사' 개념과 연결된다.
- d. 일본어에서의 '한자어+する' 구성의 어휘는 한국어의 '한자어+하다'의 구성의 의미 범주와 동일한 경우가 많다. (한자어 의미가 확연히 다르게 쓰이는 경우가 아니라면)
- e. 따라서 한국어의 '한자어+하다'가 파생적 피동을 사용할 수 없듯이, 일본어의 '한자어 + -する' 구성의 어휘 또한 한국어로 옮길 때 파생적 피동을 사용한 피동 표현을 사용할 수 없다.
- f. 일본어의 모든 자·타동사 쌍에서 자동사가 한국어 피동 표현으로 나타나는 것은 아니지만, 본고에서 말하고자 하는 것은 한국어의 피동사가 일본어의 자동사적 성격과 같이 나타나는 것이 있으며, 일본어의 피동 표현이나 기타 언어의 피동적 성격 (예를 들어 영어의 피동태) 과는 다른 측면이 있어 한국어 교육에서도 고려될 사항인 것을 언급하고 싶은 것이다.

따라서 일본어권 학습자의 한국어 피동 표현을 연구하고자 할 때는(한국어 피동 표현의 특성을 생각하면 비단 일본어권 학습자뿐만은 아닐 것이지만) 먼저 한국어 피동 표현의 특성상, 실제 피동 표현을 나타내는 것 이외에도 일본어의 자동사적인 성격을 띠고 있는 어휘가 있다는 것을 인지시켜야 할 것이다. 또한 일본어가 한국어와 마찬가지로 한자 사용이 왕성한 언어인 만큼, 일본어에서의 '한자어+する'의 구성으로 이루어진 어휘는 피동 접미사 '-이, -히, -리, -기' 를 사용한 파생적 피동 표현이 불가능하다는 것을 먼저 학습시켜야 할 것이다. '한자어+ -하다' 형식의 어휘가 파생적 피동 표현을 할 수 없다는 논의가 당연하다고 생각될 수는 있으나, 본고에서 다루고자 하는 것은 특정 언어권인 일본어 언어권 학습자의 학습을 고려한 논의이고(다국적 교실 환경을 고려한 것이 아님), 동일한 한자어권에 속하기 때문에 학습자에게 이를 먼저 인지하게 하는 것이 유용할 것이라 확신하며, 이를 일본어권 학습자에 대한 피동 교육 방안에서 가장 기초적인 지침이 되었으면 하는 의지가 있다. 일본어권 학습자가 L1을 기반으로 한국어 작문을 할 때, 일본어 자동사가 한국어의 피동 표현과 일치하는 것이 있다는 사실이 장애 요소가 되지는 않겠지만, 동일한 파생 접미사를 사용한 어휘의 의미가 자신들이 생각하는 의미와 다른 것에 대한 의문은 일소할 수 있을 것이라는 생각이다. 그리고 한국어 학습에서는 자·타동사의 대응을 학습의 중요한 요소로 삼지 않지만, 일본어 어휘에서는 그 대응이 뚜렷하게 보이는 만큼, L1 언어에서 L2 언어로의 전이라는 측면에서 한국어 피동 표현의 자동사적 성격에 대한 언급이 비단 고급 학습자에게만 유용한 개념은 아닐 것이라는 생각이다. 다만, 학령기의 학습자들에게는 학습자 수준에 따라 역효과가 있을 수 있다는 점은 고려되어야 할 것이다.

2.2.2. 일본어의 피동 표현이 한국어에서는 피동 표현으로 사용되지 않는 경우

(2) 예문에서 먼저 살펴보았듯이, 일본어에서는 빈번하게 피동 표현으로 사용되지만, 한국어에서 피동 표현으로는 사용할 수 없거나, 사용될 경우 수용성이 굉장히 떨어지게 되는 경우들이 존재한다.

- (12) 가. 広場の人々に見られて困りましたが
 hirobano hitobitoni miraretekomarimashitaga
 광장의 사람들에게 보여?/보아져?서 곤란했습니다만
 나. 母には部屋の掃除をしろってよく言われたものだ
 hahaniha heyano soujio sihrotte yoku iwareteimasu
 어머니에게는 방 청소를 하라고 자주 말해지? 고는 했다.
 다. 歌詞を間違ってみんなに笑われた
 kasio matsigatte minnani warawareta
 가사를 틀려서 모두에게 웃겼?/웃어졌?다.

이러한 경우들은 한국어와 일본어의 어휘 인식적 측면에 기인한 것으로 본다. 기본적으로 많이 사용되고, 그 의미역의 차이가 거의 없다고 해도 무방한 단어들인 보다, 말하다, 웃다 (見る `言う` 笑う)가 피동 표현이 되자 그 의미 인식의 차이가 생기는 것을 보면 그렇다고 할 수 있다. 이러한 단어들로 구성된 일본어 피동 표현의 경우, 한국어로 그 의미를 그대로 옮기는 것이 불가능하므로, 일본어권 한국어 학습자들을 대상으로 한 교육의 경우, 일본어권 화자들이 피동 표현으로 사용하는 어휘 중, 한국어 피동 표현으로 옮기는 것이 불가능한 것들을 빈도순으로 먼저 정리하여, (12)의 경우처럼 그 빈도가 가장 높은 어휘들을 기준으로 능동 표현으로 대체하여 사용할 수 있도록 지도할 필요가 있을 것이다.

- (13) 가. 광장의 사람들이 보고 있어서 곤란했습니다만
 나. 어머니에게 방 청소를 하라는 잔소리를 자주 들곤 했다.
 다. 가사를 틀려서 모두에게 웃음을 샀다.
 다'. 가사를 틀려서 모두에게 비웃음당했다.

(13)의 대체 표현 중에서도 (13 나)의 경우는 일본어의 피동 표현이 많은 경우에 '피해' 혹은 '당하'는 상황을 기저에 두고 있는 경우가 많으므로, '잔소리'라는 어휘를 통해 일본어의 '言われる'의 기저 의미를 더하는 방법을 제시한 것이며, (13 다)의 경우 또한 마찬가지로 일본어 표현에서 '웃다'를 '웃는 행위를 당하다'라는 의미로 사용하였을 때의 의미를 한국어적 표현으로 대신하여 '웃음을 사다'로 대체하거나, 혹은 한국어에서 사용할 수 있는 피동 표현으로서 '비웃음 당하다'의 표현으로 대체할 수 있도록 하여 학습자들에게 적합한 대체 방식을 제시할 수 있을 것이다.

2.2.3. 일본어 피해 피동 표현의 한국어적 표현으로의 전이

일본어의 피해 피동 표현은, 일반적으로 생각하는 영어의 수동태 표현이나, 한국어의 피동 표현과는 다른 양상을 보이고, 일본어만의 독특한 표현 양상으로 생각된다.

- (14) 가. 私は雨に降られて `服が濡れてしまった
 watashiwa ameni hurarete, hukuga nurete simatta
 나는 비에 내려져서?/내리게 되어서?, 옷이 젖어 버렸다.

- 나. 彼は夕べ赤ちゃんに泣かれてひと眠りもできなかつたらしい
 watashiwa yuube akatsyanni nakarete hitinemurimo dekinakatta
 그는 어제 아기에게 울려서?/울어져서?/울게 되어서?한숨도 자지 못했다고 한다.
- 다. 私は去年 `母に死なれた
 watashiwa kyoneN, hahani shinareta
 나는 작년엔, 어머니에게 죽어졌다?/죽게 되었다?

이러한 표현의 본질은, 여격으로 나타난 동작주가, 어떠한 형태로든 주어가 피해를 당한 경우를 말한다. 즉 (14 가)의 경우는 '나는 비가 내리는 것(행위)을 당했다', (14 나)는 '아기가 우는 것(행위)을 당했다. (잡을 못 간 것이 나의 피해 상황), (14 다)는 '어머니가 죽는 것(행위)을 당했다'(어머니가 죽어서 심리적이거나 혹은 경제적인 피해를 당하였다) 라는 의미가 된다. 이러한 독특한 표현방식은 사실 한국어의 피동 표현이나 기타 언어의 피동 표현과는 확연히 구별되는 특징을 가지고 있어서, 표현 시의 오류보다는 어떻게 이것을 한국어 피동으로 표현하느냐가 문제인데, 이는 한국적인 적절한 표현을 제시해 주는 것 이외에 문법 구조적인 문제로 해결할 수는 없을 듯하다.

- (14) 가. 나는 비를 맞아서, 옷이 젖어 버렸다.
 나. 그는 어제 아이가 울어서, 한숨도 자지 못했다.
 다. 나는 작년엔 어머니가 죽었다(돌아가셨다).

이러한 표현을 한국어의 피동 표현으로 나타낼 문법 구조는 존재하지 않는다. 일본어권 한국어 학습자는 일본어의 피동 표현이 한국어의 피동 표현의 본질과는 많이 다르다는 사실을 먼저 인지해야 한다. 한국어의 피동 표현은 일본어권 화자가 주로 사용하는 '피해', '당하는' 표현들과는 달리, 앞선장에서 언급했듯, 일본어의 자동사와 같이 행동이 '이루어지는' 류의 표현들이 많기 때문이다. 따라서 한국어로 바로 옮겨서 사용하려는 방식 보다는, 능동문을 먼저 떠올려 사용하는 것에 익숙해지도록 하는 것이 초-중급 한국어 학습자들에게는 더욱 유용한 방식이 될 것으로 생각한다. 이에 다음과 같은 기준을 통해 학습자들에게 도움을 줄 수 있을 것이다.

a. 피동 표현의 동작이 주어, 혹은 주어의 소유물(가방, 지갑, 손, 발, 어깨 등)에 직접적으로 작용하는 경우는 한국어 피동 표현을 적절히 사용할 수 있도록 한다.

- (15) 가. 私はバスで足を踏まれた
 watashiwa basude ashio humareta
 나. 나는 버스에서 발을 밟혔다.
 다. 弟は私にお尻をけられた
 otoutowa watashini oshirio kerareta
 라. 동생은 나에게 엉덩이를 차였다.
 마. 泥棒が警察に捕まれた
 dorobouwa keisatsuni tsumakareta
 바. 도둑이 경찰에게 붙잡혔다.

- b. 피동 표현의 동작의 방향이 자신이 아닌 다른 곳을 향하거나 불특정한 방향을 향하는 경우에는 가급적 능동문을 사용하도록 학습시킨다.

(16) 가. 私は仲間に行かれて `仕事を進めなかった

watashiwa nakamani ikarete, shigotoo susumenakatta

나. 나는 동료에게 가져서?/가게 되어서? 일을 진행하지 못했다.

나'. 나는 동료가 가버려서, 일을 진행하지 못했다.

다. 妹は弟に周囲を回れて `全然勉強できなかった

imoutowa otoutoni syuuio mawarete, zenzen benkyou dekinakatta

라. 여동생은 동생에게 주위를 돌아져서?/돌게 되어서? 전혀 공부할 수 없었다.

라'. 여동생은 동생이 주위를 맴돌아서, 전혀 공부할 수 없었다.

마. 彼は雪に降られて `遅刻してしまったそうだ

karewa yukini hurarete, chikokusitetimatta

바. 그는 눈에게 내려져서?/내리게 되어서? 지각하고 말았다.

바'. 그는 눈이 내려서, 지각하고 말았다.

2.2.4 일본어 사동 피동 표현의 한국어적 표현으로의 전이

일본어에는 사동 피동 표현이라는 것이 있다. 의미는 1장에서 언급하였듯이, 사동 하는 행위를 당하다. 라는 의미인데, 조금 더 풀어서 얘기하자면, 나는 하기 싫었으나(그러한 감정이 포함된다) 여격으로 나타나는 동작주가 나에게 일정한 행위를 시켰고 나는(마지못해) 그 행위를 당했다(하게 되었다)는 의미이다.

(4) 가. 居眠りしすぎて `私は先生に立たされたまま `授業を受けた

inemuri shisugite, watashiwa senseini tatasaretamama, jyugyououketa

나. 너무 좋아서, 나는 선생님에게 서게하는것을 당한? 채로, 수업을 들었다.

이것은 다른 언어에서는 찾아볼 수 없는 독특한 표현방식으로, 일본어권 학습자가 이것을 한국어로 수월히 옮길 수는 없을 것 같다. 이 경우, 본질적으로 사동 피동 표현 또한 중점은 피동(피동)에 있지만, 이것을 중점으로 말하고자 한다면 표현이 어려워지는데, 이는 '서다-세우다' 처럼 사동 표현은 용이하지만, 이것을 피동으로 표현하기가 힘들기 때문이다.

(14) 나는 너무 좋아서, 선생님께 세워진? 채로 수업을 들었다.

이 경우는 '세우다' 어휘의 특성상 피동 표현이 자유롭지 못하기 때문인데, 이와 같은 상황을 벗어나기 위해서는 부득이 피동보다는 사동의 의미에 중점을 두는 편이 낫다.

(15) 가. 내가 너무 좋아서, 선생님께서 자리에 세웠고, 그대로 수업을 들었다.

나. 내가 너무 좋아서, 선생님께서 나를 자리에 서게 한 채로 수업을 듣게 했다.

이처럼 일본어 사동 피동 표현의 경우, 일본어에서는 '억지로', '당하다', '어쩔 수 없이'라는 의미로서 피동 표현적인 의미가 강한 것이 맞지만, 이것을 한국어로 표현하고자 할 때 적절한 표현이 되지 못할 수 있기 때문에, 이때는 부득이 피동적인 표현을 빼고 표현을 하는 것이 한국어적으로는 자연스러운 표현이 될 수 있다. 다만 이것은 '세우다' 어휘만의 문제는 아니고, 다른 어휘에서도 동일한 문제가 발생함을 알 수 있다.

- (16) 가. 飲み会で私は仲間たちに帰らされた
 nomikaide watashiwa nakamatatini kaerasareta
 나. ?술자리에서 나는 동료들에게 돌아가게 되었다.
 나'. ?술자리에서 나는 동료들에게 돌아가게 해졌다.

따라서 위의 (15) 같은 해결법이 현재로서는 가장 적절한 방법인 듯하다. 이러한 특수한 표현의 경우는 오히려 오류가 발생할 가능성은 적으나, 이를 어떻게 표현할지에 대한 문제에서는 학습자들에게 도움을 줄 수 있을 것으로 생각된다.

3. 결론

본고에서는 한국어 피동 표현의 특징과 학습법을 연구하기 위해, 일본어와의 대조를 통해 그 특징을 살펴보고, 학습 전략에 대해 구상해 보았다.

첫 번째로는 한국어의 피동 표현과 일본어의 피동 표현의 특징을 살펴보았다. 한국어의 피동 표현은 '-이, -히, -리, -기'와 같은 피동 접미사가 각각의 접미사와 결합될 수 있는 어휘가 제한되어 있고, 그 생산성이 낮다고 언급하였다. 또한 행동이 '이루어진' 결과나 상태를 나타내는 표현이 많았고, 행동을 '당하거나' '받는다'는 것을 나타내는 실제적 피동 표현들과는 다르게 많이 희석된 느낌으로 자리하고 있다고 분석하였으며, 일본어의 피동 표현의 경우 조동사인 '-れる' '-られる'를 동사와 결합하여 거의 제약 없이 사용 가능하다는 점, 그리고 그 본질적인 의미에 있어서도 한국어 피동 표현과는 달리 피동 표현의 동작주가 주어에게 어떠한 행동을 하고 주어는 그 행동을 '당하'거나 '받'는다는 것에 중점을 두고 있는 듯하고 분석하였다.

두 번째로는 한국어 피동 표현과 일본어 자동사와의 관계성에 대해 분석하여, 한국어 피동 표현은 일본어의 자동사와 같은 의미를 가진 어휘들이 다수 존재하며, 한국어 피동 표현이 가지는 피동성에 대한 논의를 진행하였으며, 한국어의 경우 타 언어에 비해 피동적인 의미가 많이 희석되어 있다고 평가하였고, 이를 학습자들이 어떻게 학습에 적용하면 될지에 대한 방향성에 대해 논하였다.

세 번째로는, 일본어 피동 표현에서는 가능하지만, 한국어에서는 피동 표현으로 사용하지 않는 동사들과 이에 대한 표현방식을 어떻게 하는가에 대한 문제에 대해 논하였으며,

네 번째로는 일본어의 피해 피동 표현은 무엇이며, 피해 피동 표현과 일반 피동 표현에 있어, 어떤 식으로 일본어권 화자가 의미를 파악하고 피해 피동으로 생각되는 표현을 한국어로 전이할 때는 가급적 능동문을 사용하도록 하고, 일반 피동의 경우 한국어 피동 표현으로 옮겨서 사용하는 연습을 할 수 있도록 하자고 제안하였다.

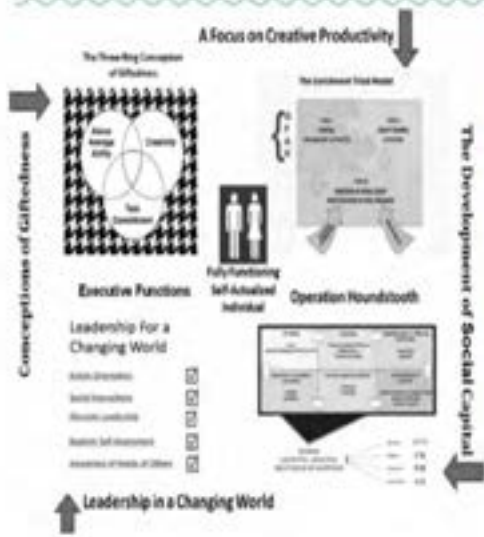
다섯 번째로는 일본어의 사동 피동 표현의 의미는 무엇이며, 사동 피해 피동 표현은 한국어로 온전히 옮기기 어렵다고 보고, 이를 한국어적 능동 표현으로 전이하는 방법에 대해서

논해 보았다.

이 연구를 통해 한국어 피동 표현에 대해 더욱 깊이 고찰하고, 한국어 피동 표현을 일본어권 화자들이 학습하고자 할 때, 부딪힐 수 있는 문제들과 피동 표현에 대한 기타 언어들의 양상과 한국어 피동 표현만의 특징, 그리고 일본어 피동 표현의 특징에 대해 심도 있게 연구하고 기존의 연구들 또한 깊이 있게 탐구할 수 있었다고 생각한다. 다만 각각의 분석 내용에 따른 구체적인 교수법을 제시하지 못한 점에 대해서는 매우 아쉽게 생각한다. 또 유사한 주제로 실제 학습 현장에서의 교수 방안이나 유의점에 대해서 더욱 심도 있게 다루는 것을 향후 과제로 남겨두고자 한다.



1. 서론



A four-part theory of talent development(Renzulli, 2012)

내러티브 창의적 글쓰기 -> 창의성 +언어 능력향상

내러티브 창의적 글쓰기는 창의력과 상상력을 발휘하게 하는 기회인 동시에 이를 향상시킬 수 있는 도구, 뿐만 아니라 이를 통해 언어적 유창성과 동사적 지식을 표현하는 기회

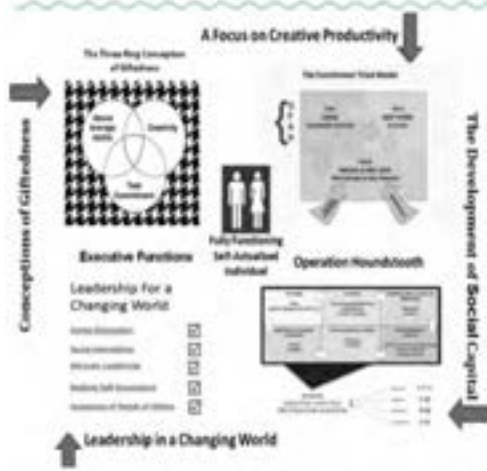
영어 고능력 학습자, 영어 영재?

