The 2024 Seoul International Conference on Linguistics (SICOL-2024)

Reenvisioning linguistic creativity: Language and AI in interaction

Presentation Materials

August 8 (Thu) - 9 (Fri), 2024
Kyung Hee University, Seoul

Co-hosted by
The Linguistic Society of Korea &
Kyung Hee Institute for the Study of Language and Information

한국언어학회
The Linguistic Society of Korea

General Information

Conference Website:

Please visit the following conference website for details about the conference including the conference program, venue, registration, and the handbook:

https://lsk202324.github.io/2024_summer/index.html

All information about the conference will be available at this site. Participants are asked to check this site to keep up to date regarding possible alternations and changes.

Preregistration:

Online participation is free. You do not have to pay registration fee to attend the conference online. We will send the information regarding our Zoom meeting to LSK members, presenters and moderators. Non-members and overseas participants are recommended to register for this hybrid conference by filling out and submitting the registration form available at the following site so we can provide the Zoom IDs, passwords and links:

https://forms.gle/xn1wNv87bYXurPHB7

The same online registration form is also available at the 'Registration' tab of the conference website. After receiving your registration form, we will send you the information regarding our Zoom meeting.

If you've registered, but have not gotten the Zoom IDs and links, please email us at lsk202324@gmail.com.

CONTENTS

SICOL-2024 Committees

Organizing Committee:

Jong-Bok Kim (Chair, LSK president, Kyung Hee University)
Youngju Choi (Chosun University)
Sumi Han (Hallym University)
Chung-Hoon Lee (Sogang University)
Suk-yeong Mun (Seoul National University)
Yunju Nam (Hanyang University)
Dongwoo Park (Korea National Open University)
See-Gyoon Park (Kunsan National University)
Byong-Rae Ryu (Chungnam National University)

Program Committee:

Se-Eun Jhang (Co-chair, Korea Maritime and Ocean University
Hanjung Lee (Co-chair, Sungkyunkwan University)
Byeongkil Ahn (Gyeongsang National University)
Kyungchul Chang (Pusan National University)
Nayoun Kim (Sungkyunkwan University)
Hansaem Kim (Yonsei University)
Eon-Suk Ko (Chosun University)
Iksoo Kwon (Hankuk University of Foreign Studies)
Jungmee Lee (Seoul National University)
Seonmin Park (Korea Advanced Institute of Science and Technology)

Conference Theme

Reenvisioning Linguistic Creativity:
Language and AI in Interaction

The theme for SICOL-2024 is 'Reenvisioning linguistic creativity: Language and AI in interaction.' Recent advances in the development of Artificial Intelligence (AI) have significant implications for academia in general and language education and research specifically. Linguistic research utilizing AIs usually focuses on what they can do with language. There are growing discussions of how AIs based on a large language model (LLM) understand and use language. Less investigated, however, has been the possible role of novel generative AI tools, such as AI LLM chatbots, in assisting researchers in the conception, development, and execution of research and data analysis.

SICOL-2024 provides a forum for the presentation of cutting-edge research focusing on the theoretical and empirical study of 'language research in the era of post-generative AI' from both perspectives: what AIs can do with language and what they can do for linguists.

The SICOL-2024 organizing committee welcomes abstract submissions for 20-minute oral presentations, followed by 10 minutes for discussion and questions. We invite papers related to the conference theme such as applications of AI in language research, language education, and communication. We also invite papers from all subfields of linguistics (morphology, syntax, semantics, pragmatics, discourse, language acquisition, variation, psycholinguistics, computational linguistics).

SICOL-2024 features four invited presentations by the following distinguished scholars:
John Beavers (The University of Texas at Austin, USA)
Benjamin Bruening (University of Delaware, USA)
Thomas Hoffmann (KU Eichstätt-Ingolstadt, Germany)
Bernd Kortmann (University of Freiburg, Germany)

It also features six themed sessions which include papers addressing issues related to the conference theme and seventeen general sessions which include papers from all areas of linguistics.

Welcome from the LSK President

Welcome to the 2024 Seoul International Conference on Linguistics (SICOL-2024).

We are excited to gather on August 8-9, 2024, for two days of insightful and inspiring discussions. This year's conference, themed 'Reenvisioning Linguistic Creativity: Language and AI in Interaction,' brings together leading experts, researchers, and enthusiasts from around the world to explore the new era of linguistics where traditional knowledge interplays with artificial intelligence.

Our program is packed with 66 presentations that cover a wide range of topics across various subdisciplines of linguistics. We are honored to feature four plenary talks by distinguished scholars, John Beavers, Benjamin Bruening, Thomas Hoffmann, and Bernd Kortmann, who will share their visionary insights and expertise. They will offer valuable perspectives and inspire deep and stimulating discussions throughout the conference. We are also deeply grateful to all the presenters for their invaluable contributions, which are at the heart of this event. We extend our heartfelt gratitude to the Organizing and Program Committee members, especially the Program Committee Co-chairs Se-Eun Jhang and Hanjung Lee, and moderators for their dedication and hard work. My thanks go to all who have helped make this conference a great one.

We hope this conference will be not only a platform for presenting creative and genuine research but a unique opportunity to build connections and foster collaborations as well. We believe that SICOL-2024 will provide you with fresh insights into the field of linguistics. Once again, welcome, and enjoy the conference!

Warm wishes,

Jong-Bok Kim
President, The Linguistic Society of Korea

Program

Program Overview (Korea Standard Time (KST) [UTC + 9:00])

Day 1: Thursday, August 8, 2024

Room No.	Room 1 (#715)	Room 2 (#704)	Room 3 (#705)	Room 4 (#706)
09:00-09:15	Registration			
09:15-09:25	**Opening Ceremony (Room 1)** Welcome Address: Jong-Bok Kim (LSK President, Kyung Hee University, Korea) Moderator: Dongwoo Park (Korea National Open University, Korea)			
09:30-11:00	**General Session 1** Syntax A Incheol Choi 1. (Withdrawn) 2. Daiho Kitaoka 3. Barry C.-Y. Yang	**General Session 2** Phonetics/Phonology A Joo-Kyeong Lee 4. Kyle Ho, Kuo-Chiao Lin 5. Hayeun Jang 6. Sheng-Chieh Huang, Yan-Bin Huang, Kuo-Chiao Lin, Zi-Xiang Lin, Ryan Wu		**Themed Session 1** Language and AI A Eunkyung Yi 7. Nguyen Van Nguyen, HoaMing Truong, Issra Pramoolsook 8. So Young Lee, Russell Scheinberg, Amber Shore, Ameeta Agrawal 9. Ching-Yu Yang, Jhih-Jie Chen
11:10-12:10	**Invited Talk 1 (Room 1)** Speaker: John Beavers (The University of Texas at Austin, USA) Moderator: Juwon Lee (Jeonju University, Korea)			
12:10-13:30	Lunch Break			
13:30-15:00	**General Session 3** Syntax B Sanghee Park 10. Zi-chun Lin (online) 11. Tatsuya Hashimoto (online) 12. Min-Joo Kim (online)	**General Session 4** Applied Linguistics A Sun-Young Lee 13. Eunjin Gye, Wooyeol Lim, Youngjoo Kim 14. Yeonkwon Jung 15. Ben Yue Shi Yang (online)	**General Session 5** Text/Corpus Linguistics A Hansaem Kim 16. Masaaki Ogura (online) 17. Salima Araar (online) 18. Ziyun Dai (online)	**Themed Session 2** Language and AI B Hye-Won Choi 19. Sarah Park, Sang-Geun Lee 20. Ye-eun Cho 21. Kaoru Amino
15:10-16:40	**General Session 6** Syntax/Semantics Sang-Geun Lee 22. Rebecca Kirkle, Eugene Chung 23. Yuki Ishihara 24. Hojeong Koo	**General Session 7** Applied Linguistics B Seonmin Park 25. Mi Kyong Kim (online) 26. Ki-tae Kim (online) 27. Xiaotian Wang (online)	**General Session 8** Text/Corpus Linguistics B Okgi Kim 28. Elena V. Beloglazova 29. Cuilin Liu, Se-Eun Jhang 30. Jinawat Kaenmuang (online)	**Themed Session 3** Language and AI C Sumi Han 31. Hisashi Morita (online) 32. Jose Belda-Medina, Sandra Abad-Bataller, Kaoutar Smaili-Bensaad, Andrea Gallardo-Ferrandez, Jorge Soto-Almela, Jose R. Calvo-Ferrer (online)
16:50-17:50	**Invited Talk 2 (Room 1; Online)** Speaker: Thomas Hoffmann (KU Eichstätt-Ingolstadt, Germany) Moderator: Iksoo Kwon (Hankuk Univerisity of Foreign Language, Korea)			

Day 2: Friday, August 9, 2024

Room No.	Room 1 (#715)	Room 2 (#704)	Room 3 (#705)	Room 4 (#706)
09:15-09:30	Registration			
	General Session 9 Semantics/Pragmatics A	**General Session 10** Phonetics/Phonology B	**General Session 11** Language Processing	**Themed Session 4** Sign Language Linguistics A
	Jungmee Lee	Inkie Chung	Jeong-Ah Shin	Se-Eun Jhang
09:30-11:00	33. You-Yun Yu 34. Dongeun Lee 35. Sue Young Chung, Eugene Chung	36. Chun Ting Chien, François-Xavier Brajot 37. Geonhee Lee 38. Riley Kent, Eunhae Oh	39. Eun Seon Chung 40. Nayoun Kim, Ziying Li, Sanskriti, Kwangsu Kim 41. (Withdrawn)	42. Haewon Jeon 43. Ahmed Mai, Soon-Bok Kwon (online) 44. Youngju Choi (online)
11:10-12:10	**Invited Talk 3 (Room 1)** Speaker: Benjamin Bruening (University of Delaware, USA) Moderator: Heejeong Ko (Seoul National University, Korea)			
12:10-13:30	Lunch Break			
	General Session 12 Syntax C	**General Session 13** Semantics/Pragmatics B	**General Session 14** Computational Linguistics	**Themed Session 5** Sign Language Linguistics B
	Jungsoo Kim	Dongsik Lim	Sangah Lee	Haewon Jeon
13:30-15:00	45. Mija Kim 46. Manabu Mizuguchi 47. Pin-Wei Li, Jen Ting	48. Taku Kumakiri 49. Ahbi Koh (online) 50. Ben Yue Shi Yang, Jenny Yichun Kuo (online)	51. Kyung Eun Lee 52. Jieun Kim	53. Charmhun Jo 54. Seonhye Lee 55. Se-Eun Jhang, Sook Ki Lee, Byeong-Chen Yoon, Haewon Jeon, Seonhye Lee
	General Session 15 Syntax D	**General Session 16** Phonetics/Phonology C	**General Session 17** Sociolinguistics	**Themed Session 6** Language and AI D
	Bum-Sik Park	Haeil Park	Eugene Chung	Seulkee Park
15:10-16:40	56. Ji-in Kang 57. Ching-Syuan Shen, One-Soon Her	58. Anna S.C. Cheung (online) 59. Zi-Xiang Lin (online) 60. Man-ni Chu, Yu-Chun Wang (online)	61. Ajay Malik, Usha Udaar 62. Balázs Horváth (online) 63. Oleg Shcherbakov (online)	64. Erika Kristine E. Arcenal, Licca Pauleen V. Capistrano, Marielle Jessie D. De Guzman, Micaela Isabel M. Forrosuelo, Janeson M. Miranda (online) 65. Evgeniya Aleshinskaya (online) 66. Matthew Galbraith, Preeti Kumari, Martina Wiltschko (online)
16:50-17:50	**Invited Talk 4 (Room 1; Online)** Speaker: Bernd Kortmann (University of Freiburg, Germany) Moderator: Suwon Yoon (University of Seoul, Korea)			
17:50-18:00	**Closing Ceremony (Room 1)** Closing Remarks: Hanjung Lee (Program Committee Co-chair, Sungkyunkwan University, Korea) Moderator: Suwon Yoon (University of Seoul, Korea)			

Presentation Lists

Day 1: Thursday, August 8, 2024

9:30-11:00	**General Session 1: Syntax A**
	Room #715
Moderator	Incheol Choi (Kyungpook National University, Korea)
Presentations	1. (Withdrawn)
	2. Daiho Kitaoka (University of Nottingham Ningbo China (UNNC), China) **Dividing Plurals in Japanese and the Heterogeneity of the Syntax of Plurals**
	3. Barry C.-Y. Yang (National United University, Taiwan) **Is Chinese a Partial Null Subject Language?**

9:30-11:00	**General Session 2: Phonetics/Phonology A**
	Room #704
Moderator	Joo-Kyeong Lee (University of Seoul, Korea)
Presentations	4. Meng-Kai Kyle Ho (Kang Chiao International School, Taiwan) Kuo-Chiao Lin (Kang Chiao International School, Taiwan) **Modeling Schwa Reduction in Directional Harmonic Serialism**
	5. Hayeun Jang (Sungkyunkwan University, Korea) **Phonological Preparation in Korean: The Influence of Shared Structures and Units**
	6. Sheng-Chieh Huang, Yan-Bin Huang, Kuo-Chiao Lin, Zi-Xiang Lin, Ryan Wu (Kang Chiao International School, Taiwan **Trochaic Template Satisfaction in Isbukun Bunun**

9:30-11:00	**Themed Session 1: Language and AI A**
	Room #706
Moderator	Eunkyung Yi (Ewha Woman's University, Korea)
Presentations	7. Nguyen Van Nguyen, Hoa Ming Truong, Issra Pramoolsook (Suranaree University of Technology, Thailand) **ChatGPT-Generated Modifications on Human-Generated TESOL Abstracts by Vietnamese Researchers**
	8. So Young Lee (Miami University, USA), Russell Scheinberg, Amber Shore, Ameeta Agrawal (Portland State University, USA) **Resolving RC Attachment Ambiguity in LLMs: Contrasting Head-Initial and Head-Final Language Processing**
	9. Ching-Yu Helen Yang (National Chung Hsing University, Taiwan) Jhih-Jie Chen (National Tsing Hua University, Taiwan) **An Ablation Study on the Automated Essay Scoring Using ChatGPT**

11:10-12:10	**Invited Talk 1**
	Room #715
Moderator	Juwon Lee (Jeonju University, Korea)
Speaker	John Beavers (The University of Texas at Austin, USA) **On the Interface of Word Meaning and Grammar**

13:30-15:00	**General Session 3: Syntax B**
	Room #715
Moderator	Sanghee Park (Hanbat University, Korea)
Presentations	10. Zi-chun Lin (National Taiwan Normal University, Taiwan) [online] **A Movement-based Analysis for Non-local Binding: Chinese _ziji_ and Its New Perspectives**
	11. Tatsuya Hashimoto (University of Tsukuba, Japan) [online] **Dividing Categorial Selection: Towards a Mechanism behind Non-Argument Taking Relation in Syntax**
	12. Min-Joo Kim (Texas Tech University, USA) [online] **Three-Way Classification of ACC-Marked External Possession in Korean and Its Implications**

13:30-15:00	**General Session 4: Applied Linguistics A**
	Room #704
Moderator	Sun-Young Lee (Cyber Hankuk University of Foreign Studies, Korea)
Presentations	13. Eunjin Gye, Wooyeol Lim, Youngjoo Kim (Kyung Hee University, Korea) **Processing Patterns of Korean Adverbial Sentences with '-(eu)ro': Comparing Fixed and Free Word Orders**
	14. Yeonkwon Jung (Kansai Gaidai University, Japan) **Japanese Academic Apology: A Case of English as a Lingua Franca Encounter**
	15. Ben Yue Shi Yang (National Chung Cheng University, Taiwan) [online] **Artificial Intelligence-Assisted Learning in High-Functioning Autistic Children for Processing Modern Chinese Proverbs**

13:30-15:00	**General Session 5: Text/Corpus Linguistics A**
	Room #705
Moderator	Hansaem Kim (Yonsei University, Korea)
Presentations	16. Masaaki Ogura (Osaka Metropolitan University, Japan) [online] **Textual Patterns in _The Rambler_: A Preliminary Stylistic Study of Samuel Johnson and Canonical Authors**
	17. Salima Araar (University of Limerick, Ireland) [online] **The Use of the Pragmatic Marker _Like_ by Algerian Speakers of English in Ireland:** **A Corpus-Based Discourse Analytic Approach**
	18. Ziyun Dai (Korea Maritime and Ocean University, Korea) [online] **Research Trends in _Linguistic Research_ from 2014 to 2023: A Bibliometric Analysis**

13:30-15:00	**Themed Session 2: Language and AI B**
	Room #706
Moderator	Hye-Won Choi (Ewha Woman's University, Korea)
Presentations	19. Sarah Park (Korea University, Korea), Sang-Geun Lee (Korea University, Korea) **Intuitive Minds vs. AI Algorithms: Anaphoric Marker Choices in Korean Bridging Contexts**
	20. Ye-eun Cho (Sungkyunkwan University, Korea) **Prompting Strategies of Generative AI for Korean Pragmatic Inference**
	21. Kaoru Amino (ShanghaiTech University, China/Japan) **Why Do Humans Feel Unsatisfied with Conversations with Unsupervised Chatbots?: A Comparative Study on the Topical Depth Using T-R Nexus and FSP**