*영재성은 세 가지 분야의 상호작용 - 평균이상의 일반능력, 높은 수준의 과제 집착력, 수준 높은 창의성 -으로 발휘된다고 보고, 수준 높은 창의성을 영재의 주요 특징으로 본다(Renzulli, 1978).
*영어영재는 이러한 일반적 영재성을 지니고, 영어에 재능이 있으며 영어로 산출물을 만드는 것에 높은 과제 집착력을 보이는 재(이)주영 김창열, 2021)로 본다.

영어영재 프로그램 교육과정 요구

영어 고능력 학습자 혹은 영어 영재의 영어영재 프로그램 교육과정에 대한 요구 -문학, 문화, 건학문적 주제 중심의 토의 토론을 통한 고차원 사고력 향상 기회 및 언어 표현 기능 향상, 비문학 창의적 사고력 향상에 대한 요구가 높음 (이주영 김창열, 2021)

1. 서론



A four-part theory of talent development(Renzulli, 2012)

중학교 영어 고능력 학습자 맞춤형 프로그램 개발

'Change'를 중심으로 문학 작품을 속 건학문적 주제를 토의 토론해 심화 개념 정립 및 과정 산출 중심의 심화 학습 활동 전개. 영재학생을 위한 통합교육과정 모델을 언어영재학생에게 적용함. 즉, 문학 작품 내용 분석을 위해 추론바퀴(Reasoning Wheel)활동, Literary Questions, Change Questions 등을 활용해 이해 중심에서 점차 표현 중심의 내러티브 창의적 글쓰기 최종 산출물 생성으로 전개. 또한, 심부 심화로 조별 활동에서 개인적 자기주도형 중심으로 글쓰기 활동 전개. 주장성과 책임감이 심화되는 과정 산출 중심의 활동 전개에 ICM 기반 프로그램 완성

통합교육과정모델(ICM) 기반 내러티브 창의적 글쓰기 프로그램 개발에 대한 효과성 분석 필요

내러티브 창의적 글쓰기 평가가 직접적인 평가 방법임. 쓰기 평가 자체가 교사에게 어려운 난제임(Discourses of writing and learning to write, 2004). 영어 쓰기 능력이 글로벌 인재의 핵심역량으로 인식되고 중학교 영어 고능력 학습자 또는 영어영재를 위한 쓰기 평가에 적합한 평가 방법 필요

연구 목적

통합교육과정모델(ICM) 기반 내러티브 창의적 글쓰기 수업을 영어영재학생들에게 실제로 적용해보고 창의성과 쓰기에 미치는 효과가 있는지 확인하고자 한다.

연구문제

1. 통합교육과정모델(ICM) 기반 내러티브 창의적 글쓰기 수업이 영어영재학생의 **창의성**에 미치는 효과가 있는가?
2. 통합교육과정모델(ICM) 기반 내러티브 창의적 글쓰기 수업 후 **코메네틱스(Coh-Mentix)** 분석 결과 영어영재학생의 내러티브 창의적 글에서 사전과 비교해 어떤 언어적 변화가 있는가?

2.1 창의적 내러티브 글쓰기에서의 창의성 평가

2.1.1 창의성

Guilford(1950) - 창의성을 다양하고 많은 정보를 새로운 시각에서 아이디어나 산출물.
 Torrance(1968) - 어떤 문제에 직면했을 때 문제를 해결하기 위해 적절한 질문을 할 수 있는 능력, 문제를 해결하기 위한 예측이나 가설을 형성할 수 있는 능력, 적절한 예측이나 가설을 검증하거나 재검증하여 가장 적절한 해결책을 찾아낼 수 있는 능력, 문제 해결력

창의적 사고- 인지적 요소, 하위 요소로 포함되는 것이 유창성, 융통성, 독창성, 정교성(Guilford, 1967, 1971)이다.

창의적 발달- 개인차적 요소, 창의성의 성격적 특성 강조함. 개인의 성격적 요인, 특성, 성격구조에 초점

창의성 발달의 쓰기에 중심적이고, 과업 기반 쓰기는 EFL 학습자에게 창의성에 도움을 준다(Ward and Dadaan, 2012)

Ward and Cheng(2015) 창의성론 제 2면에 수행과 학습 성취에 있어 등가에 영향력을 준다.

2.1.2 글쓰기와 창의성

*글쓰기와 창의성은 상호적으로 결합하여 있어 서로 연관됨
 목표로 하는 언어를 창의적으로 사용하는 능력과 새로운 발화를 무수히 다양하게 표현하고 이해하는 능력으로 주로 언어
 어떤 연구에서는 이미 알고 있는 것이 없지 못하는 것을 넘어서 구체적인 표현을 넘어서 표현하는 능력으로 보고, 언어적 능력(proficiency)에 초점을 두어 주로 언어적 창조성에 제한함. 하지만 최근에는 글쓰기 자체가 창의적 사고를 드러내는 능동적 행동의 과정(Bakhtin & Dargatzis, 2010)이며 창의적 글쓰기가 학생의 창의성을 높이는 효과적인 도구로 여김(Cademir, S. & Celik, A., 2018)

2.1. 창의적 내러티브 글쓰기에서의 창의성 평가

2.1.3 창의성 측정 도구

*Divergent thinking test가 창의성을 이루는 요인의 기능에만 치중했다는 점에서 비판받음.

*Torrance Test of Creative Thinking(TTCT)

- 1) Form A(Thinking Creatively with Words)-verbal creativity를 측정
 -문고 추측하기, 연인과 풍자 추측하기, 산출물 상상하기, 형용사 없는 활용, 가정해설하기
- 2) Form B(Thinking Creatively with Pictures) nonverbal 또는 formal creativity를 측정
 -그림 구성, 그림 완성, 언어나 한 글자

언어 없애나 언어 없애제의 창의성 측정 이들에게 좀 더 실제적이고 다양한 평가가 될 수 있도록 **학생 언어 산출물인 쓰기 샘플을 측정하는 것이 바람직함**(Pahghaderi & Mehr, 2011)

2.1.4 내러티브 창의적 글쓰기 채점 rubric

창의적 산출에 대한 독자에 관한 연구가 많이 이루어지지 않았지만(Besemer & Treffinger, 1981; Starko, 2014), Creative Product Assessment Matrix (Besemer & O'Quin, 1986)와 학생 산출 평가 양식(Johnson & Reis, 1991)에는 독창성/창의성이 구인으로 포함됨

Creative Product Assessment Matrix에서는 정교함이 창의성 구인으로 추가적으로 포함되었다. 창의성과 독창성은 문맥 독창성에서 이미 널리 인정.

Pfister(1991)는 아이들의 창의적 글쓰기는 전통적으로 독특한 정교와 의미지향, 문구, 생생한 묘사, 유머 등을 포함.

쓰기 문헌에서의 나타나는 창의성은 크게 독창성과 정교성임(Kettler, T., & Bower, J., 2017)

Suzar Kermiz(2008)가 개발한 "Evaluation Criteria Form for Creative Writing"에서는 다음과 같은 기준을 부분적에 반영함.

- 명백하지 않은 새로운 아이디어가 텍스트에 존재하는지, 특별한 방법으로 특별한 요소를 묘사하는지 실제적인 맥락에 포함하여 있는지, 새로운 아이디어가 명확하게 표현되었는지,
 - 느낌과 생각이 효과적이고 융통하게 표현되었는지, 적절한 제목이 사용되었는지
- 무엇이든에서 개발된 창의적 측정지표 및 도구개발 연구(2012)에서도 아이디어 특성상론 그대로 포함되고, 정교성은 완성도, 직관성과 학습성으로 세분화됨

2.2. 코메트릭스를 활용한 언어 분석



2.2.1. 코메트릭스(Coh-Metrix) 소개

미국 캘리포니아대학교 저능형시스템연구소에서 개발한 Coh-Metrix Linguistic을 분석 도구로 활용한다.

Coh-Metrix 3.0은 Coh-Metrix는 언어학적 특징을 세부 지문 중 206개의 지문으로 컴퓨터로 자동 분석 할 수 있는 컴퓨터 언어학 도구이다. 측정치가 나타내는 공통된 언어적 속성에 따라 11개의 그룹으로 분류된다.

이들 11개의 그룹은 기초 산출지(descriptive indices), 텍스트 용이성 요인 지수(text easability principal component scores), 참조적 응집성(inferential cohesion), 잠재적 의미 분석(latent semantic analysis), 어휘 다양성(lexical diversity), 접속사 밀도(connectives), 상황 모형(situation mode l), 문리적 복잡성(syntactic complexity), 문사 구조 밀도(syntactic pattern density), 단어 정보(word information), 그리고 문장 어휘성 지문(readability)이다(McNamara et al., 2004)

2.2. 코메트릭스를 활용한 언어 분석



2.2.2. 코메트릭스(Coh-Metrix)(McNamara et al., 2004)

어휘 다양성

어휘적 다양성은 단어(토큰)의 총 개수와 텍스트에서 발생하는 고유한 단어(유형)의 비율로 나타낸 단어 유형의 수가 총 단어 수(토큰)과 같은 경우 모든 단어가 다름.

이 경우가 어휘적 다양성 최대치임, 텍스트의 응집력은 매우 낮거나 매우 높을 가능성이 있음.

TTR(Type-token ratio)은 고유 단어(유형이라고 함)의 수를 이 단어들의 토큰 수로 나눈 것.

텍스트의 각 고유 단어는 단어 유형으로 간주됩니다. 특정 단어의 발생수는 무관하다고 함.

예) 단어 dog이 텍스트에서 7번 나오면, 유형(type)은 1, 토큰은 7, TTR = 1/7.

잠재 의미 분석(LSA)

문장 간 또는 문단 간의 의미 중복 측정용 제공.

Coh-Metrix 3.0은 8개의 LSA 지수를 제공하고, 각 측정값은 어휘론 응집력에서 1(높은 응집력)을 나타냄.

참조 응집력(Referential Cohesion)

인접 문장 간의 내용 단어의 중복 또는 공통 참조를 의미함.

공통 참조는 독자가 텍스트 기반 이해에서 현재, 잘 될 문장을 연결하는 데 도움이 되는 언어적 단서(Haliday & Hasan, 1970; McNamara & Kintsch, 1996).

참조 응집력에 대한 Coh-Metrix 측정값은 지역에서 전체(어려서까지 2차원에서 측정됨).

국소(local) 응집력은 연속적인 인접 문장 간의 중복을 평가하여 측정하는 반면, 전체(global) 응집력은 단어 또는 텍스트의 모든 문장 간의 중복을 측정.

2.2. 코메트릭스를 활용한 언어 분석



2.2.2 코메트릭스(Coh-Metrix)(McNamara et al, 2004)

Flesch Reading Ease: RDFRE (index 104)

Fleesch Reading Easy 글자의 음역은 0에서 100 사이의 숫자. 점수가 높을수록 읽기가 더 쉬움. 평균 문장의 길이가 읽기 쉬운 점수는 6에서 70 사이

Flesch_Kincaid Grade Level: RDKGL (index 105)

콜레어-킨케이드 등급 공식은 읽기 쉬운 점수(RDFRE)를 미국 초등학교 수준으로 변환 숫자가 높을수록 본문을 읽기가 더 어려움. 등급 수준은 0에서 12 사이.

통사적 복잡성(Syntactic Complexity)

통사론은 단어를 품사 범주(명사, 동사, 형용사, 접속사)에 할당하고, 단어를 구문 또는 구성 요소(명사, 동사구, 전치사구, 절)로 그룹화하여, 문장에 대한 구문 트리 구조를 구성.

어떤 문장들은 짧고, 단순한 문법을 가지고 있으며, 행위자-행동-목적어 구문 패턴을 따르고, 포함된 할미거의 없으며, 수동적이지만 능동적인 목소리를 가진 (통사적 복잡성이 낮음)

일부 문장은 작업 메모리를 더 많이 요구하는 복잡한 문법을 가지고 있다. 본문의 구문은 짧은 문장이 있을 때, 본문의 본문사 앞에 몇 개의 단어가 있을 때, 명사구 앞 몇 개의 단어만 있을 때 처리하기가 더 쉬운 문장이 있다

통사적 패턴 밀도(Syntactic Pattern Density)

구문 복잡성은 또한 특정 구문 패턴, 즉, 단어 유형 및 구문 유형의 밀도에 의해 결정됨

Coh-Metrix는 명사구(DRNP), 동사구(DRVP), 부사구(DRAP), 전치사(DRPP)의 발생에 대한 정보를 제공.

각각의 상대적 밀도는 특히 텍스트의 다른 기능과 관련하여 텍스트의 처리 난이도에 영향을 미칠 것으로 예상. 어떤 텍스트가 더 높은 명사구 동사구 발생률을 가지고 있다면, 그것은 복잡한 구문으로 정보적으로 밀도가 높을 가능성이 더 높음.

2.2. 코메트릭스를 활용한 언어 분석



이해성 (난이도)	기초산출지 (Descriptive Indices)	3. DESWC (총 단어수) 6. DESSL (평균 문장 길이)
	표준가독성 지표	104. RDFRE (FRE점수) 105. RDKGL (FKGL점수)
	통사적 복잡성 (Syntactic Complexity)	67. SYNLE (본문사 할 단어수) 68. SYNMP(명사구 할 수식어구)
어휘적 특징	통사적 패턴 밀도 (Syntactic Pattern Density)	74. DRNP (명사구 밀도) 75. DRVP (동사구 밀도) 80. DRGERUND (동명사 밀도) 81. DRINF (부정사 밀도)
	어휘 정보 (Word Information)	95. WRDAOAc (언어습득 L2)(96. WRDFAMc (친숙도) 97. WRDCNCc (구체성) 98. WRDMGc (심상성) 99. WRDMEAc (의미성)
	어휘다양성 (Lexical Diversity)	46. LOTTRc (타입 토큰 비율-내용어 기본형) 47. LDTTRa (타입 토큰 비율 - 모든 단어)
담화적 특징	참조적 의미 분석(LSA) -의미적 응집성 -의미적 유사성	38. LSASS1(인접 문장 간 의미적 응집성) 40. LSASSp(전체 문장 간 의미적 응집성) 44. LSAGN(구절보/산정보 응집성)
	참조적 응집성 (Referential Cohesion) -참조적 정합성	28. CRFN01 (인접 문장 간 명사 중복) 29. CRFA01 (인접 문장 간 논항중복) 30. CRFS01(인접 문장 간 위근 중복)
	접속사 (Connectives)	50. CNCAU(전체 접속사 수) 51. CNCAU(인접 접속사 수) 52. CNCLogic(논리 접속사 수) 53. CNCAO(역접 접속사 수) 54. CNCTemp(시간 접속사 수) 58. CNCAAdd(참가 접속사 수)
	장문(Namativity)	12. PCNARz(유러티브 장문 정도)

2.3. 선행 연구



2.3.1. 창의적 내러티브 글쓰기에서의 창의성 평가 선행연구

한정진(2020)은 고등학교 학습자를 대상으로 토론 기반 영어 쓰기 학습 활동을 진행하고 창의성, 쓰기 및 창의적 자기평가의 시후 세교에 실험집단이 비교되었다. 대부분 학생은 글을 확인한, 제작성-이론적(2016)은 언어통합프로그램의 교육학생의 영어 쓰기 능력과 창의성 향상에 교육적 인식을 확산하기 위해 고등학교 1,2학년 상위 영어 학습자 집단을 대상으로 영어 쓰기 능력 시험과 창의성 검사를 사전 설문 조사한 것을 분석해 실험집단과 통제 집단에서 쓰기 능력의 창의성이 유의한 차이가 나는 것을 확인해 프로그램 효과성을 입증.

김승희와 김형섭(2007)은 초등학교 언어 영역 하위과 및 언어 영역의 언어적 창의성 발달에 관한 연구에서 초등학교 언어 영역 아동과 일반 아동의 언어적 창의성의 발달을 글쓰기에 초점을 두어 살펴보았다. 학년별 언어 영역과 일반 아동들의 글쓰기를 측정하고 관련된 언어적 능력을 실시해 이를 두 집단의 확산적 사고와 수렴적 사고 수준을 유의한 차이가 있는지 확인했고, 그 결과 학년이 높아지면서 확산적 사고와 수렴적 사고 모두 일반 아동과 언어 영역 아동이 발달하고 있는 것으로 나타났고, 일반 아동에 비해 언어 영역의 확산적 사고와 수렴적 사고의 발달 수준이 더 높은 것으로 확인됨.

황재영의 주창서로 및 도구개발 연구(김어유, 2022)
 미국의 경우 언어를 활용한 창의적 산출물에서 나타나는 창의적인 창의적인 생각을 갖는 것과 새로운 아이디어에 대한 가치를 평가로 구분. 이는 Bloom의 original taxonomy를 개정한 Anderson and Krathwohl(2001)의 revised Bloom taxonomy에서 Create에 해당.
 Create란 창의적 사고력과 비판적 사고력 둘 다 포함하는 것으로 창의적 능력인, 활동이 창의적 산출물에 대한 기대를 동일하게 할 때 포함되는; 오소리 학생의 수준, 연령적 사고, 그리고 평가에서도 포함
 창의적 산출에 대한 특성에 관한 연구(Besemer & Taylor, 2013; Starks, 2014)
 Creative Product Assessment Matrix (Besemer & O'Quin, 1989)
 학생 산출 평가 양식(Rensell & Reil, 1991)은-특정성,정신성이 구안으로 포함되었다.
 창의적 글쓰기 평가에서의 무의미에 관한 연구(Yoon & Reibel, 2019),
 What characterizes creativity in narrative writing, and how do we assess it?(Saxos, 2021)
 Re-imagining narrative writing and assessment: a post-NAEPAN craft-based rubric for creative writing(Cery, Davidson & Williams, 2022)

2.3. 선행 연구



2.3.2. 코메트릭스 관련 선행연구

- 국내에서는 주로 읽기 지문의 난이도와 복잡성, 이해정보 등을 분석 coh-matrix를 활용해 분석했는데, 이 교재를 분석한 연구(김장영, 2014; 김현영, 임지은 2012; 김지은, 전우기 2013; 이지영 2015; 송혜성 2016; 이경민, 한태희 2013; 양인애, 전우기 2013; 전우기 2011; 전우기 2015; 전우기, 양인애 2009; 이숙은 2010; 성달우 2014)이 주로 역향에 2020)
- 대체 가능 문제나 EBS 자료 분석(김정영 2017; 김민경 2018; 신유선 2018; 최민우, 김형영, 2017; 황다수, 이계영 2020; 연혜선, 이지영 2021; 유주연 2018;정수현 2020;계정근, 이종우 2021)에 활용됨
- 최근에는 학습자 쓰기 분석에도 코메트릭스를 활용함(이재희, 김정영 2022; 한문기, 최복희 2015; 이호 2019; 이우 2022;전수현, 2020; 이은희, 2014; 김준우, 임규원 2019; 정유구, 신명미, 2021).
- 해외에서는 쓰기 분석에 있어 코메트릭스를 활용해 학습자의 언어 분석에 관한 연구가 다수 존재(Crosley, Weston, McLain Sullivan, & McNamara, 2011; Graesser, A. C., McNamara, Cai, Conway, Li, & Pennebaker, 2014; Crosley & McNamara, 2011; Chen & Shin, 2020; Aryadoust & Liu 2015; Lai & Man & Ting, 2014; Babe & Nitta, 2010; Babe & Nitta, 2012; & Letfi & Gierl, 2021)이고, 언어 창의성 평가에 있어 coh-matrix를 활용해 언어를 분석한 연구도 해외에서는 진행. (Taxaocik & Pennebaker, 2010; Zedelius, Mills & Schooler, 2019; Wolfe & Dandignac, 2011; Letfi & Gierl, 2021; Lim 2019; Plate & Hubson, 2022)하나 국내에는 평가 도구를 비교할 때 언어 평가가 글쓰기를 진행하고 이를 코메트릭스에서 분석(Suh & Bae, 2022)하는 연구가 매우 적음.

3. 연구의 방법

양적 연구 방법

3.1. 연구대상

본 연구에서는 광역시 소재의 한 고등학교에서 진행되는 중학생 대상 영어영재프로그램에 참여하는 1학년 학생 2학급으로 구성된 40여명을 대상으로 실험반 20명 비교반 20명으로 진행.

이들은 영어영재프로그램에 참여하기 전에 교사 및 학교장의 추천을 통해 1단계 선발에서 통과하고, 2단계 전국단위 창의적 문제해결력 검사(영어 및 창의력 검사)를 거쳐 영재학급 대상으로 승인을 받음. 실험 집단과 비교집단 모두 영어 문학 읽기 자료를 기반으로 영어 수업을 받았지만, 실험집단은 ICM-상부심화 모형 기반 내러티브 창의적 글쓰기 수업의 실험 처치를 받는 반면, 통제집단은 기존 영어 문학 읽기 자료에 대해 이해를 점검하고 등장 인물 이해, 독후 활동 및 감상 표현 중심의 기존의 영재수업 진행방식을 따름.

3.2 연구진행 절차

연구방법	창의 내러티브 수업 및 실험 처치		영역분석	
연구시기	2023년 4월 사전검사 내러티브 창의적 쓰기 평가	중학교교육과정모형(ICM) 기반 14주 프로그램 치치 7월 초- 창의적 산출물 제출: 조별 도의 후 연달 새롭게 창작하기 8월 중순-개인별 내러티브 창의적 산출물 프로젝트 완성본 온건한 제출	2023년 9월 사후검사 내러티브 창의적 쓰기 평가	-루브릭 활용 창의성 측정 -Coh-Metric 3.0 활용 Text Analysis
실험집단	○	○	○	○
비교집단	○	X	○	

3. 연구의 방법

양적 연구 방법

3.3. 평가도구(Change Through Choice, 2020)

A Likely Story

Look at the picture and create an interesting story related to it with your imagination.

Include what is likely to happen just before this picture and what will happen afterwards.



3.4. 내러티브 창의적 글쓰기 채점 루브릭

본 연구에서는 Carey, M. D., Davidow, S., & Williams, P. (2022)에서 호주의 The National Assessment Program - Literacy and Numeracy (NAPLAN) 루브릭과 Zedelius, C. M., Mills, C., & Schooler, J. W. (2019)에서 creative writing 루브릭에서 평가요소를 추출하여 본 연구 대상자인 영어영재학생 1학년의 creative writing에 맞게 조정함.

본 검사의 채점자는 한국에서 영어영재 학생을 대상으로 교육활동용 한 원어인 1인과 한국어인 영어교육전공 박사과정생 1 총 2명이 채점. 이들이 점수를 채점한 다음, 토를 에세이 채점 점수 조정 방식처럼, 급간이 2(10점) 이상 차이가 날 경우, 2인이 의논하여 조정 점수를 채점함.

두 채점자 간 신뢰도 계수는 크론바흐알파 사전평가 0.783 사후평가 0.733으로 채점자간 신뢰도가 확보됨

3. 연구의 방법

양적 연구 방법

평가항목	내용	급간	배점
이미지(image)	시각, 청각, 냄새, 물리적 감각, 감정 등을 포함한 선명한 심상을 떠올리게 하는가? 독자군(하어급 등장인물)의 몸과 마음에 몰입할 수 있도록 구체화되었나? 사건에 대해 풍부한 묘사를 제공하는가?	5점	20
보이스(voice)	작가만의 독특한 목소리나 글쓰기 스타일이 담겨있나? 특정 단어 선택, 새롭게 만든 어휘, 흥미로운 문장 구조나 구두점이 있는가? 문체나 풍자, 의도적으로 문장 단순성이나 복잡성 등을 통해 소설 속 화자의 독특한 관점이나 태도가 표현되었나? 단락 간 빈 공간, 대문자, 문유, 띄어쓰기 등 수학적 표현을 작가만의 스타일이 표현되었나?	5점	20
독창성 (originality)	다른 학생들과 다르게 새롭고, 혁신적이고 독창적이지만 여전히 이미가 전달해 성공적인가? 예상치 못한 놀람이나 선택이 있는가? 참신한 주제, 독창적 배경, 독창적 플롯, 명백하지 않은 이야기 구조나 평범하지 않은 연등, 스타일, 감정뿐이 있는가?	5점	20
아이디어 (idea)	복잡한 아이디어나 간학문적 주제를 담고 있는가? 아이디어가 높은 중요도를 보이는가? 아이디어가 높은 연결성, 나태지는가? 왜 만약 아이디어를 유창하게 만들어 내는가?	5a	20
읽기 몰입감 및 명료성 (engagement & clarity & lack of distraction)	Hook로 시작 부분에서 독자를 사로잡는가? 높은 응집성을 보이는가? 문장의 다양화로 몰입감 있는 흐름을 만들어 내는가? 지속적인 명확한 면에 선택과 서술적 장치의 사용으로 독자의 읽기에 몰입감을 주는가? 문법적 오류나 스타일 오류로 인한 이해 흐름을 방해하는 것이 적은가?	5점	20

3. 연구의 방법

양적 연구 방법

3.5. 영어언어분석도구(coh-matrix)

본 연구에서는 프로그램 전후 학생이 쓴 내러티브 창의적 글들 Coh-matrix의 32개 측정치로 언어적 변화를 분석함.
 사전 사후 비교가 가능하고 동일 조건에서 시험을 치른 경우의 학생의 산출물을 분석 대상으로 삼아 최종적으로 실험반 14명, 비교반 16명의 사전 사후 산출물을 선택함.
 코메트릭스를 활용한 연구에서 공통적으로 많이 활용한 측정치를 본 연구의 분석자료인 내러티브 글쓰기 분석에 맞게 재배치함. 분석할 34 측정치는 선행 연구에서 제시함.

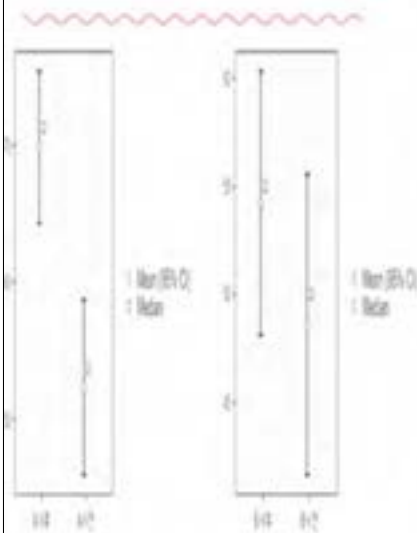
3.6. 자료분석 및 처리

Jamovi 프로그램을 사용하여 자료를 분석하고, 유의수준 5%에서 해석.

ICM기반 내러티브 창의적 글쓰기는 영어영재학습자의 언어 창의성을 향상시키기 위해 만들어진 프로그램으로 본 연구에서는 창의성 향상에 효과가 있을 것으로 가설을 세워 Flower in the Brick Wall 제목의 사진을 보여주고 사전 사후 내러티브 창의적 글쓰기를 실시했다. 실험집단과 통제집단의 사전 검사 결과가 t-test 검정 결과, 실험집단과 통제집단의 동질성이 확보되었으므로 ICM기반 내러티브 창의적 글쓰기 수업이 언어 창의성을 향상시켰는지 확인하기 위해 사후 검사 결과를 이용해 사전 사후 대응표본 t-test를 실시하였다.

4. 연구의 결과

창의성에 미치는 프로그램 효과



4.1. 창의성에 미치는 프로그램 효과

창의성 점수 기술통계량

사전-사후	Group	N	Mean	Median	SD	SE
사전	A	14	52.3	53.8	12.2	3.26
	B	16	48.6	50.0	14.2	3.55
사후	A	14	69.8	71.3	10.6	2.83
	B	16	54.2	55.0	12.5	3.12

창의성 점수의 사전-사후 대응표본 t-검정 결과

		t	df	p	Mean difference	SE difference	Cohen's d Effect Size
A사후	A사전	7.12	13.0	<.001	17.50	2.46	1.902
B사후	B사전	1.42	15.0	0.177	5.63	3.97	0.354

내러티브 글쓰기 창의성 요소별 기술통계량

			N	Mean	Median	SD	SE
이미지 image	A	사전	14	10.54	10.00	3.13	0.836
		사후	14	14.11	12.50	3.19	0.854
	B	사전	16	10.31	10.00	4.55	1.138
		사후	16	10.78	10.00	3.38	0.845
목소리 voice	A	사전	14	10.36	10.00	2.57	0.666
		사후	14	13.93	15.00	2.72	0.728
	B	사전	16	9.38	10.00	2.96	0.740
		사후	16	12.03	10.00	3.56	0.890
독창성 originality	A	사전	14	10.71	11.25	2.85	0.761
		사후	14	13.04	13.75	3.13	0.836
	B	사전	16	9.84	10.00	2.66	0.664
		사후	16	10.78	10.00	3.13	0.781
아이디어화 ideation	A	사전	14	8.93	7.50	2.34	0.626
		사후	14	12.86	12.50	2.16	0.578
	B	사전	16	7.66	7.50	1.93	0.482
		사후	16	10.16	10.00	2.81	0.702
몰입 및 명료 engagement & clarity & lack of distraction	A	사전	14	11.79	12.50	3.01	0.905
		사후	14	15.89	15.00	2.32	0.621
	B	사전	16	11.41	12.50	4.08	1.020
		사후	16	11.25	11.25	3.65	0.913

내러티브 글쓰기 창의성 요소별 대응표본 t-검정

요소	그룹	t	p	Mean difference	SE difference	Cohen's d Effect Size
이미지 image	A	3.680	0.003	3.571	0.971	0.9835
	B	0.387	0.704	0.469	1.212	0.0987
목소리 voice	A	6.276	<.001	3.571	0.569	1.6775
	B	2.868	0.012	2.656	0.926	0.7170
독창성 originality	A	2.414	0.031	2.321	0.962	0.6452
	B	0.878	0.394	0.938	1.067	0.2196
아이디어화 ideation	A	6.271	<.001	3.929	0.626	1.6780
	B	3.038	0.008	2.500	0.823	0.7596
몰입 및 명료 engagement & clarity & lack of distraction	A	6.097	<.001	4.107	0.674	1.6295
	B	-0.131	0.898	-0.156	1.196	-0.0327



4. 연구의 결과

내러티브 창의적 글쓰기 결과물인 코메트릭스 분석 결과- 담화적 특징

담화적 특징 기술 통계량

	Mean	95% Confidence Interval			
		Lower	Upper	Median	SD
LSAS51	0.185	0.1317	0.228	0.195	0.0927
LSAS51 (2)	0.174	0.1270	0.221	0.144	0.0813
LSAS5p	0.168	0.1172	0.218	0.168	0.0875
LSAS5p (2)	0.129	0.0823	0.165	0.108	0.0628
LSAGN	0.323	0.3027	0.343	0.320	0.0349
LSAGN (2)	0.334	0.3116	0.357	0.333	0.0395
CRFNO1	0.252	0.1617	0.342	0.257	0.1564
CRFNO1 (2)	0.221	0.1681	0.295	0.229	0.1097
CRFAO1	0.556	0.4617	0.650	0.521	0.1631
CRFAO1 (2)	0.512	0.4064	0.617	0.512	0.1824
CRFSO1	0.277	0.1790	0.375	0.265	0.1686
CRFSO1 (2)	0.275	0.2008	0.347	0.248	0.1252
CNCAR	101.807	85.2011	118.382	100.761	26.7079
CNCAR (2)	93.179	80.2812	106.078	95.314	22.3378
CNCaus	28.783	19.0438	38.543	25.502	16.8868
CNCaus (2)	25.308	19.0744	31.541	19.133	10.7958
CNClogic	43.802	34.3787	53.232	46.176	16.3329
CNClogic (2)	39.843	35.3968	44.290	40.513	7.8058
CNCADC	17.293	12.1628	22.423	16.592	8.8849
CNCADC (2)	17.628	13.6068	21.630	17.371	6.9302
CNCtemp	23.364	13.1050	33.624	20.998	17.7890
CNCtemp (2)	19.515	13.7582	25.270	18.441	9.9880
CNCadd	49.588	40.5793	58.558	50.362	15.5888
CNCadd (2)	47.803	37.0778	58.788	48.877	18.8006
PONARz	0.963	0.6252	1.301	0.965	0.5695
PONARz (2)	0.959	0.7446	1.174	0.984	0.3722

담화적 특징 대응표본 t 검정

		t	p	Mean difference	SE difference	Cohen's d (Effect Size)
LSAS51 (2)	LSAS51	-0.4127	0.687	-0.01129	0.0273	-0.11929
LSAS5p (2)	LSAS5p	-1.8244	0.091	-0.03914	0.0215	-0.48759
LSAGN (2)	LSAGN	0.8195	0.427	0.01157	0.0141	0.21901
CRFNO1 (2)	CRFNO1	-0.5709	0.578	-0.02057	0.0360	-0.15259
CRFAO1 (2)	CRFAO1	-0.6941	0.500	-0.04414	0.0638	-0.18550
CRFSO1 (2)	CRFSO1	-0.0391	0.968	-0.00179	0.0457	-0.01045
CNCAR (2)	CNCAR	-1.0419	0.318	-8.62793	8.2811	-0.27846
CNCaus (2)	CNCaus	-1.0643	0.307	-3.48571	3.2753	-0.26443
CNCADC (2)	CNCADC	0.1109	0.913	0.32536	3.0236	0.02964
CNClogic (2)	CNClogic	-1.0623	0.307	-3.95807	3.7261	-0.28990
CNCtemp (2)	CNCtemp	-0.6751	0.511	-3.84993	5.7025	-0.18044
CNCadd (2)	CNCadd	-0.3146	0.758	-1.63550	5.1983	-0.08409
PONARz (2)	PONARz	-0.0208	0.984	-0.00386	0.1885	-0.00556

5. 결론



첫째, ICM 기반 내러티브 창의적 글쓰기 수업은 창의성을 신장시키는 데에 효과가 있다.

실험집단-전체 창의성 점수, 이미지, 목소리, 독창성, 아이디어화, 몰입 및 영료 모든 하위 영역에서 사전과 비교해 유의한 차이가 나게 함(p<0.05)
비교집단-목소리, 아이디어화에서 사전에 비해 유의한 차로 향상(p<0.05)

둘째, 코메트릭스로 실험집단의 글쓰기에 나타난 언어적 변화

전체 어휘수 증가
읽기 가독성 지표와 통사적 복잡성-난이도 높아짐
TTR-어휘 다양성 줄어듦 (전체 어휘수 증가와 함께 고려하면 사전에 비해 실제적인 어휘 다양성이 줄었다고 말할 수는 없음)
음절성- 큰 변화없음, 내러티브 큰 변화 없음

내러티브 창의적 글쓰기를 수행함으로써 영어영재와 교사의 주요 관심사인 언어의 표현적 특성 이해와 활용에 관심을 갖게 할 뿐만 아니라, 아이디어 생성과 창의성에도 관심을 가지게 함으로써 좀 더 높은 사고 과정을 유도함
표현과 몰입의 독창성, 캐릭터와 배경에 대한 지필하고 정교한 묘사, 아이디어 정교화 등의 언어 창의성과 비판적 사고력 등의 고등 사고력을 향상

연구의 제한점 및 제언

연구 데이터 수집 시 학생들이 코로나 상황 및 학생 개인 여건으로 인해 연구 참여자에 비해 데이터 확보 수가 줄어듦.
연구 참여자 수가 작고, 선발된 집단에 프로그램 적용에 일반화 어려움.
비교집단 실험집단의 전체 프로그램 진행 시수는 같으나 1주일에 1번씩 적용되지 못하고, 주어진 프로그램 시간대로 진행됨.
내적 만족감과 자긍심, 몰입감 등의 창의적 성향 변화 확인하는 연구와 프로그램 후 창의적 행동에 대한 자기 인식 등 비인지적 창의성에 대한 후속 연구 필요

질문 및 토론



한국교원대 영어교육과
이주영 박사과정생
1442sun@naver.com

참고 문헌

김승희, & 김창섭. (2008). 초등학교 언어 장애 아동과 일반 아동의 언어지능차상 발달에 관한 연구. *초등교육연구*, 36(1), 403-421.

김지숙. *창의적성 측정지표 및 도구 개발 연구*. 서울: 한국교육개발원, 2012.7

김민철, & 안지윤. (2012). Coh-Metrix를 활용한 초-중등 영어교과서 연계성 분석. *영어교육*, 47(2), 319-342.

김지수, 천우기. (2020). Coh-Metrix 코호그램을 활용한 2015년 개정 교육과정 중학교 1학년 영어 교과서 비교 분석연구, 26, 8-9-121.

박정근, and 이은수. (2021). Coh-Metrix를 활용한 고등학교 영어 교과서의 대학수학능력시험 읽기 지문 난이도 비교 연구. *영어영문* 34(2), 473, 147-170.

박재현, & 이경희. (2016). K-M 기반 영어 언어정보로그그램이 고등학생의 영어 쓰기 능력과 읽기 성취에 미치는 영향. *영어교육*, 51(2), 29-52.

이주영, & 김정열. (2021). 영어영문정보로그그램 교과과정 요구조사 분석. *한국영어영문학*, 39(4), 213-242.

전승철. (2020). 토론 기반 영어 읽기 및 쓰기 학습 활동이 고등학교 학습자의 창의성, 영어 쓰기 및 읽기 자기평가에 미치는 영향. *Foreign Language Education*, 27(2), 47-65.

Carry, M. D., Davidow, S., & Williams, P. (2022). Re-imagining narrative writing and assessment: a peer-APPLAN (collaborative) rubric for creative writing. *The Australian Journal of Language and Literacy*, 3-26.

Geneser, A. C., McNamara, D. S., Loewen, M. M., & Cai, Z. (2004). Coh-Metrix Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, and Computers*. <https://doi.org/10.3758/BF03205544>

Cropley, A. J. (2011). Definitions of creativity. In M. A. Runco & S. R. Pritzker (Eds.), *Encyclopedia of Creativity* (pp. 511-524). San Diego, CA: Academic Press.