15:10-16:40	**General Session 6: Syntax/Semantics**
	Room #715
Moderator	Sang-Geun Lee (Korea University, Korea)
Presentations	22. Rebecca Kirkle, Eugene Chung (Korea University, Korea) **Dummy or Not? Exploring the Referentiality of the English Pronoun 'It'**
	23. Yuki Ishihara (Tokyo Institute of Technology, Japan) **The Role of Addressee in Interpreting Predicate Doubling in Japanese**
	24. Hojeong Koo (Sogang University, Korea) **Semantic Case Markers in Korean Noun Phrases: Focusing on *-ey tayhan* and *-ey uyhan***

15:10-16:40	**General Session 7: Applied Linguistics B**
	Room #704
Moderator	Seonmin Park (The Korea Advanced Institute of Science and Technology, Korea)
Presentations	25. Mi Kyong Kim (Chodang University, Korea) [online] **ChatGPT and Google Bard for Critical-PBLL in Korean University English Education**
	26. Ki-tae Kim (Keimyung University, Korea) [online] **Positioning Analysis of AI-Based Medicine in Published Expert Interviews: A Preliminary Analysis**
	27. Xiaotian Wang (Harvard University, USA) [online] **Word-order Reflexes of Information Structure under the Influence of Bilingualism: A Pilot Study on Mandarin-English Bilingual Speakers**

15:10-16:40	**General Session 8: Text/Corpus Linguistics B**
	Room #705
Moderator	Okgi Kim (Kyung Hee University, Korea)
Presentations	28. Elena V. Beloglazova (Herzen State Pedagogical University, Russia) **Translating a Culture: Evidence from the Japanese, English and Finnish Comparable Corpora**
	29. Cuilin Liu, Se-Eun Jhang (Korea Maritime and Ocean University, Korea) **A Comparative Bibliometric Analysis of Gender Equality Research Trends in the Maritime Industry and Global Domains**
	30. Jinawat Kaenmuang (Chulalongkorn University, Thailand) [online] **Integrating Cultural Contexts in AI: A Study on Thai Boxing Lexicon and Language Comprehension**

15:10-16:10	**Themed Session 3: Language and AI C**
	Room #706
Moderator	Sumi Han (Hallym University, Korea)
Presentations	31. Hisashi Morita (Aichi Prefectural University, Japan) [online] **Does Generative AI Dream of Human-like Translation?: Implementing Obviation Control Rules into an LLM through Finetuning**
	32. Jose Belda-Medina, Sandra Abad-Bataller, Kaoutar Smaili-Bensaad, Andrea Gallardo-Ferrandez, Jorge Soto-Almela, Jose R. Calvo-Ferrer (University of Alicante, Spain) [online] **AI Assessment Tools in Language Learning: Bridging the Gap between Theory and Practice**

16:50-17:50	**Invited Talk 2 (online)**
	Room #715
Moderator	Iksoo Kwon (Hankuk University of Foreign Studies, Korea)
Speaker	Thomas Hoffmann (KU Eichstätt-Ingolstadt, Germany) **Creativity in Construction Grammar: The 5C Model of Linguistic Creativity**

Day 2: Friday, August 9, 2024

9:30-11:00	General Session 9: Semantics/Pragmatics A
	Room #715
Moderator	Jungmee Lee (Seoul National University, Korea)
Presentations	33. You-Yun Yu (National Taiwan University Graduate Institute of Linguistics, Taiwan) **Hands that Speak: Metaphorical and Metonymical Insights from Korean 손 'Hand'- Related Idioms**
	34. Dongeun Lee (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany) **Understanding Embodiment in Phraseological Expressions - Body Part Nouns in German and Korean**
	35. Sue Young Chung, Eugene Chung (Korea University, Korea) **Pragmatic Analysis of Testimonial Strategies in Perjury Court Rulings**

9:30-11:00	General Session 10: Phonetics/Phonology B
	Room #704
Moderator	Inkie Chung (Sogang University, Korea)
Presentations	36. Chun Ting Chien, François-Xavier Brajot (National Taipei University of Nursing and Health Sciences, Taiwan) **Investigating Mandarin Clear Speech: Acoustic Cues and Their Applications in Technology and Therapy**
	37. Geonhee Lee (Stony Brook University, USA) **Prosodic Tones and Boundaries: Disambiguating Non-canonical *Wh*-Questions and *Wh*-Exclamatives in Korean**
	38. Riley Kent, Eunhae Oh (Konkuk University, Korea) **Perceptual Adaptation to Foreign Accented Speech by L2 Listeners**

9:30-11:00	General Session 11: Language Processing
	Room #705
Moderator	Jeong Ah Shin (Dongguk University, Korea)
Presentations	39. Eun Seon Chung (Hankyong National University, Korea) **Second Language Processing of Errors in Korean-to-English Machine-translated Output**
	40. Nayoun Kim, Ziying Li, Sanskriti, Kwangsu Kim (Sungkyunkwan University, Korea) **The Mechanism behind Understanding Anaphoric *One* and Noun Phrase Ellipsis: How Language Models Work**
	41. (Withdrawn)

9:30-11:00	Themed Session 4: Sign Language Linguistics A
	Room #706
Moderator	Se-Eun Jhang (Korea Maritime and Ocean University, Korea)
Presentations	42. Haewon Jeon (Korea Maritime and Ocean University, Korea) ***Wh*-doubling Constructions in Korean Sign Language**
	43. Ahmed Mai, Soon-Bok Kwon (Pusan National University, Korea) [online] **Epistemic Modality Expression in Korean Sign Language**
	44. Youngju Choi (Chosun University, Korea) [online] **Double Mapping Constraint and KSL Metaphors**

11:10-12:10	**Invited Talk 3**
	Room #715
Moderator	Heejeong Ko (Seoul National University, Korea)
Speaker	Benjamin Bruening (University of Delaware, USA) **Subjectless Readings of *Again* and Argument Structure**

13:30-15:00	**General Session 12: Syntax C**
	Room #715
Moderator	Jungsoo Kim (Incheon National University, Korea)
Presentations	45. Mija Kim (Kyung Hee University, Korea) **Korean Expletive Negation in a Subordinate Clause: Focusing on Its Variants** 46. Manabu Mizuguchi (Kokugakuin University, Japan) **Revisiting Phase-internal Movement** 47. Pin-Wei Li, Jen Ting (National Taiwan Normal University, Taiwan) **On the Syntactic Category of Rúguǒ 'if' and Its Implications for Clausal Structure in Mandarin Chinese**

13:30-15:00	**General Session 13: Semantics/Pragmatics B**
	Room #704
Moderator	Dongsik Lim (Hongik University, Korea)
Presentations	48. Taku Kumakiri (Shukutoku University, Japan) **Function of the Diminutive in Narrative of Tunis Arabic** 49. Ahbi Koh (Kobe College, Japan) **Japanese *Yabai* with Empathy** 50. Ben Yue Shi Yang (National Chung Cheng University, Taiwan) Jenny Yichun Kuo (National Chiayi University, Taiwan) **Human Processing of Animal Metaphors in English and Chinese**

13:30-14:30	**General Session 14: Computational Linguistics**
	Room #705
Moderator	Sangah Lee (Seoul National University, Korea)
Presentations	51. Kyung Eun Lee (Seoul National University, Korea) **A Study on the Output Type of Korean Particle '-의 [ɰi]' in Automatic Speech Recognition (ASR)** 52. Jieun Kim (University of Ulsan, Korea / University of California, Santa Cruz, USA) **Evaluating Unsupervised Parsing: DIORA and The Quest for Structural Understanding in Natural Languages**

13:30-15:00	**Themed Session 5: Sign Language Linguistics B**
	Room #706
Moderator	Haewon Jeon (Korea Maritime and Ocean University, Korea)
Presentations	53. Charmhun Jo (Southwest University, China) **Conceptual Metaphor and Double Mapping in Chinese Sign Language** 54. Seonhye Lee (Pukyong National University, Korea) **Grammaticalization of Adverbial Clauses in Spoken and Sign Languages** 53. Se-Eun Jhang (Korea Maritime and Ocean University, Korea), Sook Ki Lee (Korea Nazarene University, Korea), Byeong-Cheon Yoon (Chosun University, Korea), Haewon Jeon (Korea Maritime and Ocean University, Korea), Seonhye Lee (Pukyong National University, Korea) **Lexical Development in Young Deaf Twins: A Comparative Longitudinal Study of Korean Sign Language Acquisition**

15:10-16:10	**General Session 15: Syntax D**
	Room #715
Moderator	Bum-Sik Park (Dongguk University, Korea)
Presentations	56. Ji-in Kang (Hankuk University of Foreign Studies, Korea) **A Construction-Based Approach to the English *Is-it-just-me-or-X*? Construction**
	57. Ching-Syuan Shen (National Chengchi University, Taiwan), One-Soon Her (National Chengchi University / Tunghai University, Taiwan) **Identifying Sortal Classifiers in Korean**

15:10-16:40	**General Session 16: Phonetics/Phonology C**
	Room #704
Moderator	Haeil Park (Kyung Hee University, Korea)
Presentations	58. Anna S.C. Cheung (The Hong Kong Polytechnic University, China) [online] **Teaching Mandarin Tones: How to Make It Comprehensible and Compelling**
	59. Zi-Xiang Lin (Kang Chiao International School, Taiwan) [online] **Revisiting Extrametricality in Isbukun Bunun**
	60. Man-ni Chu (Fu Jen Catholic University, Taiwan), Yu-Chun Wang (Dharma Drum Institute of Liberal Arts, Taiwan) **The Inspiration of Seq2Seq and Its Error Analysis – The Sound Changes of *Chaoshan* and *Zhangzhou***

15:10-16:40	**General Session 17: Sociolinguistics**
	Room #705
Moderator	Eugene Chung (Korea University, Korea)
Presentations	61. Ajay Malik (Central University of Rajasthan, India) Usha Udaar (IIT Kanpur, India) [online] **A Sociolinguistic Areal Analysis of a Three-Village Micro-Linguistic Area (MLA) at the Haryana-Punjab Border in India**
	62. Balázs Horváth (Corvinus University of Budapest, Hungary) [online] **"I" or "we" in a Pandemic? A Linguistic Case Study on the Personalization of Hungarian Political Communication during the Coronavirus**
	63. Oleg Shcherbakov (Lomonosov Moscow State University, Russia) [online] **Language as Symbol in Urban Linguistic Landscapes: Comparative Analysis of China, Thailand, and Russia**

15:10-16:40	**Themed Session 6: Language and AI D**
	Room #706
Moderator	Seulkee Park (Kyung Hee University, Korea)
Presentations	64. Erika Kristine E. Arcenal, Licca Pauleen V. Capistrano, Marielle Jessie D. De Guzman, Micaela Isabel M. Forrosuelo, Janeson M. Miranda (De La Salle University Integrated School, The Philippines) [online] **Impersonally Married: A Comparative Genre Analysis of Reddit and ChatGPT-4 Spouse Appreciation Posts**
	65. Evgeniya Aleshinskaya (National Research Nuclear University MEPhI, Russia) [online] **Applying AI in Teaching Academic Writing: A Pilot Study**
	66. Matthew Galbraith (Universitat Pompeu Fabra, Spain), Preeti Kumari (Universitat Pompeu Fabra, Spain), Martina Wiltschko (ICREA, Universitat Pompeu Fabra, Spain) [online] **Grammar Plays a Role in Human-Computer Interaction, Huh?**

16:50-17:50	**Invited Talk 4 (online)**
	Room #715
Moderator	Suwon Yoon (University of Seoul, Korea)
Speaker	Bernd Kortmann (University of Freiburg, Germany) **World Englishes in the Context of World Languages**

Offline Venue Information

Location

#704, #705, #706, and #715

Chungwoon Building (청운관), Kyung Hee University, Seoul

Address: 26, Kyungheedae-ro, Dongdaemun-gu, Seoul, Korea

Registration fee (in-person)

- Faculty or PhD holders: KRW 30,000

- (Under)graduate students: KRW 10,000   [우리은행 1002-363-787147 (예금주: 이주원)]

Free Wifi

- SSID: KHU Wi-Fi Guest

- Password : vision2020

* Eduroam also available

Parking on Campus

- 5,000 KRW per day

- Purchase a ticket at the Registration booth

Cafeteria

Lunch will be ready on the B1 floor of the venue building. You can take the elevators.
Bring your coupon and show it to the employee at the entrance.

SICOL-2024 Website

You can find useful information (including Travel & Lodging) on SICOL-2024 Website.
Visit website: https://lsk202324.github.io/2024_summer/index.html

Online Venue Zoom Links

Day 1: Thursday, August 8, 2024

Room No.	Sessions	Zoom ID	Link
Room 1 (#715)	General Session 1, 3, and 6 Invited talk 1 and 2	863 8544 8163	https://khu-ac.zoom.us/j/86385448163
Room 2 (#704)	General Session 2, 4, and 7	868 0836 5390	https://khu-ac.zoom.us/j/86808365390
Room 3 (#705)	General Session 5 and 8	810 8222 3759	https://khu-ac.zoom.us/j/81082223759
Room 4 (#706)	Themed Session 1, 2, and 3	875 6732 6620	https://khu-ac.zoom.us/j/87567326620

Day 2: Friday, August 9, 2024

Room No.	Sessions	Zoom ID	Link
Room 1 (#715)	General Session 9, 12, and 15 Invited talk 3 and 4	863 8544 8163	https://khu-ac.zoom.us/j/86385448163
Room 2 (#704)	General Session 10, 13, and 16	868 0836 5390	https://khu-ac.zoom.us/j/86808365390
Room 3 (#705)	General Session 11, 14, and 17	810 8222 3759	https://khu-ac.zoom.us/j/81082223759
Room 4 (#706)	Themed Session 4, 5, and 6	875 6732 6620	https://khu-ac.zoom.us/j/87567326620

* No passwords required

* Please mute your microphone upon entry

Contact us (Technical support):

lsk202324@gmail.com

Invited Talks

On the Interface of Word Meaning and Grammar

John Beavers

(The University of Texas at Austin, USA)

It has long been thought that the meaning of a word determines aspects of its grammatical properties, and nowhere has this idea been more explored than in the domain of verbs. Theories of verb meaning usually assume that verbs are organized in a language's vocabulary either in terms of the kinds of eventualities they describe, or in terms of the kinds of entities that participate in those eventualities. These semantic concepts are furthermore thought to be organized into semantic structures — events that are decompose into subevents, or coherent casts of participants — that determine the verb's syntactic and morphological behavior. In this talk I explore the ins and outs of such theories, and suggest that in many cases the proposed semantic structures are often not motivated by anything other than the grammatical properties they were meant to explain. I show instead that a theory of a verb's grammatical properties can instead be rooted in its truth conditional content, i.e. what must be true of the world to even use the word, and the ways truth conditions can contrast. I also explore other factors that govern how words are used grammatically when truth conditions do not, including pragmatics and common conventions of word usage, and even some places where lexical semantic structure does in fact matter. The case studies I explore come from data on English, Colloquial Sinhala, Bahasa Indonesian, and the Uto-Aztecan language O'dam. I ultimately suggest that the relationship of word meaning to grammar is governed by a soup of factors, but with actual meaning at the heart of it all.

Creativity in Construction Grammar:
The 5C Model of Linguistic Creativity

Thomas Hoffmann

(KU Eichsättt-Ingolstadt, Germany)

Creativity is a design feature of human language. In this talk, I will present a cognitive model of verbal creativity that draws on insights from the psychological research into creativity – particularly Glăveanu's 5A model that distinguishes five crucial perspectives on a creative act (actors, audience, artefacts, actions and affordances). It will outline a linguistic version of this model that adopts Construction Grammar as its theoretical foundation. The resulting "5C model of constructional creativity" (Hoffmann 2024, fc.) argues that the central elements of linguistic creativity are constructors, co-constructors, constructs, constructional blending and the constructional network. As I will show, this model also allows us to address the question as to whether Large Language Models can be considered creative or not.

Subjectless Readings of *Again* and Argument Structure

Benjamin Bruening
(University of Delaware, USA)

According to Kratzer (1996), internal arguments are arguments of verbs but external arguments are not. While much recent work has adopted this model, Bale (2007) and Smith & Yu (2021) present a strong challenge to it with the claim that it makes incorrect predictions about readings of the adverb 'again'. Specifically, they claim that an external argument cannot be excluded from the presupposition of 'again' if the verb is an unergative or a stative transitive, but only if it is an eventive transitive. Since the Kratzer model treats all external arguments the same---of eventive transitives, stative transitives, and unergatives---it apparently fails to capture the data. We show that the empirical claims of Bale (2007) and Smith & Yu (2021) are incorrect: unergatives and stative transitives do allow subjectless readings of 'again' if an appropriate context is presented. Unaccusatives do not, however. This precisely matches the predictions of the Kratzer model. We conclude that internal arguments are arguments of the selecting verb, but external arguments are not.

World Englishes in the Context of World Languages

Bernd Kortmann

(The University of Freiburg, Germany)

The special status of English among the world's major languages as "hyper-central" language and "hub of the World Language System" (de Swaan 2002) is undisputed and most likely irreversible. But what about the World Englishes complex (Mesthrie/Bhatt 2008)? Which role do Englishes and English-based creoles play in the current concert of world languages, especially those (standard and non-standard) varieties and creoles at the hyper-central, super-central and central level of Mair's (2013) World System of Englishes? And what will their likely role be in the future, be it as providers of spreading morphosyntactic features and patterns, as drivers of the spread of English, or as candidates of newly emerging world languages themselves? In the latter context, one consequence this may have and is worth exploring is the more systematic inclusion of individual World Englishes and English-based creoles in AI applications on a global and regional level, especially in the relevant parts of the Anglophone world.