Gen, L., Onisley, S. A., & McNamara, D. S. (2013). Predicting human judgments of essay quality in both integrated and independent second language writing contexts: A comparison study. *Australian Writing*, 19(3), 216-236. <https://doi.org/10.1016/j.aew.2011.05.002>

Artiles, T., & Bowen, J. (2017). Measuring creative capacity in gifted students: Comparing teacher ratings and student products. *Gifted Child Quarterly*, 61(4), 290-299.

Leff, S., & Genl, M. (2021). Automated scoring of junior and senior high essays using Coh-Metrix features: Implications for large-scale language testing. *Language Testing*, 38(1), 62-85.

Huckler, J. A., Bighetti, R., & Dow, G. (2004). Why isn't creativity more important to educational psychologists? Potential, pitfalls, and future directions in creativity research. *Educational Psychologist*, 33, 63-95. [doi:10.1207/s15326985ep3302_1](https://doi.org/10.1207/s15326985ep3302_1)

Selby, C. C., Shaw, E. J., & Houtz, J. C. (2005). The creative personality. *Gifted Child Quarterly*, 49, 300-314. [doi:10.1177/17006902052692404](https://doi.org/10.1177/17006902052692404)

Zedlitz, C. H., Mills, C., & Schooker, J. W. (2015). Beyond subjective judgments: Predicting evaluations of creative writing from computational linguistic features. *Behavior Research Methods*. <https://doi.org/10.3758/s13428-015-1213-1>

한국언어과학회 2023 겨울학술대회
"언어과학의 새로운 미래"
2023년 2월 16일(목)

3분과(코퍼스/응용)



경청해 주셔서 감사합니다.

한국교원대학교 이주영

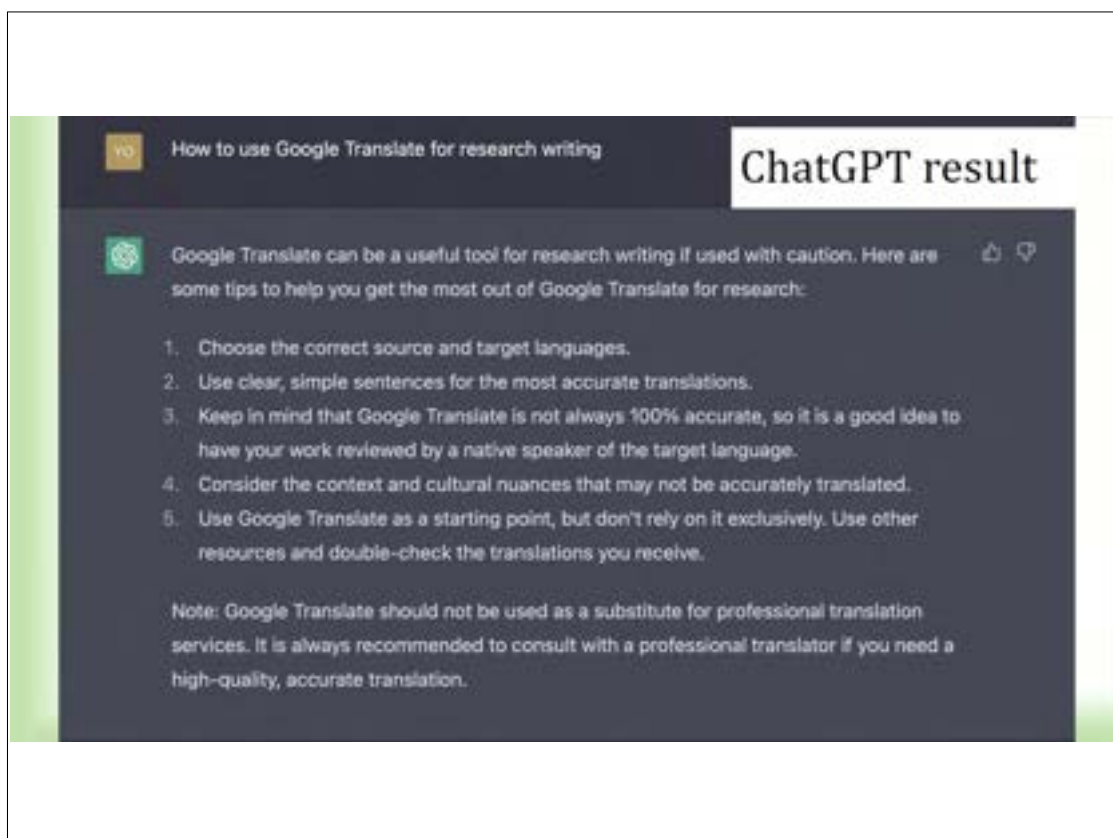
제4분과

응용/언어교육

Instructing how to use Google Translate effectively for research writing

Younghyon Heo

University of Aizu



The screenshot shows a ChatGPT interface with a dark background. At the top left, there is a small orange icon with 'v3' and the text 'How to use Google Translate for research writing'. To the right, the text 'ChatGPT result' is displayed in white. Below this, there is a green icon representing Google Translate and a paragraph of text: 'Google Translate can be a useful tool for research writing if used with caution. Here are some tips to help you get the most out of Google Translate for research:'. To the right of this paragraph are two small icons: a share icon and a heart icon. Below the paragraph is a numbered list of five tips. At the bottom of the screenshot, there is a 'Note' section with a paragraph of text.

How to use Google Translate for research writing

ChatGPT result

Google Translate can be a useful tool for research writing if used with caution. Here are some tips to help you get the most out of Google Translate for research:

1. Choose the correct source and target languages.
2. Use clear, simple sentences for the most accurate translations.
3. Keep in mind that Google Translate is not always 100% accurate, so it is a good idea to have your work reviewed by a native speaker of the target language.
4. Consider the context and cultural nuances that may not be accurately translated.
5. Use Google Translate as a starting point, but don't rely on it exclusively. Use other resources and double-check the translations you receive.

Note: Google Translate should not be used as a substitute for professional translation services. It is always recommended to consult with a professional translator if you need a high-quality, accurate translation.

English writing is never like before...



- When it comes to English writing, many non-native speakers of English do not write something from scratch (at school or at their workplaces).
- The ability to selectively use the machine-generated English sentences and modify the output is more important (writer → editor).

Me as a Thesis Writing instructor

- I have been teaching Thesis Writing for about 10 years in Japan.
- Undergrad students at University of Aizu write a graduate thesis in English.
- Most of them are **computer science majors**, with the English proficiency level between **low-intermediate and intermediate**.
- Compared to 10 years ago, it is becoming more and more difficult to spot machine-translated student drafts.

What has been happening in the past a few years?

- Since becoming widely available to the public by Google in 2006, machine translation has constantly evolved along with increasing computation power to meet the need for tools dealing with large data sets.
- Particularly after the transition from Statistical Machine Translation (SMT) to **Neural Machine Translation (NMT)** in 2016, translated texts between the world's major languages are of far better quality than ever before.
- Since 2016, there have been more and more English translation/editing services based on neural networks technology launched, and machine translation is hard to detect.
- Many language learners rely heavily on machine translation for their writing.

So what can we teach in a (English) writing class?

- Insist language learners refrain from using machine translation?
→ impossible
- (One of the) course objectives:
To help English learners polish their writing by selectively **choosing** sentences from what is generated by machine-translation and further **modifying** those sentences following the **style/purpose** of writing and also for the **clear delivery** of intended ideas.

How about Thesis Writing Class?

- We can assume that most of the students are using machine translation and teach them to polish their writing following the stylistic guidelines of research writing.
- Writers as "designers" in the past, but now writers as "stylists". We aim to teach them to become a writer (researcher) who understands:
 - 1) the tradition of research writing,
 - 2) the style accepted for research writing, and
 - 3) the critical use of machine-translated sentences.

Goals

- 1) In this preliminary report, we analyze:
 - the human-written and machine-translated academic writing and compare them, and
 - the results of the survey conducted to EFL learners on their experience with Google translate for research writing.
- 2) With the patterns noticed in 1) we create a teaching material instructing students enrolled in a English Thesis Writing class in Japan **how to use Google Translate effectively for their research writing.**

* This study was conducted as part of the research project supported by Grants-in-Aid for Scientific Research.

***Analysis and comparison of
human-written and machine-translated academic writing***

**Comparing human writing and machine writing
: Materials 1**

- Two collections of academic abstracts from the National Institute of Informatics (NII) in Japan.
 - Test Collection 1: 30,000 academic abstracts written in Japanese and English collected from 65 academic societies in Japan between **1988 and 1997**
 - Test Collection 2: 400,000 Japanese and 130,000 English documents submitted between **1997 and 1999**
- It was clear that the English abstracts were not themselves machine-translated since these abstracts were **published before Google Translate (SMT-based) began its service in 2006.**

- Two types of documents were prepared:
 - 1) Japanese abstracts translated by Google Translate, and
 - 2) English abstracts written by Japanese authors.
- In this study, a corpus of 385,184 machine-translated and 193,922 human-written English sentences was used.

Comparing human writing and machine writing : Materials 2

- 14 student writings written in Japanese (graduation thesis) were selected for the analysis.
- These human-written theses were Google-translated and analyzed for their unique machine translation features and for their comparison to the original Japanese writing to evaluate the quality of translation.

Analysis of machine-translated (Google) research writing: Method

- We reviewed the machine-translated versions of the research abstracts (data provided by NII) and student writings (graduation theses).

Analysis of machine-translated (Google) research writing: Results

- We analyzed the machine-translated versions and found that the following typically lead to inaccurate/unnatural/bad style translation:
 - lengthy sentences
 - idiomatic expressions
 - fixed expressions in Japanese formal writing

Lengthy sentence: Example

このようにハードウェアを使用して行うことは、今後ドローンの研究を行うのに、コストがかかること、安全面を考慮しなければならないこと、環境によって左右されるなどいくつかの壁に阻まれることが予想されます。

이처럼 하드웨어를 이용한 연구는 향후 드론 연구를 진행함에 있어 비용이 많이 든다는 점, 안전성을 고려해야 한다는 점, 환경에 따라 영향을 받는다는 점 등 여러 가지 장벽에 부딪힐 것으로 예상된다.

Using hardware in this way is hindered by several obstacles, such as the high cost, the need to consider safety, and the fact that it is affected by the environment, in order to conduct drone research in the future, is expected.

Idiomatic expression: Example

AI技術の発展のおかげで最近は人間とパソコンの距離が近くなっています。

AI 기술의 발전으로 최근 인간과 컴퓨터의 거리가 점점 가까워지고 있다.

Thanks to the development of AI technology, the distance between humans and computers is getting closer these days.

Set phrases in Japanese formal writing: Example

この設計手法をマクロセルを構成する101個のリーフセルと65個の標準セルに適用した。その結果、標準セルは人手と同等、リーフセルについても人手と比較して、平均10%程度の面積増に抑えられた。

이 설계 기법을 매크로 셀을 구성하는 101개의 리프 셀과 65개의 표준 셀에 적용했다. 그 결과, 표준 셀은 인력과 동등하게, 리프 셀도 인력과 비교해 평균 10% 정도의 면적 증가를 억제할 수 있었다.

This design method was applied to 101 leaf cells and 65 standard cells that make up a macrocell. As a result, the area of the standard cell was the same as that of manual processing, and the area of the leaf cell was suppressed to an average increase of about 10% compared to that of manual processing.

Comparison of human-written texts and machine-translated (Google) texts: Method

- N-gram analysis
Top 50 of unigram, bigram, trigram and quadragram in human-written vs. machine-translated texts was analyzed.
- Top 50 most common words/phrases were analyzed with two different settings: (+punctuation) and (-punctuation).

n-gram

- Human-written (+punc)
- Machine-written (+punc)
- Human-written (-punc)
- Machine-written (-punc)

Comparison of human-written texts and machine-translated (Google) texts: Results

- In the next 4 slides, top 10 of each type are provided.

Unigram

	Human (+punc)	Machin (+punc)	Human (-punc)	Machine (-punc)
1	The	the	The	the
2	This	,	This	and
3	are	and	are	of
4	paper	of	paper	a
5	for	a	for	to
6	A	to	A	it
7	which	it	which	In
8	proposed	In	proposed	this
9	systems	this	systems	method
10	new	method	new	is

Bigram

	Human (+punc)	Machin (+punc)	Human (-punc)	Machine (-punc)
1	Human (+punc)	Machin (+punc)	Human (-punc)	Machine (-punc)
2	. The	of the	This paper	of the
3	. This	. In	is proposed	and the
4	This paper	, and	for the	it is
5	. A	, we	paper describes	in the
6	is proposed	, it	has been	In this
7	for the	and the	a new	to the
8	paper describes	it is	paper presents	As a
9	. It	, the	paper proposes	as a
10	, which	in the	for a	on the
11	has been	In this	the proposed	a result

Trigram

	A	B	C	D
1	Human (+punc)	Machin (+punc)	Human (-punc)	Machine (-punc)
2	. This paper	. In this	This paper describes	As a result
3	This paper describes	, it is	This paper presents	In order to
4	. It is	. As a	This paper proposes	In this paper
5	This paper presents	, and the	is shown that	is possible to
6	is proposed .	As a result	has been developed	it is possible
7	This paper proposes	paper , we	paper describes the	a method of
8	. A new	. Therefore ,	paper proposes a	is necessary to
9	is shown that	this paper ,	paper describes a	it is necessary
10	has been developed	a result ,	paper presents a	In this report
11	paper describes the	. In the	It is shown	the time of

Quadragram

	Human (+punc)	Machin (+punc)	Human (-punc)	Machine (-punc)
1	Human (+punc)	Machin (+punc)	Human (-punc)	Machine (-punc)
2	. This paper describes	. As a result	This paper describes a	it is possible to
3	. This paper presents	this paper , we	This paper presents a	it is necessary to
4	. This paper proposes	As a result ,	This paper describes the	at the time of
5	This paper describes a	In this paper ,	This paper proposes a	it was found that
6	This paper presents a	. In order to	It is shown that	As a result of
7	This paper describes the	. In this paper	In this paper we	was found that the
8	. It is shown	. In addition ,	is shown that the	it was confirmed that
9	This paper proposes a	it is possible to	This paper deals with	propose a method to
10	It is shown that	it is necessary to	is based on the	from the viewpoint of
11	this paper , a	In this report ,	it is shown that	we propose a method

• Set phrase in Japanese: "本論文では、" 본 논문에서는,

	Human (+punc)	Machin (+punc)	Human (-punc)	Machine (-punc)
1	Human (+punc)	Machin (+punc)	Human (-punc)	Machine (-punc)
2	. This paper describes	. As a result	This paper describes a	it is possible to
3	. This paper presents	this paper , we	This paper presents a	it is necessary to
4	. This paper proposes	As a result ,	This paper describes the	at the time of
5	This paper describes a	In this paper ,	This paper proposes a	it was found that
6	This paper presents a	. In order to	It is shown that	As a result of
7	This paper describes the	. In this paper	In this paper we	was found that the

: It is obvious that the direct translation of the set Japanese phrase in formal writing such as "本論文では" appears more frequently in machine writing (#4) while using "paper" as a subject is far more frequently found in human writing (#1~#6, and #1~#4).

: "This paper..." phrase was not found in the top 50 of the machine writing list.

Student survey

Student survey on their use of machine translation : Method

- Students were asked whether they used machine translation for their thesis writing. For those who answered yes, they were asked to write freely about how they used it.
- Out of 15 undergrad students enrolled in my thesis writing class for this academic year (AY 2022), 14 students answered that they used machine translation while they worked on their graduation theses.
- The student survey response shows how they have been using it for their research writing.

Student use of machine translation: examples

- 일본어 원고를 번역한 후 부자연스러운 문장을 직접 수정해서 사용하고 있다.
- 일본어 표현과 번역한 영어 표현이 종종 다른 의미로 받아들여질 때가 있다. 그럴 때는 그 이상한 단어의 유사어를 찾는다. 또 그 단어의 실제 예문을 찾아 의도한 의미와 같은지 확인한다.
- 영어 문법에 따라 서술하면 생각한 대로 번역 결과가 나온다. 예를 들어, 주어, 술어, 목적어를 명확히 해서 쓰면 좋다.
- 너무 긴 문장을 번역하지 않는다. 간결한 문장을 번역하도록 한다.
- 번역 결과를 직접 읽어본다. 뉘앙스가 다르면 다른 단어를 찾아 사용한다. 또한 같은 단어를 많이 사용하지 않는다.

- 어떤 동사, 명사를 강조하고 싶은지에 따라 주어와 문형을 바꾼다.
- 가능한 한 스스로 영어 번역을 한다. 의미가 통하는지 구글 번역기로 확인한다(영어-일본어).
- 번역 툴 하나에 너무 의존하는 것은 위험하다고 느꼈다.
- 번역으로 나온 문장에서 단어를 하나하나 찾아보고 대체해 나가면 좋은 느낌이 든다.
- 연구에 사용하는 고유명사 등은 번역기를 쓰면 잘 표현되지 않는 경우가 많아서 고유명사는 일단 A 등으로 표현해 번역하기도 했습니다.
- 일본어에서 영어, 영어에서 일본어 번역을 두세번 정도 반복하면 더 좋은 번역 결과가 나오기 시작했다.
- 단어를 찾는 데는 효과가 있었지만, 문장을 넣으면 캐주얼한 표현(논문엔 적합하지 않은 단어)이 많은 것 같았다.

Instructing how to use Google Translate effectively for their research writing

Avoid long sentences

このようにハードウェアを使用して行うことは、今後ドローンの研究を行うのに、コストがかかること、安全面を考慮しなければならないこと、環境によって左右されるなどいくつかの壁に阻まれることが予想されます。

이처럼 하드웨어를 이용한 연구는 향후 드론 연구를 진행함에 있어 비용이 많이 든다는 점, 안전성을 고려해야 한다는 점, 환경에 따라 영향을 받는다는 점 등 여러 가지 장벽에 부딪힐 것으로 예상된다.

→ 이렇게 하드웨어를 이용해 연구를 진행하다 보면 몇 가지 벽에 부딪히게 될 것으로 예상된다. 구체적으로 향후 드론 연구를 진행하는데 비용이 많이 든다는 점, 안전성을 고려해야 한다는 점, 환경에 따라 달라질 수 있다는 점 등이 있다.

When using hardware in this way, it is expected that there will be some obstacles. Specifically, it will be costly to conduct research on drones in the future, safety must be considered, and it will be affected by the environment.

Avoid idiomatic expressions

AI技術の発展のおかげで最近は人間とパソコンの距離が近くなっています。

AI 기술의 발전으로 최근 인간과 컴퓨터의 거리가 점점 가까워지고 있다.

→ AI 기술의 발전으로 최근 컴퓨터의 사용이 인간 생활의 여러 분야에 더욱 관여하게 되었다.

With the development of AI technology, the use of computers has recently become more involved in many areas of human life.

Modify set phrases in Japanese into set phrases in English-1

この設計手法をマクロセルを構成する101個のリーフセルと65個の標準セルに適用した。その結果、標準セルは人手と同等、リーフセルについても人手に比較して、平均10%程度の面積増に抑えられた。

이 설계 기법을 매크로 셀을 구성하는 101개의 리프 셀과 65개의 표준 셀에 적용했다. 그 결과, 표준 셀은 인력과 동등하게, 리프 셀도 인력과 비교해 평균 10% 정도의 면적 증가를 억제할 수 있었다.

This design method was applied to 101 leaf cells and 65 standard cells that make up a macrocell. As a result, the area of the standard cell was the same as that of manual processing, and the area of the leaf cell was suppressed to an average increase of about 10% compared to that of manual processing.

This design method was applied to 101 leaf cells and 65 standard cells that make up a macrocell. **The result shows that** the area of the standard cell was the same as that of manual processing, and the area of the leaf cell was suppressed to an average increase of about 10% compared to that of manual processing.

Modify set phrases in Japanese into set phrases in English-2

外部記憶との入出力時間削減を図ることを試みた。その結果、計算機の使用時間で、約15倍の高速化を得ることができた。

외부 메모리와의 입출력 시간 단축을 시도했다. 그 결과 계산기 사용 시간에서 약 1.5배의 빠른 속도를 얻을 수 있었다.

An attempt was made to reduce input/output time with external memory. As a result, we were able to obtain about 1.5 times speedup in computer usage time.



An attempt to reduce input/output time with external memory **resulted in** 1.5 times improved speed in computer usage time.

OR

An attempt to reduce input/output time with external memory **led to the** 1.5 times improved speed in computer usage time.

Modify set phrases in Japanese into set phrases in English-3

本論文では、VLSIのフロアプラン設計における配線経路指定問題についての検討をおこなう。

본 논문에서는, VLSI의 플로어플랜 설계에서 배선 경로 지정 문제에 대해 고찰한다.

In this paper, we examine the wiring routing problem in VLSI floorplan design.



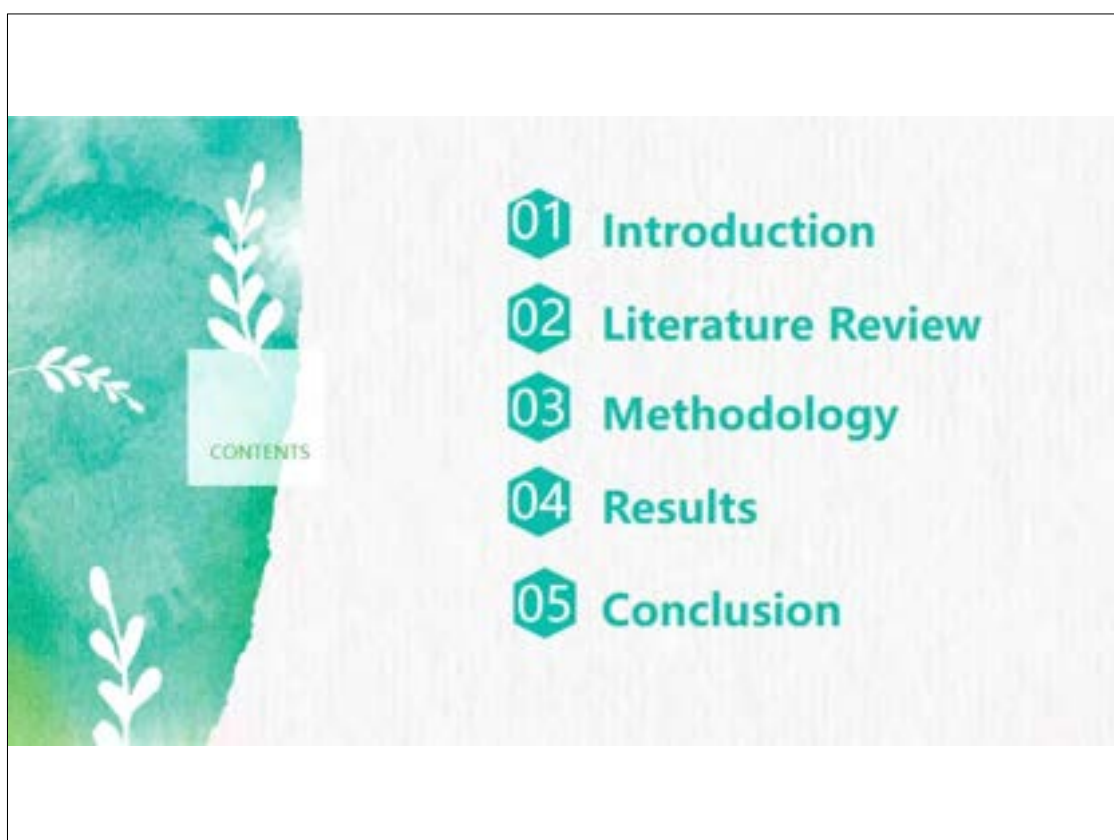
This paper examines the wiring routing problem in VLSI floorplan design.

- When students use the same expressions such as "In this paper," repeatedly throughout the paper, they can be advised to use an alternative sentence pattern that starts with the subject "This paper".

Summary and Future Works

- We reported our preliminary results and findings of analyzing and comparing the human-written and machine-translated academic texts for our development of a teaching material guiding students how to use Google Translate effectively for their research writing.
- In this project, academic writing data were predominantly research abstracts. N-gram analysis (Top 50) of the "results" part, where set phrases are frequently used, will generate another interesting set of commonly used expressions in human-written and machine-written writings.

Thank you for listening!



—01—
Introduction

1. Research Needs

Writing skill is an important measure of the language competence and proficiency of foreign language learners. By studying student writing samples, researchers can comprehensively assess the competence of foreign language learners' competence and discover pervasive issues in different levels. Therefore, this study focuses on the analysis of Chinese EFL college students' writing and investigates the errors in their English writings.

2. Research Purpose

With the participation of 30 EFL freshmen of a normal university, China, this research aims to find out the most common errors that EFL students in a specific learning environment make in their English writing and analyze them to discover the main reasons that cause the errors.

3. Research Questions

- (1) *What common errors do Chinese EFL students make in their writing and what is the frequency of each different type of error?*
- (2) *What may cause these errors to occur in the writing of the EFL students?*

—02—
Literature Review

1. Definition of Errors

Brown (2002, 2014) claimed that during L1 and L2 acquisition, errors occur when the learner's assumptions about the target language are inconsistent with the target language system and distinguished the error from the mistake.

-A mistake is a performance error and thus can be self-corrected, and just like slips of the tongue, casual bad grammar and other vocabulary or punctuation lapses are instances of mistakes.

-Errors are a significant deviation from the adult grammar of native speakers and show the language skills that a learner still hasn't acquired yet.

Jame (1998) also distinguished errors and mistakes by checking whether they can be self-corrected or not and emphasized that error analysis will help researchers and teachers understand lots of key points in the process of second language acquisition.

2. Classification of Errors

-Burt & Kiparsky (1972): global errors and local errors

-Dulay & Burt (1974): developmental errors, interference errors and unique errors

-Corder: overt errors and covert errors (1971), systematic errors and unsystematic errors (1967), intralingual errors and interlingual errors (1973)

-Richards (1971): intralingual errors, interference errors, and developmental errors

-Lennon (1991): errors related to deleting, replacing, providing, and reordering sentences in the sentence level

-Brown (2014): addition errors, omission errors, substitution errors, and arrangement errors

-After all, the classification of errors can be summarized into interlingual errors, intralingual errors, and developmental errors.

In this study, I combined previous studies of scholars and divided errors into a substance level, a text level, and a discourse level according to Zhang (2015) and later added detailed classification following Tong (2002). The classified errors here are as follows.

	Types	Examples
Substance errors	Spelling	There will be a severe compitition(competition).
	Punctuation	I think it is useful for students. (.)
	Capitalization	Now, In the campus, there is a phenomenon. (in)
Text errors	Lexical errors	It is very considerable for him to help us. (considerate)
	Grammatical errors	The advantages is that..... (are)
Discourse errors	Cohesion	Some students like the humorous teachers, and the others like the strict ones. (while)

—03—
Methodology

·Research Questions

- (1) What common errors do Chinese EFL students make in their writing and what is the frequency of each different type of error?
- (2) What may cause these errors to occur in the writing of the EFL students?

1. Research Subjects

This paper takes 30 EFL students from Guangdong Polytechnic Normal University, China as research subjects. They are EFL freshmen students of business English, and their ages are from 18 years old to 20 years old. Among them are five male students and 25 female students.

- The subjects all have very similar backgrounds in learning English:
- Their hometowns are different regions in Guangdong Province.
 - Their first language is Chinese.
 - None of them have parents who are native speakers of English or who can speak English fluently.
 - None have had any experience of living abroad before.

2. Research Processes

(1) The subjects were told to complete the writing in a limited time as part of the final exam for the writing course.

(2) Ask the subjects to write a book report or a movie review for one hour, with a 200-300 word limit.

(3) The students were not allowed to use a dictionary or phone to find unfamiliar words or get direct help from their supervisor who monitored them until they finished.

(4) The scoring standard was followed Zhang's (2013).

(5) The total score of the test item is 25 points.

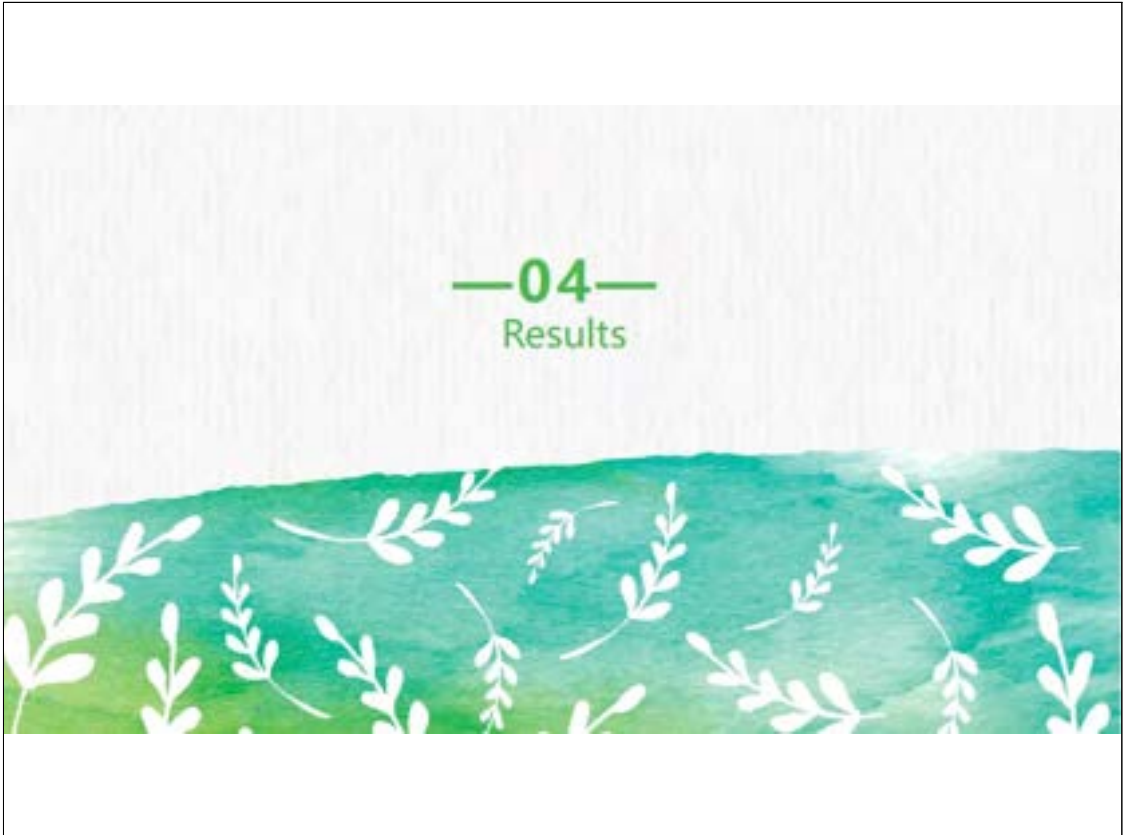
(6) The subdivisions of the students were made: the low score level from 0 to 10 points; the medium level from 11 to 15 points; and the high score level from 16 to 25 points.

(7) Two experienced teachers were invited to examined the collected data to secure the validity of identifying errors and counted the number of all the errors, ignoring a few errors made due to carelessness.

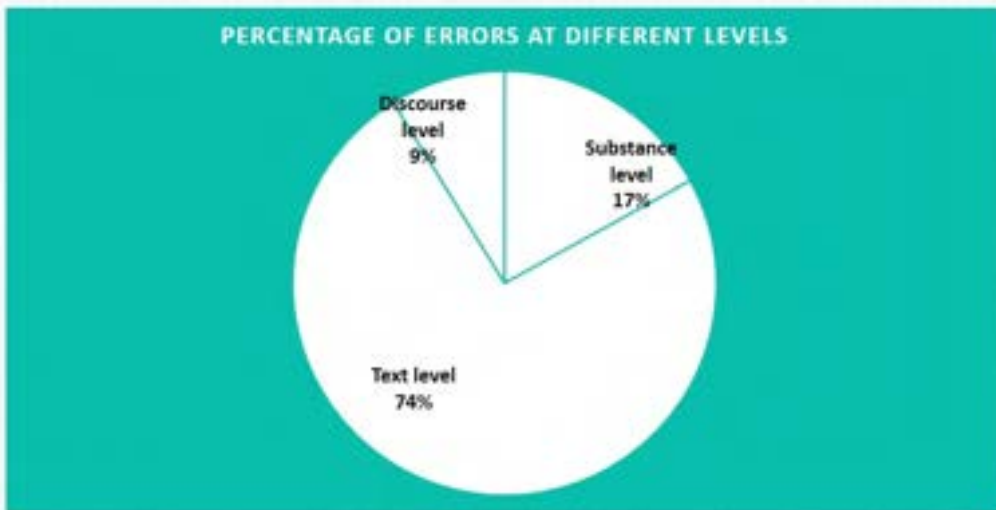
(8) I divide all the errors in the writing materials into lexical errors, grammatical errors, and discourse errors.

(9) According to the standard for evaluation, two teachers and I gave each English writing sample a final score by calculating the average amount.

—04—
Results



PERCENTAGE OF ERRORS AT DIFFERENT LEVELS




Types of Common Errors			Number	Percentage (%)
Substance Level		Spelling	23	6.97
		Punctuation	22	6.67
		Capitalization	11	3.33
Text Level	Lexical error	Parts of speech	19	5.76
		Word meaning	25	7.58
		Improper collocation	40	12.12
	Grammar Error	Lack of components	73	22.12
		Inconsistency of subject-verb	31	9.39
		Tense and voice	57	17.27
Discourse Level		Paragraph structure	13	3.94
		Reference	13	3.94
		Connection	3	0.91


-Many researchers have studied the reason that cause errors to happen in the writing of EFL students, I combine the results of previous studies and elucidate the causes for these common errors.

-The three types of causes of the errors are interlingual causes, intralingual causes, and other causes.

◆ Errors caused by interlingual causes are due to the influence of the mother tongue. The differences between Chinese and English, as well as the characteristics, habits, and rules of the two languages, have a profound influence on students' English learning and writing. In the process of writing, Chinese students tend to ignore the differences between Chinese and English thinking and those between the two languages and apply the grammatical rules and expressions of their mother tongue to English writing.



◆ Errors from intralingual causes refer to the confusion of usage by students due to the complexity and inconsistency of the English language itself. Intralingual factors can be divided into two kinds of conditions: overgeneralization and incomplete understanding of the rules. Overgeneralization is a kind of error in which students use part of the rules to generalize the whole part. In addition, when foreign language learners do not fully and comprehensively use the rules of foreign language, errors will also occur, such as inconsistencies in tenses.



◆ In addition to interlingual and intralingual causes, there are some other causes that can lead to errors. In learning English, Chinese students tend to make different errors in English writing due to different ways of thinking, cultural backgrounds, and the degree of attention on English writing, including teachers' teaching methods and students' learning styles.

—05—
Conclusion

1. Findings

This research tried to focus on the common errors that Chinese EFL students have in their writing. 30 writing pieces from participants from Guangdong Polytechnic Normal University were analyzed by two experienced professors and the researcher. The result showed that errors made by the EFL students are mainly at three levels: substance level, text level, and discourse level.

- At substance level, the types of errors are mainly errors in spelling, punctuation, and capitalization.
- At text level, errors are divided into lexical errors and grammar errors: Lexical errors include mixed use of parts of speech, misuse of word meaning, and improper collocation.
- As for the discourse level, it includes errors of paragraph structure, errors of reference, and errors of connection.
- Among all the errors, grammar errors at text level are most common errors.

- There were three kinds of causes that might make EFL students to make errors: interlingual causes, intralingual causes, and other causes. Interlingual causes include lexical errors, sentence errors and discourse errors. Intralingual causes include overgeneralization and master and application of rules. There are also some other causes that can be forwarded from a teacher's perspective and the student's perspective.

2. Pedagogical Implications for EFL Teaching

- Teachers should improve their teaching methods clearly, regard the teaching of writing as the cultivation of students' language habits, and pay more attention to the importance of teaching writing.
- They should pay attention to the basic English knowledge for students. Especially, for some students weak at learning grammar knowledge, the teacher should give timely and appropriate guidance.
- Students should enhance their extensive reading ability in the future. Extensive reading can help them understand the structure of the text and grasp some deeper understanding of writing. Also, they

- Students should enhance their extensive reading ability in the future. Extensive reading can help them understand the structure of the text and grasp some deeper understanding of writing. Also, they can get some fixed expressions through extensive reading.

3. Limitations and Suggestions

- The number of writing materials for the analysis was 30 pieces, and it was not sufficient. Thus, the result of the research cannot represent all Chinese college students.
- I analyzed the writing materials with the help of experienced teachers, but their judgement of errors cannot be 100% guaranteed.



한국언어 과학회 2023 겨울 학술대회

CAF에 나타난 한국 대학생의 말하기 특징

신순란 부산 대학교 영어영문학과

I. 서론

영어 말하기 능력의 중요성 :

- √ 인터넷을 통한 세계의 글로벌화
- √ 실시간 정보의 수집, 활용
- √ 국가 경쟁력

- 한국 대학생들의 말하기 유창성은 TOEIC, TOEFL, IELTS 등의 시험성적과 연관되어 측정되었을 뿐 **CAF를 통한 다층적인 면에서의 분석은 드문 실정.**

√ 언어의 복잡성(Complexity), 정확성(Accuracy), 유창성(Fluency)(이하 CAF)은 제 2언어 학습자의 쓰기, 말하기 능력을 측정할 수 있는 유용한 도구(Skehan, 1998; Ellis, 2003; Housen & Kuiken, 2009; Craven, 2017).

- Complexity: 정교한 언어 사용 능력
- Accuracy: 문법오류 빈도수
- Fluency : 발화

L2 학습자의 언어 능숙도를 다각적인 측면에서 측정 가능.