General Sessions
(Day 1)

(Withdrawn)

Dividing Plurals in Japanese and the Heterogeneity of the Syntax of Plurals

Daiho Kitaoka

(University of Nottingham Ningbo China (UNNC), China)

Synopsis: In this paper, I investigate the syntactic properties of two varieties of plurals in Japanese that have hitherto escaped attention in theoretical literature: the tati plurals with proper names (PN-tati) and plurals via reduplication (Redup). I propose that these plurals are dividing plurals and serve for mass-count distinction. Comparing these with the –tati plurals with common nouns, this analysis is in line with a recent view that the syntax of plurals is heterogeneous (Mathieu 2012). The existence of dividing plurals in Japanese, however, leads us to cast doubt on another recent view of the modifier-head distinction of plurals (Wiltschko 2008).

Plurals & Division: Since Ritter (1991), plurals are commonly assumed to project a nominal functional phrase (e.g., NumP). Borer (2005a) further claims that plural marking manifests the mass-count distinction (division), which is the prerequisite for participating in count syntax (e.g., *many/two ropes, *many/two rope, much rope)*. More recent analyses assert that the syntax of plurals is heterogeneous, rather than homogeneous. Namely, plurals can be realized as a head that serves for the mass-count distinction (e.g., Div in Borer 2005a), as other types of heads (e.g., #, n; Kramer 2016, Mathieu 2012), or as modifiers (Butler 2011, Wiltschko 2008). I employ the properties in (1) to determine the division of plurals; if (1a-e) apply to a plural, it divides.

(1) a. Allow an inclusive reading: *Do you have kids? If you have kids, please raise your hand.*
   b. Can be used as a nominal predicate: *They are students.*
   c. Obligatorily take a narrow scope: *I did not see spots on the floor.* (= no spots)
   d. Refer to kinds: *Dodos are extinct.*
   e.

While the English plural marker –s brings an inclusive reading and serves for division (1a), plurals can also be exclusive and just work for the singular-plural distinction (Mathieu 2012). In Arabic, as in (2), the plural can be applied to a singulativized collective noun. While the singulatives manifest division, the plural in (2) is exclusive and does not work for division.

(2) *hal   ʕindik   burtogaala-at?* [Arabic; Mathieu 2012]
   QUES   have-you   orange.FEM.PL   'Do you have oranges' (= have >1 orange?)

By separating division from counting and inclusiveness from exclusiveness, English bare plurals are divided but number neutral (and indefinite). Therefore, to determine the division of a plural, I also employ the properties of bare plurals, as in (1b-e) (Carlson 1977).

-Tati plurals: While it is widely observed that the –tati plural with common nouns does not show the properties in (1) (Nakanishi & Tomioka 2004, Kurafuji 1999, 2004), the –tati plural with proper names (PN-tati) has escaped attention in literature, although it is also well-known that –tati can be attached to proper names.

   First, I show that although –tati plurals are often linked to properties on D (e.g., specificity, definiteness, referentiality), those are not always observed; namely, -tati plurals can be non-specific/indefinite (Kurafuji 2003, Nakanishi & Tomioka 2004). For instance, by inserting scope-taking elements, the sentence forces a non-specific reading (regardless of the existence of –tati), as in (3).

(3) *Kono kooen-de-wa **itumo** kodomo-tati-ga sondeiru.*
   this park-in-TOP   always child-PL-NOM   playing

'In this park, there are always children playing.' (N&T04: 121)

Focusing on the non-specific use of –tati, I examine the division (properties in (1)) of –tati plurals with proper names. To make it easier to imagine a situation, I introduce a movie called *Riaru Oni Gokko* [Real Hide & Seek], where people named Sato are hunted and killed by demons of the king in a fictitious Japan during a specific time frame set by the king. In this story, some or many Satos are killed every day, as in (4). (Note that I here focus on the non-specific additive use of –tati, but abstract away an associative use and a specific use of -tati).

(4) *Kyoo go-nin no **Sato(-tati)**-ga koros-are-ta.*
    today 5-CL of Sato-PL-NOM   kill-PASS-PAST   'Today, five **Satos** were killed.'

I show, as in (5), that the PN-tati patterns with bare plurals in English and shows the properties in (1), with which I claim that the PN-tati is a dividing plural.

(5) a. *Kyoo   oni-wa   Sato-tati-o tukamaen-akat-ta.   *Hitori-dake da.*
    today   demon-TOP   Sato-PL-ACC catch-NEG-PAST   1.CL-only be
    'Today, demons did not catch Satos. *Only one.' (Inclusive: 1a)
   b. *Karera-wa Sato-tati-da.*   c. *Oni-ga   Sato-tati-o   sagasiteiru.*
    they-TOP   Sato-PL-COP   demon-NOM Sato-PL-ACC looking.for
    'They are Satos.' (Nominal Predicate: 1b)   'Demons are looking for Satos.' (narrow scope: 1c)
   d. *Onigokko   no zikan da-kara,   Sato-tati-wa mareda.*
    Hide-and-Seek of   time be-because   Sato-PL-TOP   rare

(6) a. *Kyoo   oni-wa   Sato-tati-o tukamaen-akat-ta.   *Hitori-dake da.*
    today   demon-TOP   Sato-PL-ACC catch-NEG-PAST   1.CL-only be
    'Today, demons did not catch Satos. *Only one.' (Inclusive: 1a)
   b. *Karera-wa Sato-tati-da.*   c. *Oni-ga   Sato-tati-o   sagasiteiru.*
    they-TOP   Sato-PL-COP   demon-NOM Sato-PL-ACC looking.for
    'They are Satos.' (Nominal Predicate: 1b)   'Demons are looking for Satos.' (narrow scope: 1c)
   d. *Onigokko   no zikan da-kara,   Sato-tati-wa mareda.*
    Hide-and-Seek of   time be-because   Sato-PL-TOP   rare
    'Because it is the Hide-and-Seek time, Satos are rare (= rarely seen).' (Kind reference: 1d)

<u>Reduplication:</u> Reduplicated nouns also denote plurality in Japanese, as in (6). Except for sporadic works (e.g., Sudo 2017), plurality via reduplication (Redup) does not enjoy a hot debate in theoretical literature. Although I acknowledge that Redup is not very productive, however, I show as in (7) that Redup shows the properties in (1) (thereby Redup is, I assume, a syntactic operation), with which I claim that Redup is a dividing plural. Thus, plurals in Japanese are heterogeneous, dividing (PN-tati, Redup) and counting (tati).

(7) a. *hito-bito* 'person'   b. *yama-yama* 'mountain'   c. *kami-gami* 'gods'

(8) a. A: *Sima-zima-o   mi-rare-masi-ta   ka.*   B: *Hai, hitotu-dake miemasita.*
    island-REDUP-ACC see-can-POL-PAST   QUES   yes 1.CL-only   visible
    'Were you able to see islands?'   'Yes, only 1 was visible.' (Inclusive: 1a)
   b. *Karera-wa kami-gami-da.*   c. *Haruko-wa   kami-gami-ni aita-gat-teiru.*
    they-TOP   god-REDUP-COP   Haruko-TOP   god-REDUP-DAT meet-want-STAT
    'They are gods.' (Nom. Pred.: 1b)   'H. wants to meet gods.' (narrow scope only:1c)
   d. *Hana-bana-wa   zetumetusita.*
    flower-REDUP-TOP   extinct   'Flowers are extinct.' (Kind Ref.: 1d)

Head-Modifier: Plurals can also be categorized as head (or inflectional) plurals or modifier plurals (Butler 2011, Kim & Meng 2022, Wiltschko 2008). Comparing plurals in English and Halomelem, Wiltschko submits the properties in (9) to distinguish these two types of plurals (listing only the relevant ones), and argues that English has head plurals while Halkomelem has modifier plurals:

(9) Head plurals vs. Modifier plurals	English (head pl)	Halkomelem (mod. pl)
a. Absence of plural marking = singular	yes	no
b. Obligatory agreement	yes	no (= pl is optional)
c. Pluralia tantum	yes	no
d. complementarity of classifiers	yes	no

In this typology, head plurals contrast with singulars, and English is a SG/PL language. Considering the inclusive reading discussed above, I revise it as a SG/PL$_{INC}$ language. Meanwhile, modifier plurals commonly contrast with number neutrality, and optional plural marking brings exclusive plurality: Halkomelem is a Neutral/PL$_{EXC}$ language. Importantly, PN-tati and Redup, although these are inclusive (and therefore they are dividing plurals, i.e., head plurals), pattern with the plurals in Halkomelem. It is noteworthy that the common noun use of PN is still number neutral. As in (4), -tati is optional for a plural interpretation (this is also true in 5). Thus, Japanese, having PN-tati and Redup, is a Neutral/PL$_{INC}$ language, which the categorization in Wiltschko (2008) fails to capture.

Selected References
Borer, H. 2005a. *In Name Only*. Oxford University Press.
Borer, H. 2005b. *The Normal Course of Events*. Oxford University Press.
Butler, L. K. 2011. *The Morphosyntax of Processing of Number Marking in Yucatec Maya*. PhD Thesis. University of Arizona.
Kim, K., & Meng, X. 2022. Variation in the syntax of plural markers: The case from plural -*men* in Mandarin. *Studia Linguistica* 76: 507-551.
Kramer, R. 2016. A split analysis of plurality: Number in Amharic. *Linguistic Inquiry* 47:527–559.
Mathieu, E. 2012. Flavours of division. *Linguistic Inquiry* 43:650–679.
Sudo, Y. 2017. Another problem for alternative-based theories of plurality inferences: The case of reduplicated plural nouns in Japanese. *Snippets* 31:26–28.
Wiltschko, M. 2008. The syntax of non-inflectional plural marking. *Natural Language and Linguistic Theory* 26:639–694.

Is Chinese a Partial Null Subject Language?

Barry C.-Y. Yang

(National United University, Taiwan)

Chinese is generally categorized as a radical pro-drop language, which permits the omission of subjects and objects freely. However, prior investigations have suggested that null subjects in Chinese are subject to certain limitations (Huang 1984, 1989, Huang, Li, and Li 2009, Liu 2014, Yang 2014, Huang & Yang 2024). This talk will demonstrate that the restrictions are attributed to locality effects and the main clause phenomenon (MCP), which prohibits empty subjects in some embedded situations.

Furthermore, an in-depth examination reveals that null subjects corresponding to discourse participants (specifically, the 1st and 2nd person) may surpass those restrictions. This prompts speculation regarding the possibility that Chinese could be considered a language with partial null subject properties. **The restrictions:** According to Huang (1984, 1989) and Huang, Li, and Li (2009), the empty subject pronoun in Chinese obeys the Generalized Control Rule (GCR), which co-indexes the subject *pro* with the nearest nominal element. For instance, in (1), the intervening subject wo 'I' makes the licensing from the topic *Zhangsan* impossible.

(1) *Zhangsan$_i$, <u>wo</u>   du-le   [$_{DP}$ xuduo [$_{CP}$ *pro*$_i$ xie] de shu].
    Zhangsan   I   read-Perf.   many   write DE book
    'Zhangsan$_i$, I have read many books that [he$_i$] wrote.'

Nevertheless, licensing the empty subjects is still limited in some constructions where GCR is irrelevant (Yang 2014, Huang &Yang 2024). In Chinese, an empty topic may license a following null subject. However, in (2), the empty topic [TOP] can serve as the nearest nominal element, whereas the sentence is still incorrect despite having no intervening subjects.

(2) *TOP$_i$ [$_{DP}$ Xuduo [$_{CP}$ *e*$_i$ xie] de   shu] dou hen changxiao.
         many   write DE book all   very sell.well
    'Many books that [he] writes sell well.'

Moreover, Liu (2014) suggested that if a sentence has an empty topic, a preceding overt topic must still be present in a monologue to license the empty topic which, in turn, licenses the null subject. This makes null subject licensing even more stringent. Nevertheless, there are still circumstances where null subject licensing becomes less acceptable even in a monologue. In (5) when *Zhangsan*, serving as a topic, is distant from the empty subject, the sentence turns out to be awkward.

(3) */??<u>Zhangsan</u>$_i$ hen ai tade nüpengyou, TOP$_i$ suoyi [dang *e*$_i$ jieshu na-duan
       Zhangsan very love his girlfriend   so when end that-CL ganqing
       shi], women dou bu xiangxin.
       relationship then we   all   not believe
    'Zhangsan$_i$ loved his girlfriend very much, so when [he$_i$] ended that relationship, we could not believe it.'

MCP and operator intervention: The aforementioned observation is reminiscent of the Root Transformation or MCP (Emonds 1970, 1976, Hooper & Thompson 1973). Building on Haegeman's (2010, 2012a,b) MCP observation that topicalization in adverbial clauses is restricted by operator movement, I propose that the above restrictions are due to (4):

(4) *[$_{CP1}$ [$_{TopP}$ [$_{CP2}$ OP[$_{TopP}$ [$_{IP}$ EMPTYSUBJECT …]]
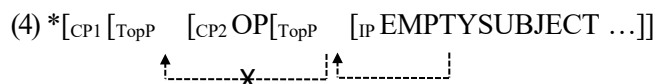
Specifically, I propose that a null subject in Chinese is a D-less ϕP that comprises a set of unvalued ϕ-features that need to be checked (Modesto 2008, cf. Holmberg 2005). When the topic is empty, the subject pro must undergo movement to the matrix topic position, as required by MCP. Such a movement is blocked by an intervening operator (introduced by the adverbial clause), leading to nonconvergence. That is, the stricter case in (3) is due to an intervening operator triggered by the adverbial clause, and it blocks the subject *pro* from moving up to the topic position.

Discourse participants: However, the following example shows that discourse participants (1st person and 2nd person) are not subject to the above constraints. In fact, these sentences can be uttered out of the blue as in (5).

(5) a. $e_{[1st]}$ hao e/ke!
          so   hungry/thirsty
      '[I] am so hungry/thirsty!'
    b. Hao jiu   bu   jian. $e_{[2nd]}$ zhang gao le   ya!
       so   long not see   grow   tall Perf. SFP
       'Long time no see. [You] have grown tall!'

These 1st person and 2nd person empty subjects are not constrained by islands, either, as demonstrated in (6), which contradicts (2).

(6) a. [$_{DP}$[$_{CP}$ $e_{[1st]}$ xiang chi] de   dongxi] name duo,   bu zhidao yao   chi
       shenme. want eat DE thing so   many not know
                    want eat what
       'There are so many things that [I] want to eat and [I] don't know what to eat.'
    b. You [$_{DP}$[$_{CP}$ $e_{[2nd]}$ xihuang] de ren]   le
                   ma? havelike   DE person Perf. Q 'Is there anyone that [you] like?'

Why are the [+participant] features so prominent/strong that they can license the null subject just like an over topic phrase? Hypothesis A: The 1st and 2nd person null subjects are unpronounced strong pronouns. Hypothesis B: A Speaker/Addressee layer is always available on top of every clause, and, therefore, the speaker and the addressee are always available as local antecedents (Sigurðsson 2004, Frascarelli 2007). Evidence from (7) suggests that hypothesis B is favored since otherwise, the 1st and 2nd person null objects would be equally licensed if hypothesis A were employed.

(7) [$_{DP}$[$_{CP}$ <$e_{[1st/2nd]}$> xihuan<*$e_{[1st/2nd]}$>] de ren] bu duo.
               like   DE person not many
    a.   'The persons whom [I/you] like are not many.'
    b. # 'The persons who like [me/you] are not many.'

This concept echoes Ross's (1970) "performative analysis," where a superior structure representing the speaker is proposed above a sentence (see revised versions in Speas and Tenny 2003, Speas 2004, Tenny 2006, Hill 2007a, 2007b, Haegeman 2012a, b, 2014a, b, Hill 2013, Haegeman & Hill 2014, among many others). Such an observation is very interesting since it suggests that Chinese may be categorized as a partial null subject language (Holmberg 2005, 2010; Holmberg & Sheehan 2010), such

as Finnish and Hebrew.

Selected References

Emonds, Joseph. 1970. *Root and structure-preserving transformations*. Doctoral dissertation, MIT.
Haegeman, Liliane. 2010. The internal syntax of adverbial clauses. *Lingua* 120: 628–648.

Huang, C.-T. James. 1984. On the distribution and reference of empty pronouns. *Linguistic inquiry* 15: 531–574. Liu, Chi-Ming. 2014. *A modular theory of radical pro drop*. Doctoral dissertation, Harvard University.

Modesto, Marcelo. 2008. Topic prominence and null subjects. In Theresa Biberauer (ed.), *The limits of syntactic variation*, 375–410. Amsterdam: John Benjamins.

Sigurðsson, Halldór Ármann 2011. Conditions on argument drop. *Linguistic inquiry* 42(2): 267–304.

Speas, Margaret, and Carol Tenny. 2003. Configurational properties of point of view roles. In A.M. Di Sciullo (ed.), *Asymmetry in grammar*, 315–344. Amsterdam: John Benjamins.

Yang, C.-Y. Barry. 2014. Chinese null subjects: A view from the top. In C.-T. James Huang and Feng-hsi Liu (eds.), *Peaches and plums*, [*Language and Linguistics Monograph Series* 54], 227–253. Tapei: Institute of Linguistics, Academia Sinica.