√ CAF의 상호관련성

1. 교환가설(Trade off Hypothesis) : Skehan(2009) 인간의 주의 집중력 시간은 제한되어 있기에 학습자의 복잡성, 정확성, 유창성이 상호 경쟁 관계
2. 인지 가설(Cognition Hypothesis) : Robinson(2005) 복잡성, 정확성, 유창성은 상호 보완관계

√ CAF의 교환가설과 인지 가설에 대한 상이한 연구결과(Crookes, 1989; Ortega, 1999; Kormos & Denes, 2004; 오미영, 2008)로 L2 학습자의 언어 수행 능력을 CAF를 통해 살펴보는 연구가 더 많이 수행될 필요가 있음.

◆ 본 논문의 연구 목적

√ CAF로 측정된 말하기 능력을 바탕으로 CAF의 두 가지 관점인 상호 관련성과 대립성을 검증

√ 한국 대학생의 영어 말하기 특징을 CAF를 통해 측정하고자 함.

→ CAF는 복잡성, 정확성, 유창성이 균형을 이룰때 L2 능숙도에 도달할 수 있기에(Fortkamp, 2000; Bygate, 2001) 종합적인 측면에서 말하기 능력을 살펴볼 수 있음.

√ 영어 말하기 향상 방법을 제시하여 교육적인 면에 기여하고자 함.

Ⅱ. 선행 연구

1. CAF의 정의 및 상호 관련성



√ L2 발달 측정을 위해 1980년대 부터 도입된 이후 L2학습자의 언어 수행력을 종합적으로 검토할 수 있기에 90년대 이후 광범위하게 활용됨.

▶ 정확성(Accuracy)

√ Lennon(1990): 무오류를 발화하는 능력

√ Wolfe-Quintero et al.(1998): 목표언어의 규칙을 따르는 것

√ Skehan, Foster(1999): 말하기 등의 수행에서 오류를 피하는 능력

▶ 정확성(Accuracy) 측정 도구

√ Yuan & Ellis, 2003 : 전체 절(clause)에서 무 오류절(error-free clause)을 구하는 방법

- 1) 통사적, 형태적, 어휘적 선택과 관련된 모든 오류
- 2) 시제, 상, 조동사, 주어-동사 일치의 측면에서 오류가 없는 동사의 비율

√ Pallotti(2009): 정확성에 원활한 의사소통이라는 개념 추가. 의사소통을 손상시키는 정도에 따라 오류에 가중치를 부여

▶ 유창성 (Fluency)

√ Lennon(2000) : 발화 속도 및 시간적 제약하에서 화자의 의도를 빠르고 효율적이며, 정확하고, 명확하게 전달하는 능력

√ Segalowitz(2000) : L2를 모국어 만큼 빠른 속도로 망설임이나 휴지(pause), 오류없이 말할 수 있는 능력. 발화속도, 발화길이, 휴지, 머뭇거림 등으로 유창성을 정의.

▶ 유창성 측정도구

√ Lennon(2000) : 발화의 속도+ 반복이나 잘못된 시작과 같은 망설임 표지

√ Kormos & Denes(2004): 1분당 발화된 음절수(발화속도) + 평균발화길이(Mean Length of Runs: MLR)

√ Skehan & Foster(2012): 발화된 평균 단어수를 나타내는 발화 길이(Length of Run: LOR) (원어민 화자: 분당 평균 4.5~6.0 단어, 비원어민 평균 3.5 단어 발화)

▶ 복잡성 (Complexity)

√ Williams & Evans, 1998 ; DeKeyser, 2008; Housen & Kuiken, 2009

- 1)인지적 복잡성(Cognitive Complexity) : 기억력, 개개인의 학습능력, 동기 등 학습자의 주관적 특징에 해당
- 2)언어적 복잡성(Linguistic Complexity) : 과제 수행에 사용되는 언어의 정교함과 다양한 정도



- ① Ortega(2003, 2009): 통사적 복잡성, 언어 사용의 정교함의 정도를 나타내는 통사적 성숙도로 정의
- ② Lintunen(2014): 통사적 복잡성
- ③ Foster & Skehan(2009): 통사적 다양성 + 언어적 정교함

▶ 복잡성 (Complexity) 측정도구

√ T-unit(minimal terminable unit) : 문어와 구어에서 전통적으로 사용해온 단위.

√ Hunt(1965): 완전한 문장을 형성하는 최소단위로 독립절과 그 독립절에 포함된 모든 절을 하나의 T-unit으로 구분.

√ T-unit의 단점: 구어에서 자주 발생하는 생략, 반복 현상 등을 포함시키지 않음.

√ AS-unit(Analysis of Speech Unit):

빈번한 생략과 반복, 수정이라는 구어의 특징에 맞는 측정 도구로

√ Foster(2000)가 제시

√ L2 말하기 복잡성 측정 도구로 광범위하게 사용.

2. CAF의 상호관련성

1) 자원제한(Trade-Off)관계 :

√ Ellis(1994): 유창성 ↑ 정확성과 복잡성 ↓

√ Skehan & Foster(2001): 사람의 집중능력과 작업기억 (working memory)는 한계 . 따라서 정확성과 복잡성이 함께 상승하지 않음. 복잡성 ↑ 정확성과 유창성 ↓

2) 인지가설(Cognitive Hypothesis):

√ Robinson(2001, 2003, 2005, 2007): 정확성과 복잡성이 경쟁관계나 반비례의 관계에 있지 않음

- 발화의 복잡성, 정확성 동시 ↑

√ Robinson & Gilabert(2007): 과제의 복잡성이 증가할수록 정확성, 복잡성 동시 ↑

√ Cadierno(2008): 복잡한 과제는 L2 학습자의 정확성과 어휘패턴의 복잡성 ↑

3. CAF와 L2 유창성

√ L2 학습에서 CAF는 말하기, 쓰기, 읽기, 듣기 네 가지 언어 영역 중 학습자의 말하기와 쓰기 수행을 측정하는 지표로 사용됨.

√ 과제의 난이도 및 유형, 준비시간의 유무에 따른 학습자의 CAF 분석이 주를 이룸.

1) L2 말하기 유창성과 CAF(국외연구)

✓ Robinson(2001): 일본인 대학생 44명을 대상으로 과제의 복잡성과 어려움, 과제의 유형에 따른 L2 학습자의 말하기 유창성을 CAF를 통해 분석함

- 과제가 어려울수록 어휘 다양성과 정확성이 동시에 증가함을 보고. 복잡성과 정확성이 대립하는 것이 아니라 지지관계에 있음을 보고.

✓ Yuan & Ellis(2003): 42명의 중국인 영어학습자를 대상으로 준비시간과 말하기 유창성의 관계를 연구.

- 과제 준비시간과 유창성 및 복잡성이 양의 관계에 있음을 보고.

✓ Larsen-Freeman(2006): 5명의 중국인 영어학습자의 말하기, 쓰기 유창성을 6개월에 걸쳐 CAF를 통해 분석

- 참가자 모두 CAF가 향상됨을 보고

✓ Crooks(1989) : 일본인 영어학습자 40명 대상. 과제 준비시간이 말하기 복잡성을 증가시키지만 정확성에는 영향을 미치지 못함.

✓ Ahangari & Abi(2011): 이란 영어학습자 40명 대상. 같은 결과 도출

✓ Ortega(1999): 준비시간을 가진 L2 학습자는 유창하고 복잡한 언어 산출 하지만 어휘범위에는 큰 차이가 없고 정확성은 일정한 결과가 없었음을 보고.

2) 국내연구

- √ **현세라(2007)**: 준비시간과 L2화자의 발화관계를 CAF를 통해 살펴봄. Ortega(1999)의 연구결과와 일치.
- √ **오미영(2008)**: 과제의 복잡성이 증가할수록 언어 복잡성을 증가시키지만 정확성과 유창성에는 유의미한 결론을 이끌어 내지 못함.

3) L2 쓰기 능숙도와 CAF(국외연구)

- √ **Frear & Bitchener(2015)**: 34명의 뉴질랜드 거주 비원어민 대상으로 과제복잡성과 쓰기 능숙도를 살펴봄.
 - 과제복잡성이 증가할수록 통사적 복잡성은 감소
- √ **Kuiken & Vedder(2012)**: 91명의 이탈리아 학생과 76명의 프랑스 대학생을 대상으로 과제복잡성이 정확성에는 영향을 미치지 않지만 통사적 복잡성, 어휘 다양성에는 영향을 미치지 않음을 보고.
- √ **Ellis & Yuan(2004)**: 42명의 중국대학생의 글쓰기 수행에서 준비시간을 가진 집단은 통사적 복잡성이 증가하지만 정확성에서는 유의미한 결과가 도출되지 않음.

4) 국내연구

√ **황은경(2012):** 글쓰기 장르에 따른 한국인 대학생의 글쓰기 특징을 CAF를 통해 살펴봄.

- 서사적 글쓰기에서 통사적, 어휘적 복잡성, 정확성이 나타났고 영어능력 상 그룹에서는 복잡성, 정확성, 유창성이 모두 일관성 있게 증가함을 보고.

√ **이민선(2018):** 고등학생을 대상으로 글쓰기 장르와 과업 복잡성이 쓰기 과업수행에 미치는 영향을 CAF를 통해 살펴봄. (같은 연구 결과 도출)

√ **주우정, 박매란(2016):** 글쓰기 과업복잡성의 정도가 성인 학습자의 CAF에 미치는 영향을 연구.

- 과업이 복잡할수록 상, 하 그룹 모두에서 정확성이 증가함 → Robinson(2001, 2002)의 인지가설을 뒷받침

- 과업이 복잡할수록 상 그룹에서만 복잡성 산출 증가

- 유창성은 상, 하 그룹 모두 과업의 복잡성 유무와 관련 없음 → 혼재된 연구결과 도출

Ⅲ. 연구 방법

1. 연구 대상

√ **광역시 소재 4년제 대학**

√ **일반 교양 영어회화와 비즈니스 영어회화 수강생 52명 (남학생 31명 (59.6%), 여학생 21명(40.4%))**

√ **평균 나이 22.8세**

√ **전공:** 패션 디자인, 영어 영문, 컴퓨터 공학, 건축학과, 전기공학, 호텔 컨벤션 경영학, 간호학, 인간시스템공학, 재활상담, 기계공학, 컴퓨터 소프트웨어, 행정학, 유아교육 등

√ **영어권 거주 경험:** 전체 참여자 중 20명(38.5%)

1년 이상 11명(55%), 1년 미만 9명(45%)

2. 분석틀

표 1
CAF 측정도구

Complexity	Accuracy	Fluency
절의 개수/ AS-unit개수	무오류의 개수/AS-unit개수	1분당 발화된 단어의 수
종속절의 개수/AS-unit 개수	수일치의 오류/AS-unit개수	
Norris & Ortega(2009)	시제오류/AS-unit 개수	Skehan & Foster(2012)
	자가 수정횟수	
	Yuan & Ellis(2003)	

√ 본 논문은 한국 대학생들의 영어 말하기에 나타난 특징을 CAF 분석을 통해 살펴보았으며

√ 빈번한 수정과 휴지, 역양, 생략 등 구어의 특징에 알맞게 구현된 AS-unit(Analysis of Speech Unit)을 사용.

AS-unit이란 :

Foster et al, 2000: 화자의 단일 발화로 독립절이나 절 이하의 단위(sub-clausal unit). 이때 독립절은 최소한의 정형 혹은 비정형 동사로 구성됨.

Ex) "I have no opportunity to visit." (정형동사 1개==> 1개의 절, 1개의 AS-unit)

Ex) "It is my hope to study crop protection." (정형동사 1개, 비정형동사 to study 목적어 crop protection → 2개의 절, 1개의 AS-unit)

Ex) "How long have you been here?"/ "Three month" → 1개의 AS-unit

3. 연구 절차

- √ 2021. 12월 4주차~ 22년 2월 셋째 주 : 관련 선행 연구 분석
- √ 2022. 3월 2주차~ 22년 4월 첫째 주: 영어 말하기 테스트 질문선정

" Which would you prefer, online shopping or offline shopping and why?"

- √ 2022. 4월 둘째 주~: 부산광역시 소재 4년제 대학교 " 비즈니스 영어회화 " 수업과 " 일반 교양 영어회화" 에 등록된 총 52명의 한국인 영어학습자를 연구 대상으로 선정. 연구목적 및 방법, 연구 참여에 대한 동의를 확인후 연구 실험 시작.

- √ 연구자와 한 명의 실험참여자가 순차적으 교실에서 진행하였고 2주간에 걸쳐 이루어짐. 영어 말하기 테스트전 30초간의 준비시간이 주어짐.

- √ 연구 참여자 52명 중 불성실하거나 영어 말하기 테스트 포기의사를 밝힌 2명을 제외한 50명을 최종 연구 대상으로 선정.

- √ 2022. 5월 둘째 주~ 넷째 주 : 연구 참여자의 1분 말하기 테스트 자료의 녹음된 내용을 모두 전사함.

- √ 2022. 6월 첫째 주~ 7월 둘째 주: 현재 4년제 대학교에서 강의하고 있는 원어민 강사가 전체 연구 참여자의 1/3에 해당하는 17명의 CAF 점수를 채점함. 연구자가 나머지 33명의 CAF점수를 채점.

√ 1분 말하기 테스트에서 추출한 CAF 점수와 수집된 설문자료는 SPSS25.0 프로그램을 사용하여 통계분석 수행.

√ CAF 점수에 따라 상,하 그룹으로 구분한 후 이들 그룹간 CAF 차이는 t-test를 통해 검증.

√ CAF간의 상관관계는 피어슨(pearson)분석 수행.

IV. 연구 결과

표 2
영어 말하기 능력(CAF측정) 상, 하 그룹의 독립표본 *t* 검정 결과

영어 말하기 능력	Levene		t-test	
	F	<i>p</i>	<i>t</i>	<i>p</i>
	7.857	.007	10.077	.000

* Levene: 등분산을 결정하지 않음

표 3
상, 하 그룹 영어 말하기 능력 기술통계

	참여자수	최소값	최대값	평균	표준편차
전체	50	11.0	134.20	56.56	29.70
상 그룹	26	50.20	134.20	79.84	21.26
하 그룹	24	11.0	50.0	31.34	10.58

표 4
복잡성(C), 정확성(A), 유창성(F)에 대한 기술통계

			참여자수	최소값	최대값	평균	표준편차
말하기 능력	복잡성	전체	50	1.0	4.80	2.64	.93
		상그룹	26	1.80	4.80	2.99	.79
		하그룹	24	1.0	4.0	2.26	0.93
	정확성	전체	50	.00	4.70	2.25	1.09
		상그룹	26	1.30	4.70	2.69	.98
		하그룹	24	.00	4.00	1.78	1.01
	유창성	전체	50	8.00	126.0	51.66	28.83
		상그룹	26	46.0	126.0	74.15	20.90
		하그룹	24	8.00	110.0	39.17	18.00

1. 복잡성(C), 정확성(A), 유창성(F) 요인간의 관계

✓ 복잡성, 정확성, 유창성 요인간의 상호 관련성을 살펴보기 위해 피어슨 상관관계 분석을 실시.

표 5
요인간 상관계수

		복잡성	정확성	유창성
말하기 능력	복잡성	1	.446	.430
	정확성		1	.398
	유창성			1

✓ Bartz(1999) .20-매우 약한 상관관계 / .20-40 낮은 상관관계 / .40-60 비교적 높은 상관관계 / .60-80 높은 상관관계 / .80 이상 매우 높은 상관관계로 정의

표 6
'복잡성' 기준 독립표본 t검정 결과

		Levene	t-test		
'복잡성'		F	p	t	p
			.026	.872	10.431

* Levene: 등분산을 결정함.

표 7
'복잡성' 기준 상, 하 그룹 기술통계

	상 하 그룹	참여자 수	최소값	최대값	평균	표준편차
'복잡성'	(상)	25	2.60	4.80	3.41	.549
	(하)	25	1.0	2.50	1.87	.493

표 8
상, 하 그룹의 '복잡성' 과 요인간 상관관계

		정확성	유창성
'복잡성'	(상)	.734	.249
		p	.000
	(하)	.506	.322
		p	.010

표 9
'정확성' 기준 독립표본 t 검정 결과

'정확성'	Levene		t-test	
	F	p	t	p
	3.032	.088	8.563	.000

표 10
'정확성' 기준 상, 하 그룹 기술통계

	상하 그룹	참여자 수	최소값	최대값	평균	표준편차
'정확성'	(상)	24	2.10	4.70	3.13	.79
	(하)	26	.00	2.00	1.45	.58

표 11
상, 하그룹의 '정확성' 과 요인간 상관관계

		복잡성	유창성
'정확성'	(상)	.190	.090
		p	.373
	(하)	.466	.319
		p	.016

표 12
'유창성' 기준 독립표본 *t* 검정 결과

	Levene		t-test	
'유창성'	F	<i>p</i>	<i>t</i>	<i>p</i>
	8.458	.005	9.937	.000

* Levene: 등분산을 결정함.

표 13
'유창성' 기준 상, 하 그룹 기술통계

	상하 그룹	참여자 수	최소값	최대값	평균	표준편차
'유창성'	(상)	26	46.0	126.00	74.15	20.90
	(하)	24	8.00	44.00	27.29	10.20

표 14
상, 하그룹의 '유창성' 과 요인간 상관관계

			복잡성	정확성
'유창성'	(상)		.323	.055
		<i>p</i>	.107	.791
	(하)		.026	.221
		<i>p</i>	.906	.229

V. 결론

√ 본 논문의 연구목적


1. 한국인 대학생들의 말하기 특징을 '복잡성', '정확성', '유창성'(CAF)의 지표로 살펴보고
2. 세 요인간의 상관관계를 분석하여
3. 한국인 대학생들의 영어 말하기 능력을 향상시키는 방안을 살펴보고자 함.

√ 본 논문의 연구 결과

1. CAF간의 상관관계

- 복잡성과 정확성 .446,  비교적 높은 상관관계
- 복잡성과 유창성 .430
- 정확성과 유창성 .398로  낮은 상관관계

2. 복잡성 기준 상, 하 분류

- 상그룹: 복잡성과 정확성: .734  높은 상관관계
- 하그룹: 복잡성과 정확성 : .506

상그룹 : 복잡성과 유창성 : .740  낮은 상관관계

√ 1, 2번의 연구결과

◆ 복잡성과 정확성은 상호 지지하는 관계임을 보고한 Robinson(2001, 2005, 2007, 2012)의 연구결과와 일치함을 확인함.

3. 유창성 기준 상, 하 분류

- 상, 하 그룹 모두 에서 복잡성, 정확성과 통계적으로 유의미한 관계가 나타나지 않음.

4. 정확성 기준 상, 하 분류

-상그룹: 정확성과 복잡성: .190
정확성과 유창성: .09 → 매우 낮은 상관관계

-하그룹: 정확성과 복잡성: .466
정확성과 유창성: .319 → 비교적 높은 상관관계

√ 4번의 연구결과 :

·상 그룹에서는 정확하게 발화하는 문장은 복잡성이 낮은 반면, 발화문장이 복잡해지면 오류가 잦음을 유추가능

-하 그룹의 경우 문장구조의 복잡성은 낮으나 낮은 복잡성을 가진 문장을 발화할 경우 오류가 적음을 유추가능.



그룹에 따른 상이한 단계별 전략을 채택할 필요를 보여줌.