Modeling Schwa Reduction in Directional Harmonic Serialism

Meng-Kai Kyle Ho, Kuo-Chiao Lin
(Kang Chiao International School, Taiwan)

Paiwan (Formosan) has a right-aligned trochee system where stress falls on the penultimate syllable with coda consonants being nonmoraic (1a). However, when the penult's nucleus is a schwa, the schwa undergoes reduction and the stress shifts to the final syllable with vowel lengthening (1b-c) (Shih 2018, 2019; cf. Chen 2009a, 2009b; Yeh 2011). Therefore, the schwa in Piuama Paiwan has three variants on the surface: i) nonmoraic [ə] in the non-head position of a foot, ii) bimoraic [əː] in the head syllable of a foot, and iii) monomoraic [ə] elsewhere (Shih 2018, 2019; cf. Yeh 2017).

(1) *Piuma Paiwan stress* (adapted from Chen 2009a, 2009b; Yeh 2011)

    a. Default stress on penult:

      ['kaka] 'sibling'   ['vuvu] 'grandparents' ['gadu] 'mountain'['ˈtsaviʎ] 'year'

      ['l̩igim] 'needle'   ['tutaŋ] 'aluminum'   [tsaˈʎiŋa]   'ear'   ['piku] 'elbow'

    b. Vowel lengthening in final stressed syllable preceded by a syllable with schwa reduction:

      [kᵊˈriː] 'small'   [qurᵊˈpuːs] 'cloud' [cᵊˈvuːs] 'sugarcane' [qapᵊˈduː] 'gall'

      [kᵊˈmaːn] 'to eat' [kəmᵊˈlaːŋ] 'to know'

    c. Schwa lengthening in final stressed syllable preceded by a syllable with schwa reduction:

      [l̩ᵊˈʎəːt] 'lip'   [ʎisᵊˈqəːs] 'nit'   [tsᵊˈməːl] 'grass'

      [masᵊŋˈsəːŋ] 'to make something'

In this presentation, we will present a python program to model serial directional evaluation of the pattern with revised TROCHEE (2) and IAMB (3).[1] Directional evaluation only evaluates candidates by directionality, i.e., where the violations occur, instead of the number of violated loci (Eisner 2000, 2002; Lamont 2022). Thus, as shown in (4), w.r.t PARSE(σ)⇒ for example, (σ)σσσ is more optimal than σ(σσ)σ and σσσ(σ) more optimal than σ(σσ)σ. Other crucial constraints required for the analysis are stated in (5)-(8).

(2) TROCHEE: Assign one violation for a foot's child that dominates i) a stressed foot-final mora, or ii) an unstressed foot-initial mora.

(3) IAMB: Assign one violation for a foot's child that dominates i) a stressed foot-initial mora, or ii) an unstressed foot-final mora.

(4) Violation vectors and harmonic ordering

    a. Violation vectors and harmonic ordering by PARSE(σ)⇒:

      1111 > 1110 > 1101 > 1100 > 1011 > 1001 > 0111 > 0011

      σσσσ <σσσ(σ) <σσ(σ)σ <σσ(σσ) <σ(σ)σσ <σ(σσ)σ <(σ)σσσ <(σσ)σσ

    b. Violation vectors and harmonic ordering by PARSE(σ)⇐ :

      1111 > 0111 > 1011 > 0011 > 1101 > 1001 > 1110 > 1100

      σσσσ <(σ)σσσ <σ(σ)σσ <(σσ)σσ <σσ(σ)σ <σ(σσ)σ <σσσ(σ) <σσ(σσ)

(5) *STR/ə: Assign one violation to a stressed schwa.

---

[1] Both revised constraints, though penalizing (ˈL$_μ$), permit a heavy syllable to be parsed into its own foot, as a surface (ˈH) may be construed as two distinct representations, (ˈH$_{μμ}$) or (ˈH$_{μμ}$), satisfying respectively (2) and (3).

(6)   *LONG-V: Assign one violation to a long vowel.

(7)   *μ/ə: Assign one violation for a moraic schwa.

(8)   HD(*ft*): Assign one violation for a nonmoraic foot.

As the schwa generally avoids bearing stress, *STR/ə⇒ must outrank TROCHEE⇒ to account for stress assignment on the ultimate, i.e., (9d), rather than on the penult, i.e., (9b-c), in the first iteration of the derivation.[2] The derivation proceeds in the second iteration wherein final stressed syllables undergoes vowel lengthening, i.e., (9f), with violated *LONG-V, ranked lowest. The vowel lengthening is driven by the higher ranked TROCHEE⇒ and IAMB⇒, as both penalize monomoraic feet. Note in particular that this also shows how TROCHEE⇒ and IAMB⇒ step in for the role FTBIN (Prince & Smolensky 1993/2004) usually plays in a non-directional analysis. In the third iteration, harmonic improvement is made via schwa reduction induced by sacrificing the lower ranked *μ/ə in favor of the higher ranked *Stressed/ə⇒, TROCHEE⇒, and IAMB⇒. The derivation then converges in the fourth iteration, for no further harmonic improvement is possible because HD(*ft*) outranks PARSE(σ)⇐ and therefore forbids parsing a syllable containing a nonmoraic schwa into its own foot.

(9) Schwa reduction and final stressed syllable with vowel lengthening, e.g. [[ə̥('ʎə:t)]]

/[ə₁ʎət₂]/	*STR/ə⇒	TROCH⇒	HD(*ft*)⇒	IAMB⇒	PRS(σ)⇐	*μ/ə⇐	*L-V⇐
a. ([ə₁ ʎət₂])	$0_1 1_2$	$W1_1 1_2$		L	L	$1_1 1_2$	
b. ('[ə₁ʎət₂])	$W1_1 0_2$	L		$W1_1 1_2$	L	$1_1 1_2$	
c. ('[ə₁])ʎət₂	$W1_1 0_2$	$W1_1 0_2$		$W1_1 0_2$	$W1_2 0_1$	$1_1 1_2$	
☞ d. [ə₁('ʎət₂)]	$0_1 1_2$	$0_1 1_2$		$0_1 1_2$	$0_2 1_1$	$1_1 1_2$	
e. [ə₁('ʎət₂)]	$0_1 1_2$	$W0_1 1_2$		$W0_1 1_2$	$0_2 1_1$	$1_1 1_2$	L
☞ f. [ə₁('ʎə:t₂)]	$0_1 1_2$				$0_2 1_1$	$1_1 1_2$	$0_1 1_2$
g. [ə̥₁('ʎət₂)]	$0_1 1_2$	$W0_1 1_2$		$W0_1 1_2$	$0_2 1_1$	$L0_1 1_2$	L
h. ('[ə₁])('ʎət₂)	$W1_1 1_2$	$W1_1 1_2$		$W1_1 1_2$	L	$1_1 1_2$	L
i. [ə₁('ʎə:t₂)]	$0_1 1_2$				$0_2 1_1$	$W1_1 1_2$	$0_1 1_2$
☞ j. [ə̥₁('ʎə:t₂)]	$0_1 1_2$				$0_2 1_1$	$0_1 1_2$	$0_1 1_2$
k. ('[ə₁])('ʎə:t₂)	$W1_1 1_2$	$W1_1 0_2$		$W1_1 0_2$	L	$W1_1 1_2$	$0_1 1_2$
☞ l. [ə̥₁('ʎə:t₂)]	$0_1 1_2$				$0_2 1_1$	$0_1 1_2$	$0_1 1_2$
m. ([ə̥₁])('ʎə:t₂)	$0_1 1_2$		$W1_1 0_2$		L	$0_1 1_2$	$0_1 1_2$

The current analysis not only shows that the empirical scope of Lamont's (2022) theory of quantity-insensitive footing can extend to quantity-sensitive phonological adjustments but is more parsimonious than previous ones like Yeh (2017) and Shih (2018, 2019) in dispensing with alignment constraints (McCarthy & Prince 1993) and FTBIN, whose effects are subsumed, respectively, under directionally evaluated TROCHEE/IAMB and PARSE(σ).

Selected References

Lamont, Andrew. 2022. "A Restrictive, Parsimonious Theory of Footing in Directional Harmonic Serialism," *Phonology* 39(1), 41-78.

---

[2] Note that *STR/ə⇒, must also dominate *Stressed/ə⇐ (not shown here), since (9d) wins over (9b-c), because * STR/ə⇒ prefers violations, if any, occur as far from the left edge as possible, whereas * STR/ə⇐ favors the other way around.

Phonological Preparation in Korean:
The Influence of Shared Structures and Units

Hayeun Jang

(Sungkyunkwan University, Korea)

The smallest segment of speech that speakers initially choose during the phonological encoding in speech production is known as the "phonological preparation unit" or "proximate unit". This unit has a big impact on how quickly and efficiently speech is produced. This unit can vary between languages: phoneme in English, Dutch and Vietnamese (Schiller 1998; Roelofs 1999; Verdonschot et al. 2022), syllable in Chinese and Cantonese (Chen et al. 2002; O'seaghdha et al. 2010; Wong et al. 2012), and mora in Japanese (Verdonschot et al. 2011).

The conclusions of recent experimental research on the Korean phonological preparation unit are inconsistent: phoneme (Han & Verdonschot 2019), CV body (Li et al. 2022), or syllable (Verdonschot et al. 2021). The experimental methods could be the reason behind the differing conclusions. Choi et al. (2017) suggest that tasks requiring excessive cognitive processing, such as implicit priming, might make it challenging for speakers to focus on phonological elements, leading to a different preparation unit. Another possible reason is the structural similarity of the experimental word pairs, which may consistently share the same syllable structure in their first syllables, never share it, or vary according to conditions.

In order to better understand the phonological preparation unit of Korean speakers, this study used an object-naming task in a priming paradigm, which is less cognitively demanding. Along the phonological dimension, the prime-target relationship was adjusted to compare all possible conditions, as in (1). The pictures used in the experiment were from a standardized set of black-and-white drawings, MultiPic (Duñabeitia et al. 2018). A total of 288 pairs were made with 54 pictures including heterogeneous pairs with no shared units.

(1) Structure of Experimental Stimuli: Sets with Shared Phonological Units

1st Syllable		Shared Unit	Prime Word	Target Word
Prime	Target			
CV, CVC	CV	C	[pʌ.sʌt] 'mushroom'	[pu.ʌk] 'kitchen'
	CVC		[pan.dʑi] 'ring'	[poŋ.tʰu] 'envelope'
	CV	V	[tʰa.dʑo] 'ostrich'	[sa.ɡwa] 'apple'
	CVC		[mal.bʌl] 'wasp'	[tɕaŋ.mi] 'rose'
	CV	CV	[sa.dʑin] 'photograph'	[sa.sɨm] 'deer'
	CVC		[sam.gak] 'triangle'	[saŋ.ʌ] 'shark'
CVC	CVC	VC	[mɛk.tɕu] 'beer'	[pɛk.tɕo] 'swan'
	CVC	CVC	[koŋ.dʑu] 'princess'	[koŋ.njoŋ] 'dinosaur'

In the first block, participants learned the names of objects by viewing pictures, saying the names, and checking the correct answers. participants viewed every picture, said the names, and checked their answers. In the practice block, they exercised three pairs. In the experimental block, they viewed two sequential pictures: covertly naming the first (prime) and naming aloud the second (target) as quickly and accurately as possible. Each trial consisted of a fixation cross for 200 ms, the prime picture for 300 ms, a white screen for 450 ms, and the target picture for 2000 ms. Reaction time was recorded at the onset of the voice input. In conditions with a significant priming effect indicated by short reaction times, the unit shared by the prime and target can be considered the phonological preparation unit utilized by

Korean speakers in the initial stage of speech production.

In total, 102 Korean speakers participated in the experiment. After data with z-scores for reaction times larger than three and those with multiple incorrect responses were removed, data from 61 participants (37 female, 24 male, average age 23.09y) were analyzed.
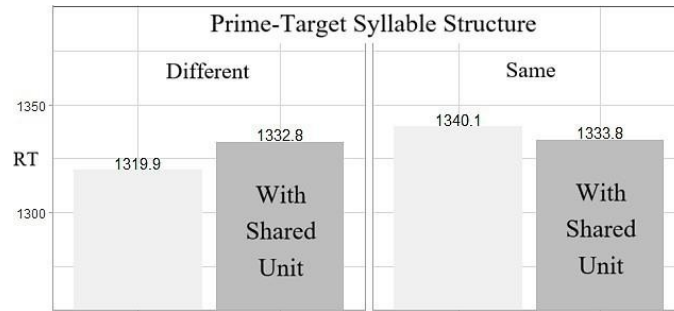

Figure 1. Reaction Time: Existence of Shared Unit by Prime-Target Syllable Structure

Figure 1 illustrates that in the shared unit condition, reaction times (RT) are similar whether the prime and target words have different or same structures for their first syllable. However, the RTs for the no shared unit condition differ significantly, with faster RTs when syllable structures differ. This suggests that interference gets stronger when there is greater structural similarity between the prime and target words. This illustrates that when designing an experiment, consideration must be given to both the cognitive load imposed by the task and the cognitive interference, which varies according to the structural similarity between the prime and the target. The linear mixed model analysis, accounting for random effects by participants, revealed a significant effect of prime-target syllable structure on reaction time (p=0.0478) and a near-significant interaction between prime-target syllable structure and the shared unit (p=0.0969).
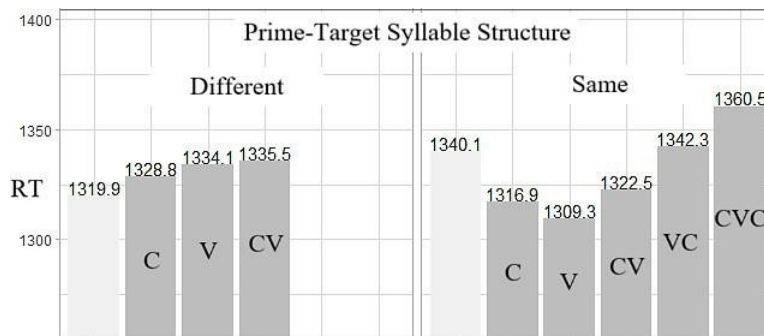
Figure 2. Reaction Time: Type of Shared Unit by Prime-Target Syllable Structure

A closer look reveals that when the target and prime words have the same initial syllable structure, conditions where the target and prime words' CV, C, and V components are the same showed faster reaction times than conditions where the two words do not share any unit, as shown in Figure 2. With random effects by participants, the linear mixed model analysis for the same syllable structure condition showed a significant contrast between the no shared unit and the shared C, V, or CV (p=0.00117). The contrast between the no shared unit and the shared CVC was also significant (p=0.00558). This suggests that in Korean, the CV body structure serves as the phonological preparation unit, and that C and V can function independently within this structure.

Selected References
Han, Jeong-Im, and Rinus G. Verdonschot. 2019. "Spoken-word production in Korean: A non-word

masked priming and phonological Stroop task investigation," *Quarterly Journal of Experimental Psychology* 72(4), 901-912.

Li, Chuchu, Min Wang, Say Young Kim, Donald J. Bolger, and Kelly Wright. 2022. "Phonological preparation in Korean: Phoneme, or syllable or another unit?," *Language and Speech* 65(2), 337-353.

Choi, Tae-Hwan, Sujin Oh, and Jeong-Im Han. 2017. "The influence of task demands on the preparation of spoken word production: Evidence from Korean," *Phonetics and Speech Sciences* 9(4), 1-7.

Verdonschot, Rinus G., Jeong-Im Han, and Sachiko Kinoshita. 2021. "The proximate unit in Korean speech production: Phoneme or syllable?," *Quarterly Journal of Experimental Psychology* 74(1), 187-198.

Trochaic Template Satisfaction in Isbukun Bunun

Sheng-Chieh Huang, Sheng-Chieh Huang, Yan-Bin Huang, Kuo-Chiao Lin,
Zi-Xiang Lin, Ryan Wu (National United University, Taiwan)

This paper presents a streamlined versioned of template satisfaction (TS) in the sense of McCarthy et al. (2012) and uses Isbukun Bunun (IB; Formosan) for demonstration. In IB a single, weight-sensitive trochaic foot is aligned at the right edge of a prosodic word (Huang, 2008; Lin, 2019), including in bimoraic reduplication wherein coda consonants count as nonmoraic (Huang, 2008), while long vowels, offglides, or post-consonantal onglides are heavy and attract stress when in word-final syllables (Huang, 2005; Lin, 2019; Ouyang, 2012). We follow McCarthy et al. (2012) in assuming that a reduplicative affix is a template (Marantz, 1982) that can be a syllable (σ) or foot (*ft*) (McCarthy & Prince, 1986/1996). While an unsatisfied template violates HD(X) (1), the template is satisfied (hence template satisfaction) by an operation COPY(X) included in GEN in *Optimality Theory* (OT; Prince & Smolensky, 1993/2004). The operation COPY(X) violates *COPY(X) (2) by creating a copy of a string of *contiguous* constituents of type X along with their contents and integrates the string into the pre-existing template (McCarthy et al., 2012). Moreover, an operation that copies a string and places the copy adjacent to the original is more faithful than one that places the copy further away, as demanded by COPYLOC(X) (3) (McCarthy et al., 2012). Crucially, we propose to revise TROCHEE and IAMB as (4) and (5), respectively.

(1) HD(X) (Selkirk, 1995): Assign one violation for a foot's child that dominates i) a stressed foot- final mora, or ii) an unstressed foot-initial mora.
(2) *COPY(X) (revised; cf. McCarthy et al., 2012): Assign a violation for every X in an application of COPY(X).
(3) COPYLOC(X) (McCarthy et al., 2012): To a candidate produced by COPY(X), assign as many violations as there are Xs intervening between the original X string and its copy.
(4) TROCHEE: Assign one violation for every foot whose rightmost child is its head, or whose leftmost child is not its head.
(5) IAMB: Assign one violation for every foot whose leftmost child is its head, or whose rightmost child is not its head.

Assuming that ALLFTR (6) is undominated and responsible for the single, right-aligned foot, (7) shows that for the (ˈLL) reduplicative pattern to win out (7a), TROCHEE must outrank both IAMB and *COPY(σ). However, TROCHEE must succumb to HD(*ft*) since the foot template cannot stay unsatisfied (7c). Importantly, ALLFTR and WSP kick in when either of the stem's last two syllables is heavy. The fact that (8a) wins over (8b) reveals WSP's dominance over TROCHEE. Nevertheless, WSP must yield to ALLFTR as exhaustive footing of heavy syllables is ruled out (8d-e). [1] Likewise, (9) shows that undominated ALLFTR would not allow for stressed heavy syllables (9c-e), except for the one as the head of the disyllabic reduplicative trochee, as the optimal candidate (9a).[2]

(1) ALLFTR (McCarthy & Prince, 1993): Every foot stands at the right edge of a prosodic word.

---

[1] Note that (8e) also violates HD(*ft*).
[2] Like (8e), (9e) breaks HD(*ft*). Note also that (9b) violates COPYLOC(σ) as well.

(1) (ˈLL) reduplication: HD(*ft*) >> TROCHEE >> IAMB, *COPY(σ) || COPYLOC(σ)

/na.lu.lu-( )/ 'to pay attention repeatedly'	HD(*ft*)	TROCHEE	IAMB	*COPY(σ)	COPYLOC(σ)
☞ a. na.lu.lu.-(ˈlu.lu)			2	2	
b. na.lu.lu.-(ˈlu)		W1	L1	L1	
c. na.lu.lu-( )	W1	L	L	L	
d. na.lu.lu.-(ˈna.lu.lu)			2	W3	
e. na.lu.lu.-(ˈna.lu)			2	2	W1

(2) (ˈH) reduplication: ALLFTR >> WSP >> TROCHEE >> IAMB

/maχ.sjal-( )/ 'a bit slippery'	ALLFTR	WSP	TROCHEE	IAMB	*COPY(σ)
☞ a. maχ.sjal.-(ˈsjal)		1	1	1	1
b. maχ.sjal.-(ˈmaχ.sjal)		W2	L	W2	W2
c. maχ.sjal.-(maχ.ˈsjal)		1	W2	L	W2
d. maχ.(ˈsjal.)-(ˈsjal)	W1	L	W2	W2	1
e. maχ.(ˈsjal)-( )	W1	L	1	1	L

(3) (ˈHL) reduplication: ALLFTR >> WSP >> TROCHEE >> IAMB

/ma.χaj.ðu-( )/ 'a bit sour'	ALLFTR	WSP	TROCHEE	IAMB	*COPY(σ)
☞ a. ma.χaj.ðu.-(ˈχaj.ðu)		1		2	2
b. ma.χaj.ðu.-(ˈχaj)		1	W1	L1	L1
c. ma.(ˈχaj.)ðu.-(ˈχaj)	W2	L	W2	2	L1
d. ma.(ˈχaj.ðu.)-(ˈχaj)	W1	L	W1	W3	L1
e. ma.(ˈχaj.ðu)-( )	W1	L		2	L

The revised TS analysis has four theoretical implications. First, TROCHEE and IAMB as defined above play crucial roles in determining the ultimate reduplicative shape. Second, both TROCHEE and IAMB may be sacrificed in favor of a reduplicative shape predetermined by the affixed template. Third, FTBIN (Prince & Smolensky, 1993/2004), whose role is subsumed by TROCHEE and IAMB, can be eliminated (cf. Lamont 2022) for more parsimonious TS. Finally, neither weight condition nor templatic allomorph is required to capture the (ˈH) reduplicative pattern.

Selected Reference:

Lin, Hui-shan. 2019. "Quantity Sensitive Foot Reduplication in Isbukun Bunun," *Taiwan Journal of Linguistics* 17(2), 1-32.

A Movement-based Analysis for Non-local Binding:
Chinese *ziji* and Its New Perspectives

Zi-chun Lin

(National Taiwan Normal University, Taiwan)

This paper has two aims: on the one hand, we reexamine some facts about syntax of Chinese *ziji* about blocking effect and long-distance binding (hereforth LD-binding). On the other hand, we discuss a movement-based alternative supplementing for earlier the literature.

Traditionally, generativists establish that anaphors have to be bound in its binding category. At the same time, people have also observed many non-local binding situations, where *ziji* is bound by a binder outside the binding domain. A number of studies (Huang 1984, Tang 1989, Huang & Liu 2001, among others) have been committed to the account for LD-binding. First, Huang (1984) and his following believe that blocking effect holds responsible for the interpretation of LD-binders. However, we found the grammaticality judgements are no longer shared by younger generation of Chinese speakers. Second, another important attempt is the logophoricity person restrictions, positing logophoric binders must be human. See Tang (1989), Choi (2000), Huang & Liu (2001), and Chen (2018). For them, as long as the animacy requirement is finely met, the reflexive pronoun counts as a logophor. However, for the present paper, we argue against this logophor definition as the sole criteria.

To this point, it is clear that the gap to fill is a clear system for LD-binding. In response, this presentation adapts a movement-based analysis for Chinese *ziji*. With reference to Reuland et al. (2020), we consider *zi* to be a reflexivizer and thus *ziji* is a reflexivized pronoun. Also, based on movement theory of binding (Hornstein 2001, Kayne 2005, McKewon 2013, Nishigauchi 2014, Baker & Ikawa 2024, a.o.), we maintain that local *ziji* is derived by movement and non-local *ziji* through "contexts" and "binders within islands". In terms of representation, we proposed a fixed *ziji*-containing NP (i.e., *Zhang-ta-ziji*) consisting of three components: DP (*Zhangsan*), telic (*ta*), and a head noun (*ziji*), of which the structural representation is shown in (1). In (2a), *Zhangsan* moves to be Topic and *ta* "clitic- moves" to attach on and reveal the structural Subject *pro*. This mechanism is quite intuitive, for the Topic-Subject sequence manifests an emphatic reading. In (2b), *Zhangsan* moves to be Subject, and *ta* as a clitic remains in-situ to form NP *taziji*. Crucial to both local-binding cases, Binding Condition A is satisfied interior to the *ziji*-NP because the PRO left by movement of *Zhangsan* may c-command and co-index with *ziji*. Also, Example (2b) is no violation of Binding Condition B, in that *taziji*, but not clitic-*ta*, is c-commanded by and co-indexed with PRO left by movement of *Zhangsan*.

(1)   The syntactic representation of *ziji*-containing NPs:
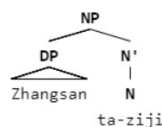
(2)   Local-binding: (moved PRO)

   a.   Zhangsan$_i$ ta   kandao   PRO $_i$ ziji$_i$.
        Zhangsan   he   see   self
        "Zhangsan himself saw himself."

   b.   Zhangsan$_i$ kandao   PRO$_i$ ta-ziji$_i$.
        Zhangsan   see   he-self
        "Zhangsan saw himself."

---

[1] M.A. linguistics student in National Taiwan Normal University. Contact: h89950501@gmail.com

Furthermore, we determine LD-binding cases as in (3). Since Chinese short passive does not have a syntactic logical subject (Ting 1998, Huang et al. 2009), we argue for no movement for *ziji* here. Namely in (3a), the *pro-ziji* sequence needs to retrieve the identity for Spec-NP *pro* from the context, leading to LD-bound *ziji* with no clitic attaching. Likewise in (3b), it is straightforward that neither *ziji* is coreferential with any syntactic entity, and thus no movement should apply, which gives rise to LD-binding. Please note that LD-binding in relation to non-obligatory control (NOC) can help explain how Binding Principle A is observed as in (2) – in that *pro* is derived by NOC and then this null subject c-commands and co-indexes with *ziji*. Lastly in (3c), it is for our argument, yet against Huang (1984)'s blocking effect, that *ziji* can take either a LD-binder 'xueshengmen' or a local binder 'xiaomao' as its antecedent. For us, "locality" has to do with the control phenomena.

(3)   Long-distance binding: (base-generated *pro*)

   a.   Mianbao$_i$ bei   na.lai   *pro*-ziji$_{*i/j}$ chi.
         bread   BEI   take.come   self   eat
         "Bread is taken to eat by oneself."

   b.   *pro*$_i$-Ziji$_i$ xie   [$_{DP}$[*pro*$_i$-ziji$_i$ de] zuoye]. DE
         self   write   self   homework
         "A person oneself does the homework of oneself."

   c.   Xueshengmen$_i$ jiao   [$_{CP}$ xiaomao$_j$ buyao   yao   *pro*-ziji$_{i/j.}$]
         students   ask   small.cat   not   bite   self
         "Students ask the small cat not to bite themselves/itself."

In conclusion, the contribution of our proposal is that we do not heavily rely on logophoricity and binding effect to account for Chinese LD-binding. Rather, the movement PRO and the non-movement *pro* can separate the local and non-local binding situations. Last but not least, the fixed *ziji*-containing DP proposal shown in (1) obeys the Binding Theory elegantly.

Selected references

Baker, M.C., Ikawa, S. 2024. Control theory and the relationship between logophoric pronouns and logophoric uses of anaphors. *Natural Language and Linguistic Theory*.

Huang, C.-T. James, and C.-S. Luther Liu. 2001. Logophoricity, attitudes and ziji at the interface. In *Long distance reflexives* (Syntax and Semantics 33), eds. Peter Cole, Gabriella Hermon, and C.-T. James Huang, 141–195. New York: Academic Press.

Huang, Cheng-Teh James, Yen-Hui Audrey Li, and Yafei Li. 2009. *The Syntax of Chinese*. Cambridge: Cambridge University Press.

Huang, Yun-Hua. 1984. Chinese Reflexives. *Studies in English Literature and Linguistics* 10:163–188. National Taiwan Normal University, Taipei.

Nishigauchi, Taisuke. 2014. Reflexive binding: Awareness and empathy from a syntactic point of view. *Journal of East Asian Linguistics* 23: 157–206.

Reuland, Eric, Sally Chi Ho Wong, and Martin Everaert. 2020. How the Complexity of Mandarin Zi-ji. Simplifies the Grammar. *Linguistic Inquiry* 51 (4): 799–814.

Tang, C.-C. Jane. 1989. Chinese Reflexives. *Natural Language and Linguistic Theory* 7: 93–121. Ting, Jen. 1998. Deriving the Bei-construction in Mandarin Chinese. *Journal of East Asian Linguistics* 7 (4): 319–354.

Dividing Categorial Selection: Towards a Mechanism behind Non-Argument Taking Relation in Syntax

Tatsuya Hashimoto

(University of Tsukuba, Japan)

This study investigates the availability of J-theory in Zhang (2023) and explores some phenomena in coordination and modification construction, say, coordination of unlike categories, Coordinate Structure Constraint (Ross 1967), and Left Branch Condition (Ross 1986). Moreover, it is proposed that category selection needs to be divided into two types, one in Narrow Syntax (NS) and the other in C-I interface. One of the key discussions of recent generative research is the computational difference between forming argument relations and non-argument ones. Zhang (2023) tackles this issue, advocating that the only element that distinguishes the two relations is a functional element she dubs J-element. The J-elements are realized as coordinators and, in some languages such as Mandarin Chinese, modification markers[1] and take complements (e.g., the second conjuncts in coordinate structures and the modifying elements in modification constructions) as other functional heads do. It is worth noting that the resulting SO, J-set, does not have a categorial feature (for a discussion, see Zhang 2023), and then the first conjunct or a modified element merges with it as a categorizer. The overall structure can be roughly represented as follows:

(1) [[$_{XP}$ XP (Conj1 / modified element)] [$_{J\text{-set}}$ J (Coordinator/modification marker) [$_{YP}$ YP (Conj2 / modifier)]]

The J-theory leads us to the theoretically desired consequence, especially for coordinate structures: there is no need for construction-specific assumptions (e.g., Boolean Phrase (Munn 1992, 1993), multi-dominance analysis (de Vries 2005), and sister movement (Chomsky 2013)) since the configuration in (1) is pervasive in coordinate structures and modification structures. Then again, J-theory holds that the categorizing elements are the only elements that are c-selected, which poses a serious empirical issue. This is what we focus on below.

   Bruening (2023) illustrates that coordinate structures mismatching categories are allowed as long as all conjuncts in a coordinate structure obey a selectional requirement from a selector (e.g., V and P) (though a few types of apparent exceptions exist).

(2)   a.   Danny became a political radical and very antisocial.
       b. * Danny became a political radical and under suspicion.   (Bruening and Al Khalaf 2020:1)

Note that the verb *become* subcategorizes for NPs and APs, not PPs. Therefore, leaving his detailed proposal aside, (2a) is grammatical since the two conjuncts are correctly selected by the verb, while (2b) is not because the PP in the second conjunct does not satisfy the selectional requirement. However, if the second conjuncts are excluded from c-selection in J-theory, how can we account for the contrast we have just seen above?[2] Zhang 2023 does not seem to answer the question.

---

[1] Na-ge   pangpang-**de**   nahai   you   lai-le.
   that-CL   fat-**modification marker**   boy   again   come-
   PRF 'That fat boy came again.'

[2] Bruening and Al Khalaf (2020) make a similar claim to Zhang's (2010) line of analysis.

The fact that the c-selectional relation of selectors is established not only with a categorizing element but also with a complement of a J-element indicates that we have two c-selectional strategies, namely one in NS and the other in C-I interface. That is, non-argument elements introduced by J-elements are subject to c-selection in C-I interface (shading represents a checking domain).

(3) a.   NS [[selector] [$_{XP}$ XP] [$_{J\text{-set}}$ Conj [$_{YP}$ YP]]]
     b.   C-I [[selector] [$_{XP}$ XP] [$_{J\text{-set}}$ Conj [$_{YP}$ YP]]]

Considering (2a), each conjuncts obeys selectional requirements in NS and C-I, respectively. On the contrary, in (2b), the selector fails to select the second conjunct in C-I since, as we have just seen, the verb does not subcategorize for PPs. This reasoning that there is another type of c-selection is supported by the well-known fact that modifying elements such as adjectives and adverbs are segregated based on what they can modify: adjectives can modify only nouns while adverbs modify categories other than nouns. It is natural to assume that there can be another type of selectional relation other than one in NS.[3]

(4) a.   NS [[ZP (selector)] [$_{XP}$ XP] [$_{J\text{-set}}$ modification marker [$_{YP}$ YP (modifier)]]]
     b.   C-I [[ZP] [$_{XP}$ XP (selector)] [$_{J\text{-set}}$ modification marker [$_{YP}$ YP (modifier)]]]

In modification constructions, the c-selection between selectors and modified elements is established (e.g., John likes a kind girl.) in NS. Regarding the c-selection in C-I, the modified element is to be a selector for adjectives as modifiers (e.g., John likes a kind girl.).

Dividing c-selection into the two types has a further consequence for Coordinate Structure Constraint, which bans extraction of conjuncts or elements contained in conjuncts. Oda 2017, 2021 reports that in some languages (Japanese, Korean, etc.) conjuncts can be moved out of a coordinate structure. In addition, he notices that such an extraction is allowed only from the first conjunct.

(5) a. ? Kyodai$_i$-to   kanojo-wa [t_i Toodai]-ni   akogareteiru.
        Kyoto.University-and she-$_{TOP}$ Tokyo.University-$_{DAT}$ admire
        'She admires Kyoto University and Tokyo University.'   (Oda 2021:606)
     b. * Todai-ni$_i$ kanojo-wa [Kyodai-to t_i]   akogareteiru.
        Tokyo.University-$_{DAT}$ she-$_{TOP}$ Kyoto.University-and   admire
        'She admires Kyoto University and Tokyo University.'   (Oda 2021:627)

To the extent that the proposal here is on the right track, the prohibition of extracting the second conjunct is explained in the following way: the complement of the J-element (i.e., the second conjunct) needs to be in a c-selectional relation with the selector (here, akogareteiru 'admire') in C-I interface (see also fn.3). If this element is extracted to the surface position in NS, this checking will not properly be conducted in C-I interface because the selection is a local relation between the selector and the complement of the J-element. In other words, the selector cannot find the second conjunct in its search domain.

The current proposal can also give an account to the following contrast, which shows a well-known constraint as Left Branch Condition (LBC).

(6) a.   How tall a man did Sheila marry?
     b. * How tall did Sheila marry a man?   (Ross 1986:130)

---

[3] The type of selection in C-I interface may be affected by the choice of J-element. We leave this open for future research.

In a modification construction as in (4), a modified element and a modifier cannot be separated until they are transferred to C-I because of the c-selectional relation. Therefore, (6b) is ungrammatical since the adjective *how tall* is extracted from the noun *man* in NS. That is, it is naturally followed that LBC is, in fact, a reflection of the selectional requirement of the modified elements in C-I interface.

Selected References
Bruening, Benjamin and Eman Al Khalaf. 2020. "Category Mismatches in Coordination Revisited," *Linguistic Inquiry* 51(1), 1-36.
Oda, Hiromune. 2021. "Decomposing and Deducing the Coordinate Structure Constraint," *The Linguistic Review* 38(4), 605-644.
Ross, John R. 1967. *Constraints on Variables in Syntax*, PhD thesis, MIT, Cambridge.
Zhang, Ning. 2023. *Coordinate Structures*, Elements in Generative Syntax, Cambridge University Press, Cambridge.