상 그룹: '복잡성'을 바탕으로 한 유창성 향상이

하 그룹: 문법적인 '정확성'을 함양할 수 있는 학습지도를 통해 영어 문장 구조를 확대할 수 있는 '복잡성'의 단계로 나아가야 함을 함의

√ 본 논문의 기여

- 기존 선행연구에서 보고되지 않았던 '복잡성', '정확성', '유창성'의 개별 요소를 기준으로 한 요인간 상관관계를 세부적으로 연구하여
- 상, 하 그룹에서 영어 말하기에 영향을 미치는 주요 요소를 통계적으로 확인함
- '복잡성', '정확성', '유창성' 요인의 중요성에 대한 인식을 제고할 수 있었다.
- 영어 말하기 능력의 세 가지 요소 중 가장 중요한 요인으로 판명된 '정확성' 요소함양을 위해서는 교수자의 일방적 문법 설명보다 일상생활에서 사용되는 문장을 통해 자연스러운 문법 규칙의 노출 및 그룹별 상호 피드백을 통한 flipped learning을 대안으로 제시함.

√ 본 논문의 한계

- 참여자의 표본수가 제한적임으로 연구 결과를 일반화하는데 한계가 있을 수 있음.
- 정확성 측정시 Pallotti(2009)가 제안한 바와 같이 문법적인 오류는 없으나 의사소통을 저해하는 요소를 포함시켜 측정의 정확성도를 높일 필요가 있음.

◆ 참고 문헌

고재현. (2004). 고등학생 영어 학습자의 "교과 교사로서" 일반 교사: 수업관, 교원관, 학습관을 중심으로. (박사학위논문), 연세대학교.

김수연, 김병호. (2010). 한국어 학습자의 구어 산출에 대한 통사적입성 발달. *한국언어교육학회지*, 28(1), 99-117.

신현영. (2008). 전화영어 수업이 한국어 학습자 발달에 미치는 영향, 정확성, 유창성에 미치는 영향. *The Journal of Studies in Language*, 36(1), 99-119.

이정연. (2003). 국문 학습자의 회소로 산출로 교사의 역할에 대한 교사의 역할에 대한 영향. (박사학위논문), 연세대학교.

서희재. (2003). 한국어 학습자의 회소로 산출로 교사의 역할에 대한 영향. *한국언어교육학회지*, 21(1), 11-21. 연세대학교.

유수영, 박재민. (2010). 화상 학습자의 구어 산출에 대한 교사의 역할에 대한 영향. *한국언어교육학회지*, 28(1), 139-155.

한재원. (2007). 한국어 학습자의 회소로 산출에 대한 통사적입성 발달. *한국언어교육학회지*, 25(1), 11-21. 연세대학교.

오석영. (2005). 한국어 학습자의 회소로 산출에 대한 통사적입성 발달. *한국언어교육학회지*, 23(1), 11-21. 연세대학교.

Johnson, R., & Adams, M. (2011). The effect of pre-task planning on the accuracy and complexity of second L2 learners' oral performance. *Proceedings - Social and Behavioral Sciences*.

Colwell, T. (2005). Learning to self-orient within a foreign language. In P. Robinson and N. C. Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*. Routledge.

Colwell, T. (2007). Monitoring language performance: complexity, accuracy and fluency measures. In 2007 TESOL International Academic Conference Proceedings.

Corbett, G. (1999). Planning and second language variation. *Studies in Second Language Acquisition*, 11, 267-281.

DeKeyser, R. (1998). The complexity of defining complexity. Paper presented at AAAA 2001, April 24, 2001, Washington D.C.

DeKeyser, R. (2012). *Second L2 On-Lap Learners: Linguistic Features, Grammar and Argumentative Philosophy*. (박사학위논문), 연세대학교.

Hansen, A., & Kalkan, T. (2009). Complexity, accuracy and fluency in second language acquisition. *Applied Linguistics*, 30(4), 441-471.

Ellis, R. (1984). *A theory of instructed second language acquisition in SLA*. (박사학위논문), 연세대학교.

Ellis, R. (2003). *Task-based Language Learning and Teaching*. Oxford University Press.

Ellis, R. and G. Barkhuizen. (2002). *Assessing Learner Language*. Oxford University Press.

DeKeyser, R. (2012). *Second L2 On-Lap Learners: Linguistic Features in Grammar and Argumentative Philosophy*. Washington State University degree of Doctor of Philosophy.

Fraser, P., & Skehan, P. (1990). The influence of Planning and Task Type on Second Language Performance. *Studies in Second Language Acquisition*, 12(5), 289-321.

Fraser, P., Taylor, A., and Fitzgerald, G. (2000). *Measuring English Language: A Guide for all Purposes*. *Applied Linguistics*, 21(5), 554-575.

Hansen, A., & Van Der Stoep, H. (2005). *Task complexity and the effectiveness of explicit grammar instruction*. In Hansen, A. and H. Van Der Stoep (eds.), *Investigation in Instructional Second Language Acquisition*. Maastricht University.

Kalkan, T. and H. Diken. (2014). Exploring accuracy and perception of fluency in the speech of second language learners. *System*, 52(2), 145-164.

Kim, M. (2010). The Development of Complexity, Accuracy and Fluency in L2 Written Production Through Mediated Participation in Online Activities. *CALL Research Journal*, 16(1), 73-86.

Kalkan, T., & Van der Stoep, H. (2007). Task complexity and measures of linguistic performance in L2 writing. *IRAL - International Review of Applied Linguistics in Language Teaching*, 45(3), 261-284.

Larsen, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, 40(1), 167-167.

Larsen-Freeman, D. (2006). The co-occurrence of complexity, fluency, and accuracy in the oral and written production of first-class learners of English. *Applied Linguistics*, 27(4), 596-619.

Larsen-Freeman, D. (2008). *Adopting Experience: The Study of Complexity, Accuracy, and Fluency in Second Language Acquisition*. *Applied Linguistics*, 29(4), 579-589.

Ortega, L. (1998). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, 11(1998), 139-148.

Ortega, L. (2005). Systemic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, 26(4), 482-518.

Hansen, A. & Ortega, L. (2008). Towards an Organic Approach to Investigating CAF in Second L2: The Case of Complexity. *Applied Linguistics*, 29(4), 551-575.

Pallant, G. (2006). CAF, Defining, Refining and Differentiating Constructs. *Applied Linguistics*, 27(4), 589-601.

Robinson, P., Tang, S., Wu, S., & Sun, J. (2005). Investigating Second Language Task Complexity. *TESOL Journal*, 16(1), 42-76.

Robinson, P. (2001). Task complexity, task difficulty, and task production: exploring relationships in experimental SLA. *Applied Linguistics*, 22(1), 17-21.

Robinson, P. (2001). *Attention and memory during SLA*. Blackwell Publishing Ltd.

Robinson, P., & Odell, R. (2007). Task complexity, the Cognitive Complexity and second language learning and performance. *IRAL - International Review of Applied Linguistics in Language Teaching*, 45(3), 161-176.

Robinson, P., Callahan, T., & Skehan, T. (2008). Time and fluency: Measuring the Effect of the Conceptual Demands of Tasks on Second Language Speech Production. *Applied Linguistics*, 29(4), 529-559.

Segalowitz, N. (2000). *Automaticity and Attentional Shift in Fluent Performance*. In P. Eggins (ed.), *Perspectives on Fluency: Aue Author*. University of Michigan Press.

Skehan, P. (1998). *A Cognitive Approach to Language Learning*. Oxford University Press.

Skehan, P. (2008). *Measuring Second Language Performance Integrating Complexity, Accuracy, Fluency, and Lexis*. *Applied Linguistics*, 29(4), 533-552.

Skehan, P. and P. Foster. (2012). Complexity, accuracy, fluency and form in oral second language performance. In Hansen, A., Diken, T., Kalkan, T. (eds.), *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA*. John Benjamins Publishing Company.

Tokkyo, S. (2017). Short-term changes in complexity, accuracy and fluency: Developing progress monitor proficiency test in Yasuoka et al. (eds.)

Loehrer, P. and Nöcker, M. (2013). Measuring syntactic complexity in spoken and written second language: Comparing the SImplicity? Research in Language, 12(1), 177-206.

Wills-Gomes, E., S. Saeed and H. T. Kim. (1998). Second Language Development in writing: Measures of Fluency, Accuracy, and Complexity. Toronto, ON: University of Toronto, Second Language Teaching and Curriculum Centre.

Yuan, F. and Ellis, R. (2003). The Effects of Pre-Task Planning and On-Line Planning on Fluency, Accuracy and Anxiety in L2 Monologue Oral Production. Applied Linguistics, 24(1), 1-27.

😊 감사합니다.



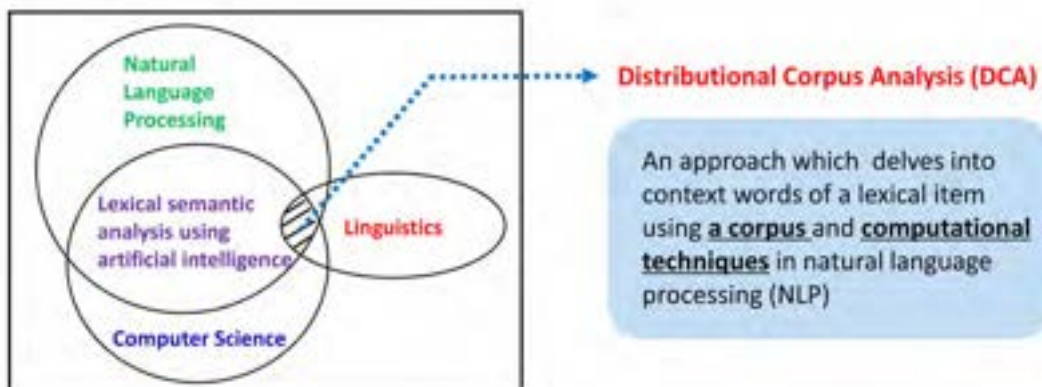
Lexical Semantic Analysis by means of Artificial Intelligence

Feb.16.2023

Wonbin Kim
wk32657@uga.edu

Department of Linguistics
University of Georgia

DCA in Linguistics



DCA and Lexical Semantics

Semantic structures

- (i) the relationship of semantic similarity
- (ii) lexical relations such as synonymy, antonymy, and hyponymy
- (iii) **syntagmatic lexical relations**

↓

Collocational Analysis
Distributional Corpus Analysis

3

Objective of Doctoral Research

[Dissertation]
Distributional Corpus Analysis of Korean Neologisms using Artificial Intelligence

[Target Words]
Three Korean neologisms: *leyal* (레이얼) & *lwuce* (루스) & *kay-* (개-)

[Research Objective]
To track semantic change of the three Korean neologisms by analyzing **their context words** using **Korean Twitter corpora** and **natural language processing (NLP) techniques**

4

Summary of Doctoral Research

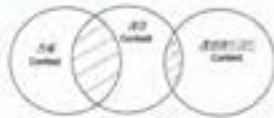
DCA

- ▶ Corpus: Twitter corpus for each neologism
- ▶ NLP techniques: AI methods
- ▶ Context words

(1) 레알

e.g. *이거 레알이에요?* (의미: '이거 진짜예요?')

- Word2vec & Cosine Similarity
- "The contexts of semantically similar words are similar to each other." (Harris 1954)
- Comparison of context: *레알마드리드* vs *진짜*



(2) 루저

e.g. *키가 174인데 그럼 저는 루저인가요?*

- Latent Dirichlet Allocation (LDA)
- Number of context ↑ Number of topics ↓
- Comparison of numbers of topics: '패배자' corpus vs '키 180이하 남자' corpus

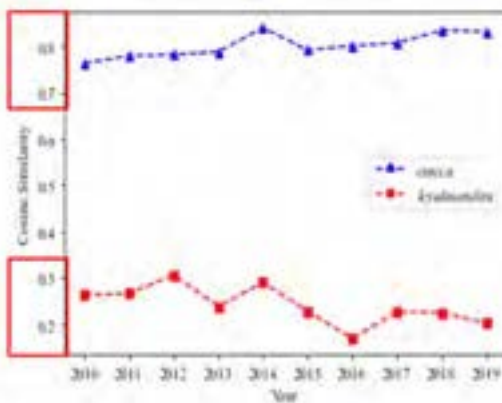
(3) 개

e.g. *개피곤하다, 개예쁘다* (의미: '너무 피곤하다', '정말 예쁘다')

- Long Short-Term Memory (LSTM)
- LSTM: sorting words into positive and negative
- Comparison of frequencies of positive and negative words: *개나쁘다, 개짜증난다, 개멋지다...* → semantic prosody

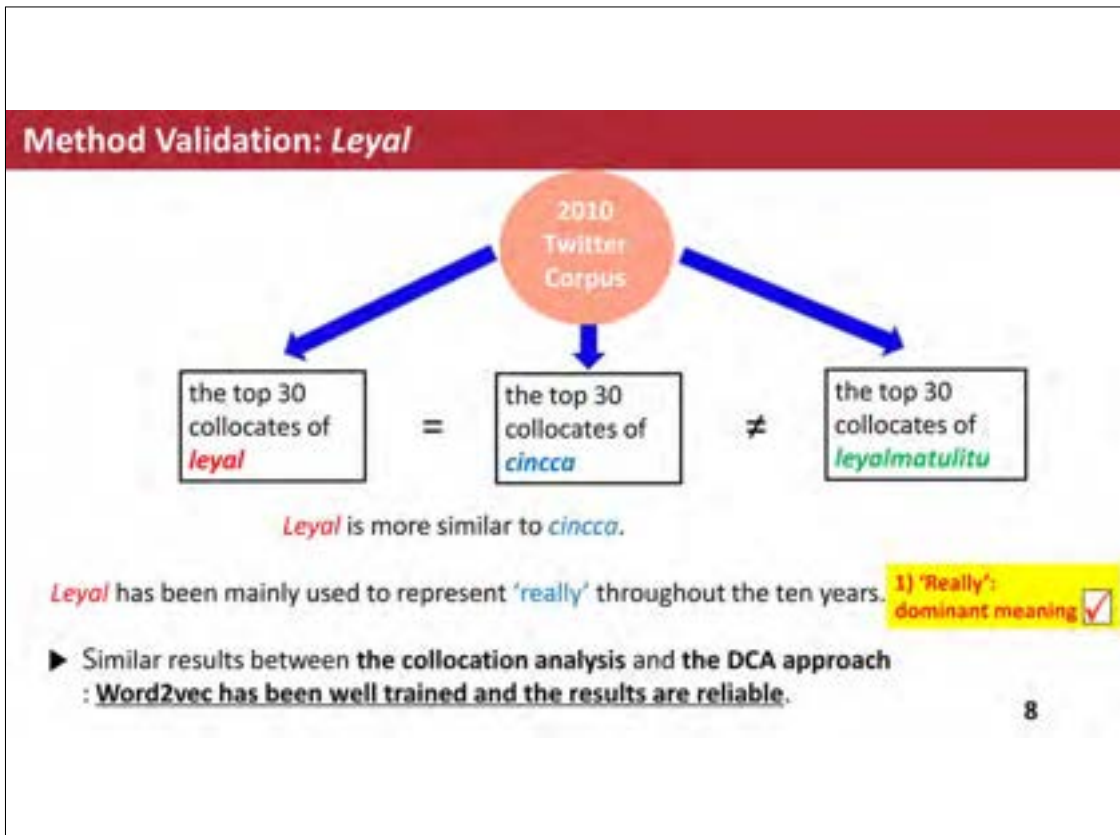
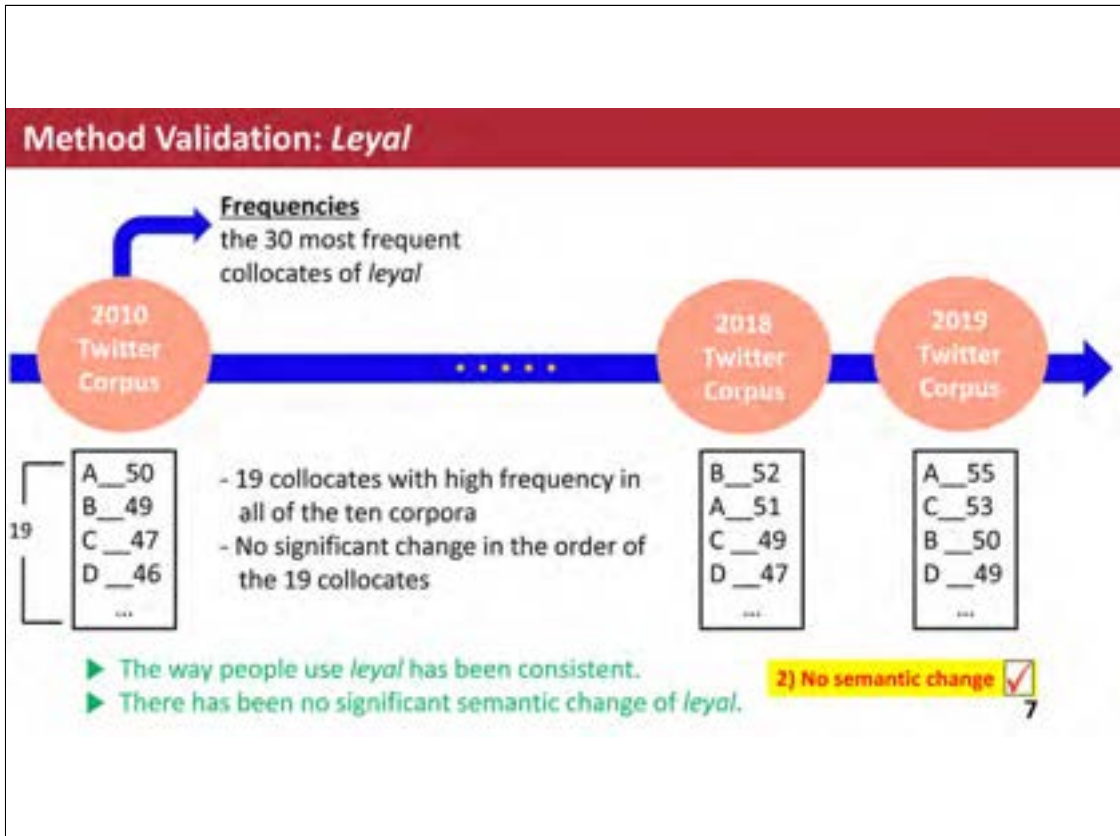
5

Results: *Loyal*

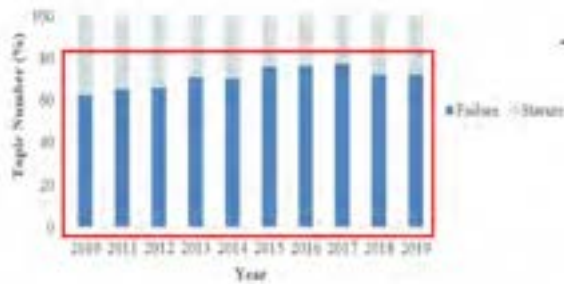


- The meaning of the neologism *loyal* is more similar to that of *cincca*, which implies that the neologism *loyal* has been more used to represent 'really' than 'Real Madrid'. **1) 'Really': dominant meaning**
- In terms of the relationship between the two meanings, there is no semantic change. **2) No semantic change**

6



Results: *Lwuce*



- The proportion of number of topics from the existing meaning subcorpus is always higher than that from the new meaning subcorpus, which implies that the existing meaning of *lwuce* is more dominant throughout the ten years.