Three-Way Classification of ACC-Marked External Possession in Korean and Its Implications

Min-Joo Kim

(Texas Tech University, USA)

Typological research shows that human language has a way of encoding external possession, in addition to internal possession (e.g., Payne and Barshi 1999). In an internal possession construction (IPC), the possessor (PR) occurs inside the maximal nominal projection containing the possessum (PM), typically bearing genitive (e.g., (1)). In an external possession construction (EPC), the PR occurs *outside* the maximal nominal projection containing the PM, bearing a core grammatical relation to the verb, but still semantically connected to the PM (e.g., (2)). The existence of EPCs in human language raises the question of why they may exist in addition to IPCs. This paper is an attempt to help answer this question by looking at the accusative (ACC) marked EPC in Korean, illustrated in (2).

(1) Cwuni-ka	**Mina-uy**	(*kyosil-eyse)	**phal-ul**	cap-ass-ta.	(IPC)
C.-NOM	M.-GEN	classroom-LOC	arm-ACC	grab-PST-DECL	

'Cwuni grabbed Mina's arm (in the classroom).'

(2) Cwuni-ka	**Mina-lul**	(kyosil-eyse)	**phal-ul**	cap-ass-ta.	(EPC)
C.-NOM	M.-ACC	classroom-LOC	arm-ACC	grab-PST-DECL	

'Cwuni grabbed Mina by the arm (in the classroom).'

As comparison between (1) and (2) shows, in the ACC-marked EPC in Korean, the PR and the PM form separate nominal projections but they occur bearing the same case, namely, ACC, so this construction is treated as instantiating the multiple object or multiple accusative construction in the literature (see, e.g., Yoon 2001; Park 2013). Korean also has two additional EPCs, where the PR and the PM occur bearing nominative or dative cases (see, e.g., Kim 1990; Maling and Kim 1992; Kim 2000; Ryu 2014; Yoon 2015). Yet the ACC-marked EPC has received the most attention because exactly what licenses it is still poorly understood. One predominant view has been that in the ACC-marked EPC, the PR and the PM must stand in an *inalienable relation* (e.g., Yoon 1997). But such a view cannot capture data like (3) where the PR and the PM do not stand in an inalienable relation. Another influential view has been that the ACC-marked EPC in Korean is subject to what is called the *affectedness condition* (e.g., Yoon 1989; Kim 1989; Tomioka and Sim 2005, 2007), as has been held for numerous other languages (e.g., French, Spanish, German, Flemish, Hebrew, Mandarin, Swahili). However, postulating the workings of this condition alone cannot capture the full range of facts, since in sentences like (4), the PR is not construed as an affected theme (Yoon 1997; Yoon 2015; Shin 2022). Authors like Tomioka and Sim (2007) argue that data like (4) do not instantiate the EPC. But, in such sentences, too, the two ACC-marked nominals stand in a PR-PM relation; moreover, they have IP counterparts. Therefore, offering an analysis that captures data like (4), as well as data like (2) and (3), will be more desirable; filling this gap in the current literature is a more specific goal of the present paper.

(3) Cwuni-ka	**Mina-lul**	**catongcha-lul**	kocangnay-ss-ta.
C.-NOM	M.-ACC	car-ACC	break-PST-DECL

'Cwuni broke Mina$_i$'s car on her$_i$.'

(4) John-i   **ku   ai-lul**   **ilum-ul**   al-n-ta.
       J.-NOM   that kid-ACC   name-ACC   know-IMPRF-DECL
       'John knows that kid's name.'   (Adapted from Shin 2022: ex. (22))

In this paper, I suggest that no previous analysis has been able to successfully capture the full range of data because, upon closer inspection, sentences like (2), (3), and (4) turn out to behave differently, despite the fact that they all fit the description that has been used to characterize EPCs in Korean. In short, in all grammatical cases, both the PR and the PM bear ACC; PR > PM is maintained both on the surface and the hierarchical levels; and all acceptable EP sentences have an IP counterpart, modulo a meaning difference between them, which may have to do with affectedness, physical contiguity, and/or pragmatic salience (see Shin 2022 for relevant discussion). When one applies various syntactic/semantic diagnostics as I have done here, however, it turns out that there are at least three types of ACC-marked EPC in Korean, which are respectively exemplified by (2), (4), and (3):

Table 1. Three-Way Classification of the ACC-Marked EPC in Korean: A New Analysis

	Type I (e.g., (2))	**Type II** (e.g., (4))	**Type III** (e.g., (3))
Verb Meaning	[+Physical]	[-Physical]	[+/-Physical]
Productivity	Yes	Not really	No
Answering 'Where' vs. 'What' Q	'Where'	'Where'/'What'	'What'
Omittability of PM	Yes	Yes(?)	No
Depictive Modification	PR	N/A	PM
Case on PM in a Passive	ACC/NOM	NOM	NOM
Fronting of PR Alone	Yes	Yes	Yes
Fronting of PR-PM Pair	Yes	Yes?	Yes??/No
Relativization of Subject	Yes	Yes?	Yes??/No
PR as a Reflexive Pronoun	Yes	Yes/No	No
Compatible with Short Negation	Yes	Yes/No	No
Affected PR	Yes	No	Yes
Inalienable PM	Yes	Yes/No	Yes/No
Part-Whole Relation	Yes	Yes/No	No

Based on the differences between the three types of ACC-marked EP sentences summarized above, I argue that in Type I and Type II cases the PR is the true object of the verb but in Type III cases the PM is the true object, and that in the former two cases, the PM is an adjunct which specifies which part of the PR the meaning of the verb holds true of, and in the latter case, the PR is an applied argument which does not participate in the event under description but which is the *locus* of the resultant state of the event. The present analysis correctly predicts that Type I and Type II cases do not behave identically because only in the former environments do the PR and the PM stand in a mereological relation in the physical domain. Additionally, on the present analysis, the affectedness in Type I cases and that in Type III cases are not of the same kind; in the latter, affectedness is less direct because the actual object of the verb is the PM, not the PR. If correct, the proposed analysis helps settle some of the longstanding issues surrounding the ACC-marked EPC in Korean, in particular which between the PR and the PM is the true argument of the verb (see, e.g., Yoon 1989, 2001; Maling and Kim 1992) and whether all occurrences of PMs are modifiers of some kind (see, e.g., Kim 1989; O'Grady 1998, 2002; Shin 2022).

Processing Patterns of Korean Adverbial Sentences with '-(eu)ro': Comparing Fixed and Free Word Orders

Eunjin Gye, Wooyeol Lim, Youngjoo Kim
(Kyung Hee University, Korea)

This study examines the processing of Korean adverbial case marker '(eu)ro' sentences among native Korean speakers and advanced learners using a self-paced reading task. Korean adverbial sentences are characterized by the relatively free positioning of components in a 'subject-object-adverb-verb' word order sentence. However, some adverbs have fixed positions within sentences. According to studies such as Lee(1989), Nam & Kim(1993) and Choi(1995), certain uses of the adverbial case marker '(eu)ro' do not permit positional swapping.

(1)   a. I knew him as a genius. (나는 그를 천재로 알았다.)
      b. I knew as a genius him. (나는 천재로 그를 알았다.)

Native Korean speakers feel (1a) is grammatically correct whereas (1b) is not(Lee, 1989). Sentence processing related to word order has been discussed in various languages. Studies such as Tomaoka et al.(2014) and Dröge et al.(2016) on Japanese and German word order revealed that sentences with altered word order increased processing complexity and significantly slowed processing speed. Recently, Baik & Lee et al.(2023) examined the processing of Korean nominative and accusative sentences, while Baik & Gye et al.(2023) examined the processing of Korean dative sentences. They found that processing speed decreased, and error rates increased for scrambled word orders, similar to findings in other languages. The analysis of response tendencies to comprehension questions confirmed a syntactic hierarchy of "meaning > word order > marker" for both groups. In addition, Baik et al.(2024) also examined the processing of Korean adverberbial sentences with '(eu)ro' and cognitive verbs. They found that the native Korean speakers had trouble with the sentence processing when the unmarked word order was scrambled, but the learners were not.

27 native Korean speakers and 27 advanced Korean learners participated in this study. The sentences for this study were selected based on prior research by Nam & Kim(1993), Byun(2002), Yoo(2004), Lim(2007), focusing on uses of the adverbial case marker '(eu)ro' with low flexibility in word order such as 'cognition' and high flexibility in word order such as 'destination'. Verbs were chosen from studies by Hong et al.(1998), supplemented by *Korean word frequency dictionaries* by Kang & Kim(2009) and the *International common Korean standard curriculum vocabularies*(2017), resulting in four verbs: 'know(알다)', 'consider(여기다)', 'take(가져가다)' and 'send(보내다)'. Using these verbs, sentences were created in six word order conditions (SOAV, SAOV, OSAV, OASV, ASOV, AOSV) and eight case marker omission conditions (no, nominal, accusative, adverbial, nominal & accusative, nominal & adverbial, accusative & adverbial and all case marker omission), generating a total of 96 sentences. The experiment was conducted using E-prime 3.0 for the self-paced reading task.

The results were as follows: First, native speakers, unlike advanced learners, found it more difficult to process non-canonical word orders ('adverb-object') compared to canonical word orders ('object-adverb'). Second, both groups found it more challenging to process non-canonical sentences of the

'cognitive' than those of the 'destination'. Third, advanced learners did not utilize case marker information as effectively as native speakers.

As a follow-up to the studies by Baik & Lee et al.(2023), Baik & Gye et al.(2023) and Baik et al.(2024), this study aims to reveal whether Korean adverbial sentences with '(eu)ro' exhibit similar processing patterns to basic and dative sentences. This research is significant as it comprehensively examines the sentence processing differences between native Korean speakers and L2 learners using the self-paced reading task.

Selected References

Baik, J., Gye, E. J., Lee, S. Y., Jeong, H. J. & Kim, Y. J.(2023), Sentence processing of native and non-native Korean speakers on dative sentences, *The Korean language and literature 205*, The study of Korean language and literature, pp.67-105.

Baik, J., Lee, S. Y., Jeong, H. J. & Kim, Y. J.(2023), The effects of word order on sentence processing of native and non-native Korean speakers: a comprehensive comparison on syntactic, semantic, and morphological factors, *Journal of Korean language Education 34(1)*, The International Association For Korean Language Education, pp.75-105.

Baik, J., Jeong, H. J., Lee, S. Y., Nam, Y. J. & Kim, Y. J.(2024), Processing Patterns in Korean Adverbial Sentences : A Comparative Analysis of Native and Non-Native Speakers of Korean, *Journal of Korean Language and Literature 207*, The society of Korean Language and Literature, pp.147- 173.

Byun, J. M.(2002), The category of cognitive verb, *Korean linguistics 16*, The association for Korean linguistics, pp.307-332.

Choi, H. C.(1995), About Korean complements, *Korean linguistics 2*, The association for Korean linguistics, pp.447-490.

Dröge, A., Fleischer, J., Schlesewsky, M. & Bornkessel-Schlesewsky, I. (2016), "Neural mechanisms of sentence comprehension based on predictive processes and decision certainty: Electrophysiological evidence from non-canonical linearizations in a flexible word order language", *Brain Research 1633*, pp.149-166.

Hong, J. S. et al.(1997), *Modern Korean Verb Lexicon dictionary*, Doosandonga.

Kang, B. M. & Kim, H. G.(2009), *Korean word frequency dictionaries*, Seoul: Hangukmunhwasa.

Lee, K. K.(1989), Scrambling and Acceptability in Korean, *Language research 25(1)*, Language education institute, pp.141-150.

Lim, H. B.(2007), Language typological approach to word order and basic word order of Korean, *Sogang Humanities Research 22*, Humanities Research Institute Sogang University, pp.53-120.

Nam, K. S. & Jo, E.(1993), A collocation study on the usage of the marker '-ro' (II), *The dong bang hak chi 76*, Institute of Korean Studies, pp.231-252.

Tamaoka, K., Asano, M., Miyaoka, Y., & Yokosawa, K.(2014), Pre-and post-head processing for single- and double-scrambled sentences of a head-final language as measured by the eye tracking metho, *Journal of psycholinguistic research 43(2)*, pp.167-185.

Yoo, H. W.(2004), The study of the transitive verb with the obligatory argument 'N-ro', *Korean linguistics 24*, The association for Korean linguistics, pp.189-220.

Japanese Academic Apology:
A Case of English as a Lingua Franca Encounter

Yeonkwon Jung

(Kansai Gaidai University, Japan)

This study theorizes and analyzes an aspect of Japanese politeness at the level of asymmetrical interactions between individuals. It explores perceptions of the need for an apology and what motivates undergraduates to apologize in the Japanese academic encounters (Spencer-Oatey and Kadar 2021). It is about assessments regarding the behavior that triggers the potential need for an apology in a certain English as a lingua franca (ELF) encounter. Data of the study are composed of authentic ELF email correspondence written by the Japanese undergraduates to inform the course instructor know their absence of the class. This study investigates whether class absence "is" the behavior that triggers the potential need for an apology.

This study exemplifies whether there is straightforward link between culture, contextual factors, and participants' reactive assessments. Given the claim, it tries to illustrate whether apology is inherently reactive politeness. This illustration expects to address whether apologies are for committing offence and/or whether apologies link to high level of awareness or mindfulness of offensiveness. This study to a lesser extent discusses whether reactive assessments relate to participants' perceptions of the degree of severity of a given offence (If there are, which type of apology; Even in case of awareness of offensiveness, does a situation outside the individual's control make no performance of apology? (e.g. overwhelming reason(s))).

On the other hands, this study exemplifies whether cultural factors can play a significant role in participants (L2 learners)' perception of offence and the need to apologize even in ELF encounters. The learners' first language sociocultural expectations might influence their L2 pragmatic performance. This study tries to address whether the choice of English is indeed a marked behavior among local Japanese in the performance of apology (In ELF encounter, does pragmatic transfer occur in sensitive moments such as apologies? Do the Japanese undergraduates observe the Japanese sociocultural values there?). This study explicates whether Japanese undergraduate students follow a Japanese social norm, *tachiba* (social roles and positions of participants), in giving such unwelcome news in English. If they do, this study illustrates that the use of English apology is also role-bound (Ide, 1998).

Major findings of the study illustrate that the norms of apology in Japan need to be considered as contextually contingent and discursively enacted. Although sincerity is often mentioned as an important element of an apology (Jung 2022), the realization of apology seems no link to perceived weight of the offence in the notification of class absence. This may be supported by the simple semantic components of apologies in the study. Namely, the present apology is not heavily loaded with upgraders. Even many of data make no performance of apology. In this respect, this study argues that there is no straightforward link between culture, contextual factors, and participants' reactive assessments. In this way, besides culture-specific (emic) perspective, culture-general (etic) perspective can also be considered in the study of Japanese apology.

Selected References

Jung, Y. 2022. *Korean Business Communication: A Comprehensive Introduction*. Routledge.

Ide, R. 1998. "Sorry for your kindness": Japanese interactional ritual in public discourse. *Journal of Pragmatics*, 29/5, pp.509-29.

Spencer-Oatey, H. and Kadar, D. 2021. *Intercultural Politeness: Managing Relations across Cultures*. Cambridge University Press.

Artificial Intelligence-Assisted Learning in High-Functioning Autistic Children for Processing Modern Chinese Proverbs

Ben Yue Shi Yang

(National Chung Cheng University, Taiwan)

The aim of this study is to investigate the potential of ChatGPT 4.0 to assist high functioning autistic (HFA) children in learning and processing modern Chinese proverbs. The research utilizes data from the Modern Chinese Proverbs Dictionary to evaluate the perceptions and comprehension abilities of both HFA children and ChatGPT 4.0. The study involves six HFA children from grades three to six and the ChatGPT 4.0 language model. Participants were tasked with completing the Metaphorical Proverbial Story Test, which is designed to assess their understanding of idiomatic expressions in context.

The Metaphorical Proverbial Story Test consists of twenty-four sets of materials, each including a short story of approximately 150 words. The proverbs or metaphors are categorized into three types: those with a human subject, those with an animal subject, and those related to daily life scenarios. This categorization helps in understanding the specific areas where HFA children may face more difficulties. Table one shows the examples of the Metaphorical Proverbial Story Test. For instance, the metaphorical proverb "熱鍋上的螞蟻" (re guo shang de ma yi) involves an animal subject and signifies a person being nervous or anxious, rather than the literal scenario of ants on a hot pot. Such metaphors are chosen to test whether the participants can grasp the non-literal meanings embedded in everyday language. After each story, a comprehension question related to the metaphor is provided. Participants are asked to explain what the metaphorical proverb means.

ChatGPT 4.0's role in this study is to provide clear, concise explanations of these proverbs and metaphors in English, making their non-literal meanings more accessible to HFA children. By contextualizing the idioms with simple, relatable examples and offering interactive, repetitive practice, ChatGPT 4.0 aims to cater to the specific learning needs of HFA children.

This study explores how artificial intelligence tools like ChatGPT 4.0 can bridge the comprehension gap observed in idiomatic expressions. The ultimate goal is to enhance the understanding of the linguistic challenges faced by HFA children and improve educational approaches. This research aims to provide insights into the effectiveness of using AI-based tools in educational settings, particularly for children with unique learning needs, and to suggest ways to integrate such technologies into traditional teaching methods.

Keywords: Metaphorical Proverbial Story Test, High-functioning autistic children, Artificial intelligence assisted learning, Learning outcomes, Modern Chinese proverbs
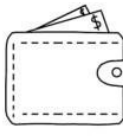
Table 1: The Metaphorical Proverbial Story Test

Question in Chinese with Chinese Pinyin	Question translated into English
小明暑假學了 26 個英文字母，開學後找老師挑戰英文單字。同學笑說他是在關公面前耍大刀。 請問: 在關公面前耍大刀是什麼意思? Xiao ming shujia xue le ershiliou ge yingwen zimu, kaizue hou zhao laoshi tiaozhan yingwen danzi. tongxue xiaoshuo ta shi zai guangong mianqian shua da dau. Qingwen: Zai guangong mianqian shua da dau shi sheme yisi?	During the summer vacation, Xiaoming learned the 26 letters of the English alphabet and challenged the teacher with new English words after school started. His classmates laughed, saying he was "brandishing a big sword in front of Guan Yu." Question: What does it mean by "brandishing a big sword in front of Guan Yu?"

王先生是有錢人，但是他卻很吝嗇。因此,大家都說王先生是鐵做的公雞。 請問:鐵做的公雞是什麼意思? Wang xiansheng shi youqianren, danshi ta que hen linse. Yinci, dajia dou shuo wang xiansheng shi tie zuo de gongji. Qingwen: Tie zuo de gongji shi sheme yisi?	Mr. Wang is a wealthy man, but he is very stingy. Therefore, everyone says Mr. Wang is "a rooster made of iron." Question: What does it mean by "a rooster made of iron?"

Table 2: Examples of ChatGPT 4.0 assisted learning of metaphorical outcomes

Explanation in Chinese with Chinese Pinyin	Explanation in English
這個說法「鐵做的公雞」暗示一個人非常不願意花錢，像是一隻鐵製的公雞，因其僵硬和不易變動的特性，連一根羽毛都捨不得脫落。 Zhege shoufa tiezuo de gongji anshi yige ren feichang buyuanyi huaqian, xiangshi yizhi tiezhide gongji, yinqi jiangying han buyibiandong de texing, lain yigen yumao dou shebede tuoluo.	The expression "a rooster made of iron" evokes the idea of a rooster made of iron, which cannot even shed a feather due to its rigid and inflexible nature. It implies that the person is so unwilling to spend money.
「在關公面前耍大刀」比喻在專家或行家面前炫耀自己的技能，結果顯得自不量力和愚蠢。這個成語源自三國時期名將關羽，以高超的武藝著稱，因此在他面前炫耀刀法顯得很可笑。 Zài Guān Gōng miànqián shuǎ dàdāo" shì yī jù chéngyǔ, bǐyù zài zhuānjiā huò hángjiā miànqián xuànyào zìjǐ de jìnéng, jiéguǒ xiǎnde zì bù liànglì hé yúchǔn. Zhè gè chéngyǔ yuánzì Sānguó shíqī míng jiàng Guān Yǔ, yǐ gāochāo de wǔyì zhùchēng, yóuqí shàncháng shǐyòng dàdāo. Yīncǐ, zài tā miànqián xuànyào dāofǎ xiǎnde hěn kěxiào.	To brandish a broadsword in front of Guan Yu" is an idiom that means showing off one's skills in front of an expert, making oneself appear foolish and overconfident. This idiom originates from the Three Kingdoms period, where the general Guan Yu was renowned for his exceptional martial arts skills, especially with the broadsword. Therefore, showing off one's swordsmanship in front of him would be ridiculous.

Table 3. Metaphorical Picture Book

32

Selected References

Chahboun S, Vulchanov V, Saldaña D, Eshuis H, Vulchanova M (2016). Can You Play with Fire and Not Hurt Yourself? A Comparative Study in Figurative Language Comprehension between Individuals with and without Autism Spectrum Disorder. *PLoS ONE*,11(12): e0168571. doi:10.1371/journal.pone.0168571

Qin Zheng, Zhongheng Jia, Dandan Liang(2015). Metaphor and metonymy comprehension in Chinese-speaking children with high-functioning autism. J*ournal of Psychology research in Autism Spectrum Disorder*, 10(1), 51-58.

Tsai, L. Y., & Scott-Miller, D. (1988). Higher-Functioning Autistic Disorder. *Focus on Autistic Behavior*, 2(6), 1-8. https://doi.org/10.1177/108835768800200601

Textual Patterns in *The Rambler*: A Preliminary Stylistic Study of Samuel Johnson and Canonical Authors

Masaaki Ogura

(Osaka Metropolitan University, Japan)

This study examines which authors, among those whom Samuel Johnson (1709-1784) regarded as ideal, have writing styles that resemble Johnson's in his periodical, *The Rambler* (1750-1752). Samuel Johnson, an 18th-century writer and lexicographer, aimed to fix the English language through his *Dictionary of the English Language* (1755). He initially aimed to fix English by drawing illustrative examples from the works of esteemed authors, as indicated in his *Plan* (1747). However, he later acknowledged the difficulty of fixing the language in the *Preface to the Dictionary* (see Kolb & DeMaria, 2005).

Despite this, Johnson continued to strive for grammatical purity and elegance. In *The Rambler* No. 208, he still expressed his desire to refine English. While scholars such as Wimsatt (1941), McIntosh (1998), and Damrosch (2019) have explored stylistic aspects of Johnson's language, the specific ideal language use he pursued in *The Rambler* remains to be researched. Specifically, research on his language use compared to the esteemed writers he frequently cited has not been conducted, thus leading to the research question stated at the beginning.

The methodology involved compiling corpora that included *The Rambler* and the Wells of English Undefiled Corpus (WEUC), which the author created. This corpus comprises works by frequently cited authors revealed by Stenberg (1994), including John Milton, Robert South, Philip Sidney, Thomas Browne, John Locke, Francis Bacon, John Dryden, Roger L'Estrange, William Shakespeare, Matthew Prior, and Edmund Spenser[i]. In addition, the *King James Bible* (*KJV*) was included as a corpus[ii]. Using AntConc,[iii] 100 p-frames (phrase frames), which are recurring patterns of words, were extracted in descending order of their range. The frequency of these p-frames was calculated per 100,000 words. The p-frame frequency data from The Rambler and the other corpora were converted into vectors, and cosine similarity was used to measure stylistic similarity. Values closer to one indicate higher similarity.

Figure 1 The cosine similarity indicating similarity between texts

---
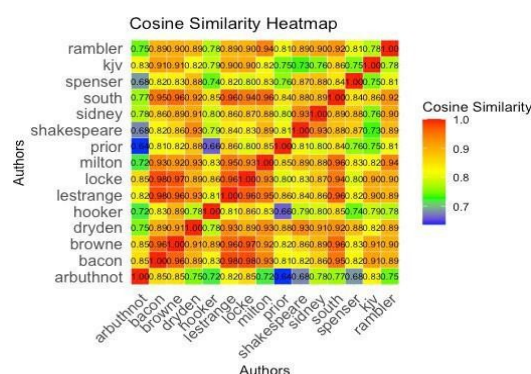
[i] The word count for each author is as follows: Arbuthnot (11,909), Bacon (1,475,454), Browne (454,920), Dryden (1,261,211), Hooker (692,851), L'Estrange (1,699,834), Locke (1,668,936), Milton (908,452), Prior (78,740), Shakespeare (1,539,846), Sidney (339,042), South (570,830), Spenser (273,521).

[ii] The data was obtained from Project Gutenberg, and the total number of words is 821,119.

[iii] Anthony, L. (2024). AntConc (Version 4.3.0) [Computer Software]. Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software

Figure 1 represents the cosine similarity values between *The Rambler* and the other authors' works. It indicates that the p-frame usage trends in *The Rambler* are closest to those in the texts of John Milton, Robert South, and Philip Sidney, with cosine similarity values exceeding 0.9 for these authors' texts. Additionally, the heatmap shows other similarities among authors, such as between South and Milton, and South and Browne. Note that these results are based on the analysis of 100 p-frames. Future research should investigate the impact of varying the number of p-frames on similarity values. By increasing the number of p-frames analyzed, it would be possible to determine whether the observed stylistic similarities hold across a broader range of p-frame features. Additionally, the effect of using n-grams instead of p-frames to measure stylistic similarity should be considered. Comparing p-frames with n-grams may provide insights into the extent to which topics or content influence the results.

The findings of this study could contribute to *Dictionary* research and the issue of authorship attribution. Comparing the language in *The Rambler* with the examples used in Johnson's Dictionary can reveal if the modifications made to quotations match the stylistic tendencies found in *The Rambler*. For example, examining whether Johnson altered quotations to better fit his stylistic ideals could provide a picture of his editorial policy. Additionally, the study could address the issue of authorship attribution, such as whether Samuel Johnson wrote parts of Charlotte Lennox's *The Female Quixote* (1752), as discussed by Small (1935/1969). This study is significant because, despite Johnson's distinctive style, often labeled as 'Johnsonese,' having been studied, no quantitative analysis had explored the stylistic similarities between Johnson and the authors he considered canonical until now.

Selected References

Bate, W. J., & Strauss, A. B. (Eds.). 1969. *The Rambler* (The Yale edition of the works of Samuel Johnson, Vols. 3–5). New Haven: Yale University Press.

Damrosch, L. 2019. *The club*. New Haven: Yale University Press.

Kolb, G. J., & DeMaria, R. (Eds.). 2005. *Johnson on the English language* (The Yale edition of the works of Samuel Johnson, Vol. 18). New Haven: Yale University Press.

McIntosh, C. 1998. *The evolution of English prose, 1700–1800: Style, politeness, and print culture*. Cambridge: Cambridge University Press.

Small, M. R. 1969. *Charlotte Ramsay Lennox: An eighteenth century lady of letters*.

Hamden: Archon Books. (Original work published 1935)

Stenberg, T. 1944. Quotations from Pope in Johnson's Dictionary. *Studies in English* 24, 197– 210.

Wimsatt, W. K. 1963. *The prose style of Samuel Johnson*. New Haven: Yale University Press. (Original work published 1941).

The Use of the Pragmatic Marker *Like* by Algerian Speakers of English in Ireland: A Corpus-Based Discourse Analytic Approach

Salima Araar

(University of Limerick, Ireland)

Pragmatic Marker (PM) *Like*, amongst other PMs, is considered a key linguistic feature of the speech of native speakers of English fulfilling a variety of communicative functions (Aijmer 2004; Ament *et al*. 2020; Almuways 2022). PM *Like* is seen to be indicative of socio-pragmatic competence in native speakers' (NS) English (Yong-Ping 2003; Aijmer 2013). Accordingly, non-native speakers (NNS) of English also need to acquire and use PMs as part of their communicative competence development in English speaking settings (Aijmer 2002). There exists a growing body of corpus-based research in the area of PMs generally and PM *Like* in particular in NS English (Fraser 1990, 1999; Andersen 2001; Brinton 2017; Aijmer 2002, 2004, 2013; Beeching 2016) but far less research has been undertaken in relation to the use of PMs including PM *Like* in NNS English (Santos 2019; Diskin 2017). Therefore, the present study investigates the use of PM *Like* by Algerian speakers of English in conversational talk. A 60,000-words corpus of conversational English produced by 20 Algerian speakers of English living and studying in Ireland was compiled and thoroughly analysed through a corpus-based discourse analytic approach.

The quantitative results have revealed that PM *Like* was substantially used by the Algerian participants. These results were compared to NS use of this marker from American, British, and Irish backgrounds. Comparisons indicated that PM *Like* was used more significantly by the Algerian speakers compared to the NSs from the different background as indicated in (1).

(1)

	Algerian use	American use	British use	Irish use
Percentage of PM *Like*	84%	63%	65%	69%

The functional analysis of PM *Like* by the Algerian participants explored both the range of functions for which the marker was being used and the extent of use of each function observed. Based on the functional classification system adapted from Andersen (2001) and Farrell (2004), an overall range of seven pragmatic functions for PM *Like* was identified in the corpus of the Algerian participants, that is the use of PM *Like* as i) a *Focuser*, ii) a *Discourse Link*, iii) a *Hesitational Filler*, iv) an *Approximator*, v) an *Exemplifier*, vi) a *Quotative* and vii) as a *Hedging Device*. Examples from the Algerian corpus are illustrated in (2):

(2) **Focuser**: *I prefer to study at home because I feel **like** more relaxed*
     Discourse Link: *I don't have sisters or brothers **like** I'm an only child*
     Hesitational Filler: *We had **like** umm the break of the winter*
     Approximator: *I got to visit New York so it was **like** ten years ago*
     Exemplifier: *This is **like** a family dinner*
     Quotative: *she was **like** wow what's going on?*
     Hedging Device: *Is that true **like**?*

Several key findings were revealed in the present study. First, PM *Like* was overly used by the Algerian participants compared to its grammatical use as well as in comparison to NS use. In this regard, as non- native speakers of the language, the Algerian participants are thereby expected to use PM *Like* more frequently to overcome their limited command over the language. They use PM *Like*

as a tactic to fill in hesitations and avoid stuttering and reluctance (Romero-Trillo 2017; Santos 2019). Moreover, PM *Like* was used by the Algerian speakers saliently and for a variety of functions which indicates that this group of NNSs have developed the sophisticated pragmatic competence to use the marker appropriately to navigate social exchanges. Furthermore, the analysis indicated the influence of proficiency level on the extent to which PM *Like* was used by the Algeria participants, in that advanced speakers showed more use than the upper-intermediate participants. In addition to proficiency level, other variables have influenced the use of PM *Like* by the Algerian participants such as: i) the experience of living abroad which increased exposure to natural language input, ii) the influence of TV shows and social media, with the advent of native English speaking platforms and contents, iii) the influence of age, given that these participants belong to a young age group (twenty-something), they are expected to overly use this marker due to its popularity among their age group (Andersen 2001).

The discourse analysis (DA) approach drew on Relevance Theory (Sperber and Wilson 1986) and Politeness Theory (Brown and Levinson 1987) to investigate how the Algerian participants used PM *Like* to express relevance and politeness in their conversational English. The DA analysis indicated that the Algerian participants have reached native-like use of the marker to express relevance and therefore their pragmatic competence had developed in this particular area.

Selected References

Aijmer, Karin. 2002. *English Discourse Particles*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Aijmer, Karin. 2004. "Pragmatic Markers in Spoken Interlanguage." *Nordic Journal of English Studies* 3 (1): 173-190.

Aijmer, Karin. 2013. *Understanding Pragmatic Markers: A Variational Pragmatic Approach*. Edinburgh: Edinburgh University Press.

Almuways, Yasir. 2022. "A Corpus-Based Comparative Pragmatic Analysis of Irish English and Canadian English." Ph.D. diss., University of Limerick.

Ament, Jennifer, Pérez-Vidal, Carmen, and Parés, Júlia Barón. 2020. "A Study on the Functional Uses of Textual Pragmatic Markers by Native Speakers and English-Medium Instruction Learners." *Journal of Pragmatics* 156: 41-53.

Andersen, Gisle. 2001. *Pragmatic Markers and Sociolinguistic Variation: A Relevance Theoretic Approach to the Language of Adolescents*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Beeching, Kate. 2016. *Pragmatic Markers in British English: Meaning in Social Interaction*. Cambridge: Cambridge University Press.

Brinton, Laurel J. 2017. *The Evolution of Pragmatic Markers in English: Pathways of Change*. Cambridge: Cambridge University Press.

Brown, Penelope, and Stephen C. Levinson. 1987. *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.

Diskin, Chloé. 2017. "The use of the discourse-pragmatic marker 'like' by native and non-native speakers of English in Ireland." *Journal of Pragmatics* 12: 144-157

Farrell, Angela. 2004. *The Use of the Pragmatic Marker 'Like' by Irish Teenagers: Is It a Girl Thing?* M.A. thesis, University of Limerick.

Fraser, Bruce. 1990. "An Approach to Discourse Markers." *Journal of Pragmatics* 14: 383-395. Fraser, Bruce. 1999. "What Are Discourse Markers?" *Journal of Pragmatics* 31: 931-952.

Romero-Trillo, Jesús. 2017. "Corpus Pragmatics." *CrossMark* 1: 1-2. doi:10.1007/s41701-017-0005-z. Santos, Giovani. 2019. "Second Language Pragmatics: A Corpus-Based Study of the Pragmatic Marker Like". *ResearchGate*, 12(4): 1-16.

Sperber, Dan, and Deirdre Wilson. 1986. *Relevance: Communication and Cognition*, 2nd ed. Oxford: Blackwell.

Yong-Ping, Ren. 2003. "The Pragmatic Functions of Discourse Marker 'Well'." *Journal of Foreign Languages*, no. 3: 58-64.

Research Trends in *Linguistic Research* from 2014 to 2023: A Bibliometric Analysis

Ziyun Dai

(Korea Maritime and Ocean University, Korea)

Bibliometrics, a term coined by Pritchard (1969: 348), refers to 'the application of mathematical and statistical methods' to the analysis of scientific publications. There have been many bibliometric studies of science or social science as a general field, and there have also been some discipline-specific studies, such as those of computer science and economics. However, these earlier studies focused exclusively on examining and comparing research productivity and impact.

Meanwhile, for academic journals, all of them have their stated goals and scopes, i.e., the specific research issues (themes/topics) and types of research methodologies they each cover. It is important to know whether and how the journal successfully achieves its goals and stays within its scope, not only for its publishers (editors and editorial boards) but also for its potential readers and researchers. Such information helps professionals in the former group to fulfill the journal's mission more effectively, while it enables individuals in the latter group to make more informed decisions about whether to read and/or submit research to the journal.