1) Existing meaning:
always more dominant
-> no semantic change

9

Results: *Lwuce*

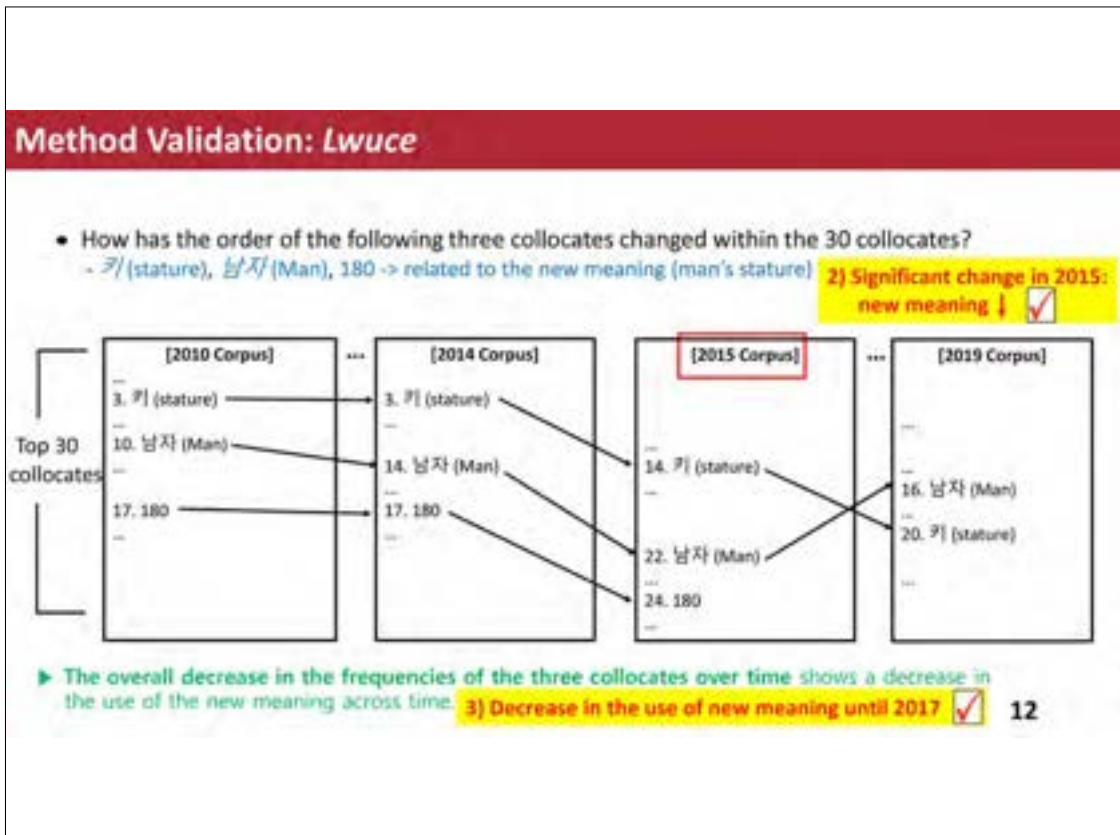
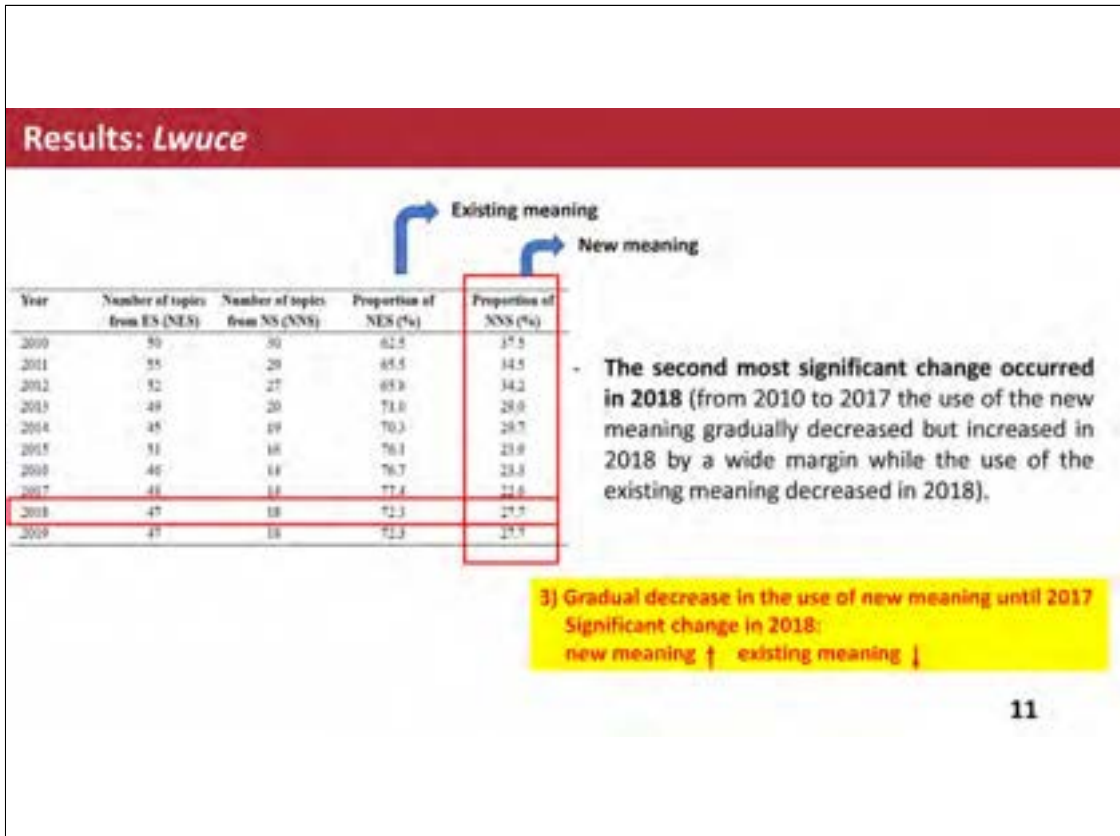
↩ Existing meaning ↪ New meaning

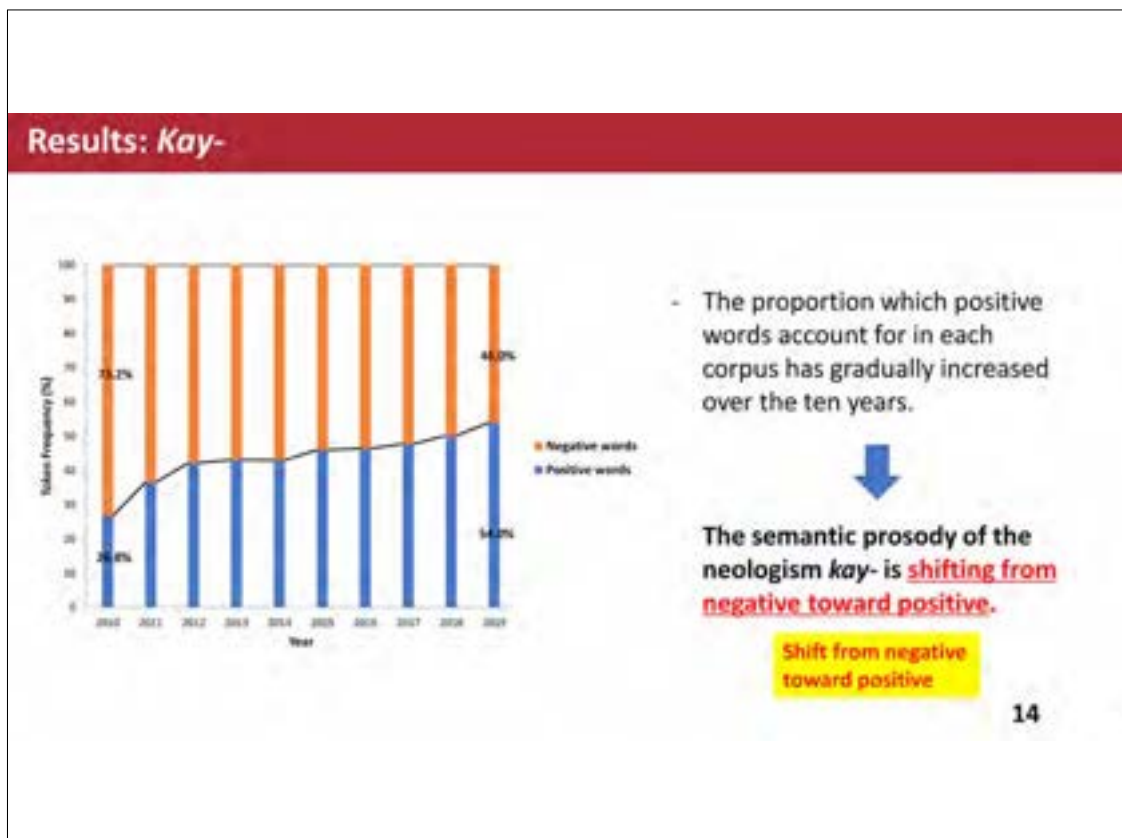
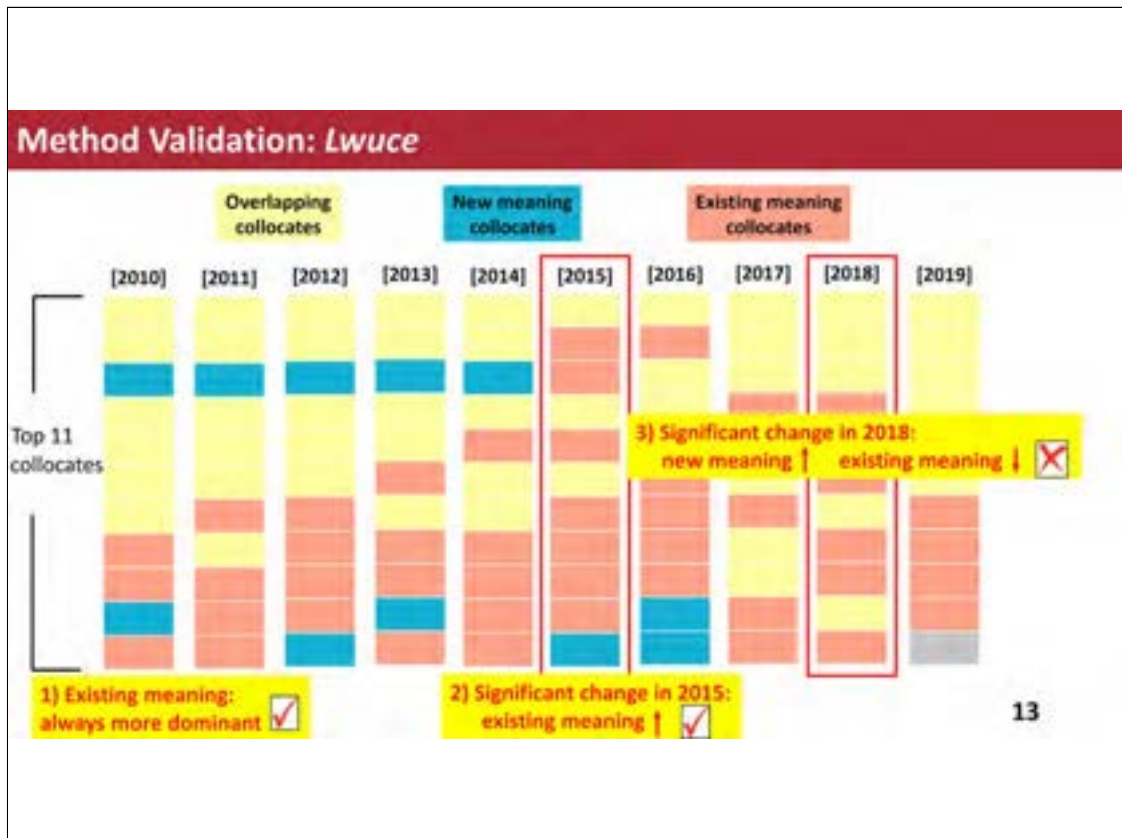
Year	Number of topics from ES (NES)	Number of topics from NS (NNS)	Proportion of NES (%)	Proportion of NNS (%)
2010	50	30	62.5	37.5
2011	55	29	66.5	33.5
2012	52	27	65.8	34.2
2013	68	20	71.0	29.0
2014	41	18	75.3	24.7
2015	51	16	76.1	23.9
2016	48	14	70.7	29.3
2017	48	14	77.4	22.6
2018	47	18	72.3	27.7
2019	47	18	72.3	27.7

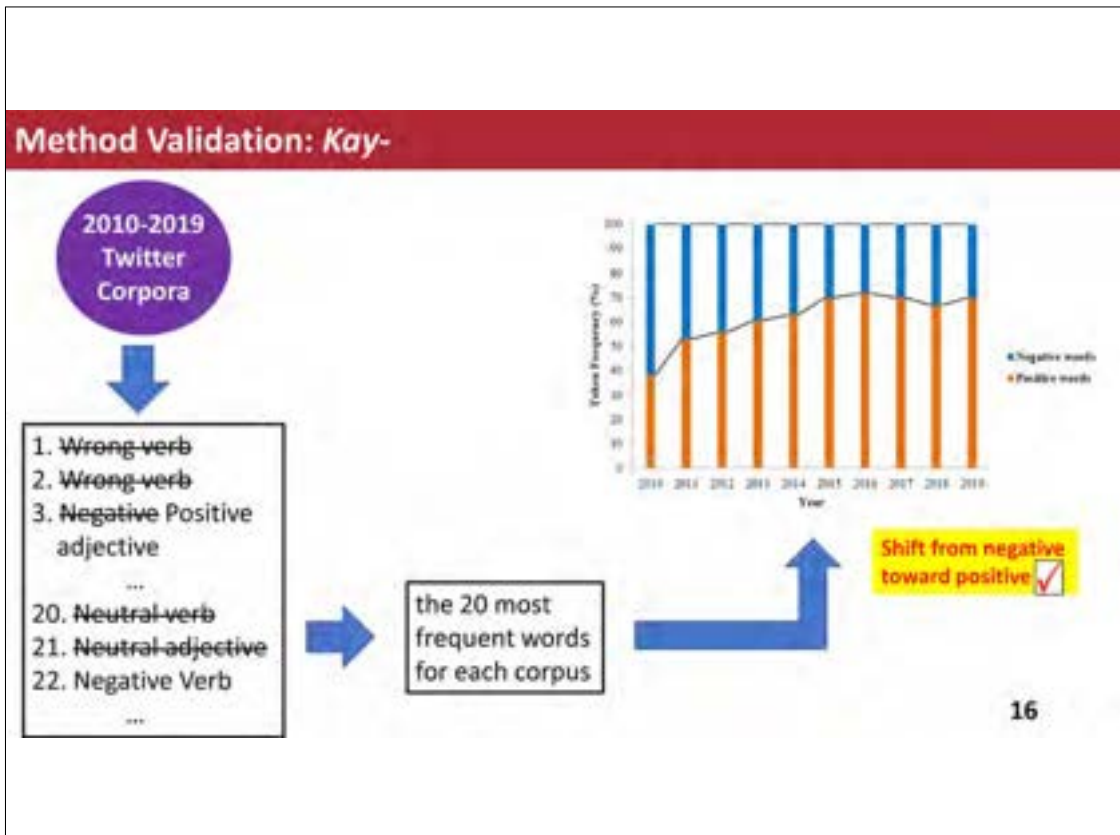
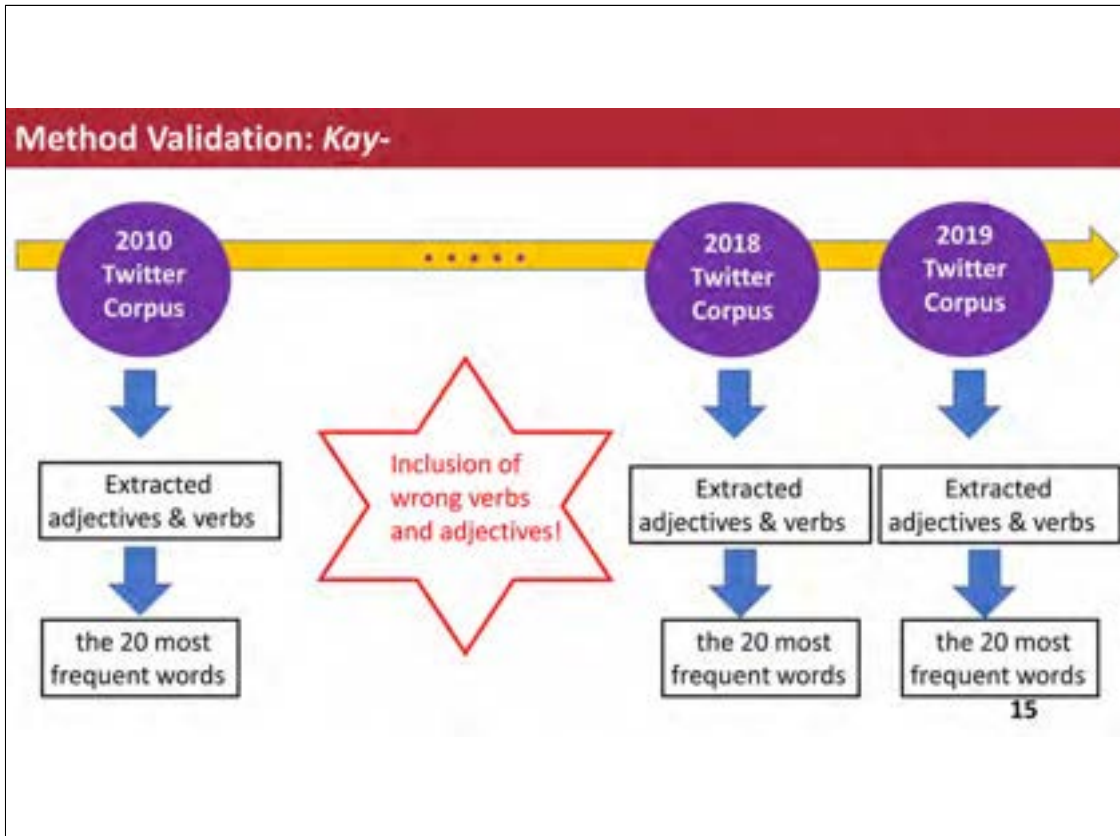
- The most significant change in the use of the two meanings occurred in 2015 (the difference of proportions between 2014 and 2015 is greatest).

2) Significant change in 2015:
new meaning ↓
existing meaning ↑

10







Conclusion

[Limitation of DCA]

The DCA approach still needs method validation. It is needed to check whether the used models have been well trained and their results are reliable.

- ▶ Consideration of variables and exceptions — Understanding of language

Expertise of linguists

- ▶ Prudent use of computational techniques in DCA & Careful acceptance of results

17

Conclusion

[Advantages of DCA]

Efficiency: when we need to handle a large amount of data
(faster & more convenient process and analysis + less time & cost)

Quantification: when we need to observe the degree of language change visually
(clear and objective analysis)

- ▶ Few DCA studies in Linguistics — Improvement and Development of DCA

Active research of linguists

18