Linguistic Research is a Scopus-indexed international journal that provides a forum for the discussion of theoretical research on natural language data. The journal publishes high- quality articles that make a clear contribution to the current debate in all branches of theoretical and applied linguistics. The journal embraces both synchronic and diachronic perspectives, and publishes articles that address language-specific as well as cross-linguistic and typological research questions. In light of this fact, this study aims to do so by analyzing the publications in the journal *Linguistic Research* over the last 10 years.

For the purpose of this study, information about the articles in the journal *Linguistic Research* from 2014 to May 2024 was downloaded from Scopus. Only research articles were chosen within the research scope. Two sm corpora are created by compiling article abstracts from various periods in time for the purpose of topic analysis and research.

Table 1. Number of articles by period and number of abstracts available

Types of item	Periods		Total number	Abstracts
	2014-2018	2019-2024		
Research articles	138	140	278	278

After part-of-speech tagging of two small corpora using R language, lemmatization is performed in this study. The monogram is then screened by topic according to POS (where only the noun class is selected as the topic of study). Then, 2-5grams are generated and n-grams are screened out except for those containing stop words. Based on the number of selected articles and the size of the corpus, only n-grams that are used more than 10 times are considered research topics. For the frequency of topic occurrence in different periods, this study adopts standardized frequency for comparative analysis.

$$\text{normalized frequency} = \frac{ttttttttttttttttttt\ rrrrrrrrrrr\ ffffrrrrffffffffffffffffffffffftttttffff}{\begin{array}{c}ttttttttttttrrrrttttt\ fffffffffffffffffffffffrrrr\ tttttffff\\rrrrrrrrrttttttttttttttttttffffrrrr\ tttttffff\ ffffrrrrrtttttth\\tttttffffffrrrrttttttttttttttt\end{array}} *100$$

The results show that in the past five years, English and Korean language teaching research has been a hot topic in journals. However, the interest of Korean language research is increasing, while the interest in English language research is decreasing. At the same time, the study of linguistics and metaphor has gradually increased in recent years.

Table 2. Comparison of research topics in *Linguistic Research* in each period

2014-2018		2019-2024	
N-grams	Normalized frequency	N-grams	Normalized frequency
Korean	117.39	Korean	140.00
English	76.09	English	71.43
ellipsis	17.39	ellipsis	16.43
vowel	14.49	vowel	42.86
movement	14.49	movement	20.00
L2 learner	11.59	L2 learner	7.14
speech rate	8.70	speech rate	0.71
wh-phrase	7.25	wh-phrase	3.57
collocation	9.42	collocation	0.71
island	8.70	island	4.29
motivation	7.97	motivation	1.42
textbook	7.97	textbook	3.57
native Korean listener	0.72	native Korean listener	7.14
gender	4.35	gender	7.86
wh-question	0.29	wh-question	7.86
assimilation	5.80	assimilation	7.14

Selected References

Lei, Lei and Dilin Liu. 2019. "Research Trends in Applied Linguistics from 2005 to 2016: A Bibliometric Analysis and Its Implications," *Applied Linguistics* 40(3), 540-561.

Liu, Xiaodong, Qi Xu and Meina Li. 2015. "A Comparative Analysis of Scientific Publications in Management Journals by Authors from Mainland China, Hong Kong, Taiwan, and Macau: 2003-2012," *Scientometrics* 105, 135–143.

Liu, Yanhua and Guangwei Hu. 2021. "Mapping the Field of English for Specific Purposes (1980-2018): A Co-citation Analysis," *English for Specific Purposes* 61(1), 97-116.

Pritchard, Alan. 1969. "Statistical Bibliography or Bibliometrics?" *Journal of Documentation* 25(4), 348–349.

Dummy or Not? Exploring the Referentiality of the English Pronoun 'It'

Rebecca Kirkle, Eugene Chung
(Korea University, Korea)

The English pronoun *it* can be divided into two main usage categories, referred to in the present study as anaphoric *it* and non-anaphoric *it.* Anaphoric *it* displays co-referential ties to an anaphor, or other entity named in the discourse.

(1) She said she was going to pack **her bag** and leave **it** in the entrance. [Gardelle, 2011]

Meanwhile, non-anaphoric *it* regularly occurs without linguistic co-reference in several contexts, including as the object of certain verbal expressions.

(2) a. From Aberdeen to Edinburgh we trained **it** [...] [Gardelle, 2011]
     b. She said she was going to pack her bag and beat **it** back to Tennessee. [Gardelle, 2011]
     c. It's really kind of hard to tell the government to go shove **it**. [Mondorf, 2016]

While the reference of anaphoric *it* in (1) can be syntactically tied back to *her bag*, the non-anaphoric *it* in
(2) cannot be tied back to *Aberdeen* or *Edinburgh* in (2a), *her bag* in (2b), or *the government* in (2c), nor does it appear to have any definable meaning of its own. Much of the literature on non-anaphoric object *it* therefore refers to this *it* as a dummy, operating under the assumption of one unified construction which encompasses all of the verbal expressions in which this type of *it* is commonly found. However, Kaltenböck (2002) shows that this type of *it* can be divided into three separate usage contexts, each with a differing level of reference.

In its first use, non-anaphoric *it* can be seen to function as the 'pseudo-object' of various intransitive verbs like *go* and weakly lexicalized denominals/deadjectivals like *train* (2a) (Mondorf, 2016). In this context, *it* appears to take on a low, context-dependent level of referentiality, as evidenced by its ability to alternate with a limited number of lexical items with vague, context-dependent meanings, like the noun *things*.

In its second use, non-anaphoric *it* can be combined with certain transitive verbs to form a small number of phrasemes with highly lexicalized meanings, as in *beat it* (2b). Much like a particle occurring as part of a phrasal verb, *it* does not seem to have any real level of reference or identifiable meaning of its own in these phrasemes, and this is evidenced by the fact that this type of *it* cannot easily be replaced by other lexical items or noun phrases, even very general ones like *things*.

In its third and final use, non-anaphoric *it* appears as the object of a number of transitive verbs like *shove* (2c) in a way that is suggestive of collocation. These verbs typically permit relatively free alternation between non-anaphoric *it* and a variety of both lexical items ranging from fairly general to highly specific. This suggests that speakers have a wide array of linguistics options to choose from if asked to specify what they mean by *it,* implying that non-anaphoric *it* may in fact maintain a definable level of referentiality when occurring with such verbs.

This leads to the question of whether previous studies have acted too broadly in categorizing the *it* of this third category as a non-referential dummy. To date, no quantitative evidence has been presented to justify such a position. The present study therefore attempts to confirm whether this type of *it* does in fact lack a measurable level of referentiality by using psycholinguistic methodology to probe the mental representation of *it* as it used in each of the three contexts outlined above. To this end, we report the results of a self-paced reading experiment in which 150 native English speakers were asked to read

sentences containing 15 common 'verb + non-anaphoric' *it* expressions, five from each context. Each sentence was constructed in three conditions which varied according to whether *it* was preceded by a singular noun with the potential for co-referential interpretation (3a), a plural noun with the potential for co-referential interpretation (3b), or no noun with the potential for co-referential interpretation (3c).

(3) a. We boarded **the taxi** and cabbed **it** to the airport.
   b. We boarded **the taxis** and cabbed **it** to the airport.
   c. We gathered ourselves together and cabbed **it** to the airport.

The reading times for *it* in each condition were then compared using ANOVA. In the case that *it* was assigned a referential interpretation, it was expected that reading times would be significantly slowed in conditions (b) and (c) as compared to condition (a), presumably due to processing difficulties caused by participants' failure to locate a grammatically matching (i.e., singular) noun with the potential to resolve the co-reference of the pronoun. By contrast, there would be no need for such resolution in the case of a non-referential interpretation, meaning that *it* should be processed in a similar manner regardless of the preceding content. Therefore, it was expected that a non-referential interpretation of *it* would result in a similarity in reading times across conditions.

The results of the experiment, which showed no significant differences in reading times for *it* across conditions for any of the 'verb + non-anaphoric *it*' expressions under study, suggest a preference for a non- referential or dummy representation of *it* regardless of usage context. However, while no obvious differences were found, there may be more subtle or fine-grained differences in processing or representation that were not captured here. Therefore, we highly encourage the continued use of quantitative methodologies in further attempts to shape a more robust psycholinguistic profile of the non-anaphoric *it* phenomenon.

Selected References

Gardelle, L. (2011). *Whoop her up, hit it, go it alone*: The role of the personal pronoun in the fossilization process. In K. Kuiper (Ed.), *Yearbook of phraseology*. Mouton De Gruyter.

Kaltenböck, G. (2002). That's It? On the unanticipated 'controversy' over anticipatory It: A reply to Aimo Seppänen. *English Studies 83*, 541–550.

Mondorf, B. (2016). "Snake legs it to freedom": Dummy it as pseudo-object. *Corpus Linguistics and Linguistic Theory 12*(1), 73–102. https://doi.org/10.1515/cllt-2015-0071

The Role of Addressee in Interpreting Predicate Doubling in Japanese

Yuki Ishihara

(Tokyo Institute of Technology, Japan)

The roles of Speaker and Addressee have been examined in syntactic research since Ross (1970), with proposals elaborating syntactic structures above CPs (Speas and Tenny 2003, Miyagawa 2022, among others). The present paper examines a colloquial syntactic construction in Japanese involving the doubling of predicates and supports the syntacticization of discourse based on its addressee-oriented interpretations.

Ishihara (2019) observes that predicate doubling emphasizes either the doubled element or a constituent containing it, depending on the aspectual types of predicates and contextual factors. For example, iteration of stative or activity predicates emphasizes the degree of the state/activity, as in (1a). When a predicate is doubled in an answer to a negative question, the truth of the proposition is focused (verum focus), as in (1bB).

(1)   a. Ah, <u>atsu-i</u>   <u>atsu-i</u>.   Moo   tae-rare-na-i.
          oh hot-NPST hot-NPST any more   bear-can-NEG-NPST   'Oh, it's very hot. I cannot stand
          it any more.'
      b. A: Taroo-wa   mada kite-na-i-yo-ne?
              Taro-TOP yet   come-NEG-NPST-SFP- SFP 'Taro hasn't come yet, has he?'
          B: Uun, <u>ki-ta</u>   <u>ki-ta</u>.
              no   come-PST come-PST 'No. He HAS come.'

However, not all instances of predicate doubling yield an emphatic interpretation of the content of propositions.

(2) a. A: Waka-tta?
            understand-PST
           'Did you understand it?'
        B: (<u>Haihai,</u>)   <u>waka-tta-(yo)</u>   <u>waka-tta-(yo)</u>. yes yes understand-PST-SFP
                         understand-PST- SFP '(Yes, yes,) I got it.'
    b. A: Hontooni gomen-ne. really sorry-SFP 'I'm really sorry.'
        B: <u>Ii-{kara/yo/*ne}</u>   <u>ii-{kara/yo/*ne}</u>.
            OK-because/SFP/SFP   OK-because/SFP/SFP 'It's OK. Never mind.'

In (2aB, 2bB), though a verum focus reading is available, the dominant reading does not emphasize the truth of propositions. Instead, Speaker B uses iteration to indicate successful reception of A's message. This interpretation is addressee-oriented because it focuses on whether the message conveyed by the addressee (i.e. the previous speaker) was understood, rather than emphasizing the content of the message itself. In certain contexts like (2aB), this can evoke a sense of annoyance.

Ishihara claims that an emphasis feature resides in a Speech Act projection above CP, triggering head movement of a predicate, and that pronunciation of a copy left behind by movement yields predicate doubling (Nunes 2004, Martins 2007). To accommodate both addressee-oriented interpretation and verum focus interpretation, I propose modifying Ishihara's analysis, following Miyagawa (2022), who posits Speaker Phrase, Addressee Phrase and Commitment Phrase above CP. I argue that when an emphasis feature (EMP_1 in (3)) is carried by the Commitment head, commitment to the proposition is amplified, leading to a verum focus interpretation. Conversely, when an emphasis feature (EMP_2) is

carried by the Addressee head, predicate doubling is interpreted as addressee-oriented.

(3) [$_{SpkP}$ Speaker [[$_{AddrP}$ Addressee [[$_{CommitP}$ CP Commit-[EMP$_1$]] addr-[EMP$_2$]]] spk]]

Miyagawa (2022:105) contends that *yo*, a Sentence Final Particle (SFP) amplifying commitment to the proposition, is merged onto the Commitment head, while *ne*, an SFP that directs expression to the addressee to ascertain the truthfulness of the proposition, is merged onto the Addressee head.

(4) [$_{SpkP}$ Speaker [[$_{AddrP}$ Addressee [[$_{CommitP}$ CP Commit-*yo*] addr-*ne*]] spk]]

If the emphasis feature in the Commitment head triggers head movement of a predicate and blocks occurrences of other elements within the head, it is predicted that a verum focus reading is not available with iterated *yo* or *ne*. This prediction is supported by Ishihara's observations; SFPs cannot be iterated along with a predicate easily in cases like (1b'B).

(1b') A: Taroo-wa   mada kite-na-i-yo-ne?
       B: Uun, <u>ki-ta-(??yo/*ne)</u>   <u>ki-ta-(??yo/*ne)</u>. no   come-PST-SFP come-PST-SFP 'No.
            He HAS come.'

On the other hand, SFP *yo* can be iterated under the addressee-oriented interpretation, as indicated in (2aB, 2bB), because EMP$_2$ is located higher than *yo* in the structure, and a predicate picks up *yo* on the way to its attracting head, Addressee. Note that doubling of *ne* is not available even under the addressee-oriented reading, because it occupies the same position as EMP$_2$.

Doubling of categories other than predicates is beyond the scope of Ishihara's study, but as shown in (2aB), interjections can be doubled as well. Since doubling of *hai* 'yes' is often interpreted as annoyance rather than emphasis of agreement in response to questions like "Do you understand?" just like doubling of predicates under the address-oriented reading, I propose an analysis similar to the one for predicate doubling. An emphasis feature occurs in the Addressee head, triggering movement of *hai* with the same feature to Addressee, and both the tail and the head of the movement chain are pronounced. To derive the word order in (2aB), we postulate two conjoined Speaker Phrases with EMP in each Addressee heads.

The doubling of other syntactic categories, such as nouns (5a) and adverbs (5b), can be explained similarly.

(5) a. Hora,   neko   neko.
        hey   cat   cat
        'Hey, (look at) a cat.'
     b. Ah,   soo   soo.
        oh   so   so
        'Oh, that is so. I agree.'

In (5a), a noun *neko* 'cat' is iterated to catch the addressee's attention rather than emphasize the fact that it is a cat. The utterance of (5b) is often accompanied by nods, indicating agreement with the addressee. These examples demonstrate that iteration emphasizes interactions between a speaker and an addressee, resulting in varied pragmatic effects.

In conclusion, this study sheds light on the addressee-oriented interpretations of syntactic doubling and supports Miyagawa's analysis of articulated syntactic structure of the expressive component.

Selected References

Ishihara, Yuki. 2019. *Syntactic Doubling of Predicates in Japanese*, Doctoral dissertation, the
        University of Tokyo.
Miyagawa, Shigeru. 2022. *Syntax in the Treetops*, Cambridge, MA: MIT Press.
    Nunes, Jairo. 2004. *Linearization of Chains and Sideward Movement*, Cambridge, MA: MIT Press.

Semantic Case Markers in Korean Noun Phrases
: Focusing on *-ey tayhan* and *-ey uyhan*

Hojeong Koo
(Sogang University, Korea)

In this paper, I present case (and/or voice) alternation phenomena in Korean noun phrases, where morphological cases exhibit a constrained distribution depending on semantic characteristics of the predicates (e.g., Aktionsarten; Vendler, 1975). In turn, I suggest that there is a semantic Case marker in Korean, which is conceptually close to structural Case (Woolford, 2006). This will be presented in detail with contrasting data. It is known that there is typically no direct relationship between the accusative case marking and the semantic, especially aspectual, properties of the verb (Butt, 2006; Kagan, 2020). However, in Finnish, aspectual properties of the verbs affect the realization of Case (Kagan, 2020:104-106):

(1) Group 1 (inherently *atelic* verbs): partitive object Rakastan   sinua   /*sinut. love.1.SG
              you.PART
              /*you.ACC 'I love you.'
(2) Group 2 (inherently *telic* verbs): accusative object Tapoin karhun   /*karhua. killed.1.SG
              bear.ACC   /*bear.PART
(3) Group 3 (both *telic* and *atelic* event predicates): partitive and accusative object
    a. Join   vettä. drank.1.SG water.PART
       'I drank some water / I was drinking water.'
    b. Join   veden. drank.1.SG water.ACC 'I drank the water.'

The similar pattern is found in Korean, especially in noun phrases, as shown below.

(4) a. *Activities*
    pyenhoin-uy phikoin-ey tayhan pyenlon attorney- GEN defendant-PART
                              argument
    'The attorney's argument on behalf of the defendant.'
    b. *Accomplishments*
    *mokswu-uy uyca-ey tayhan   ceycak carpenter- GEN   chair-PART
    production 'The carpenter's production of the chair.'
    c. *Achievements*
    *senswu-uy   khosu-ey tayhan   wancwu athlete-GEN course-PART
                 completion 'The athlete's completion of the course.'
    d. *States*
    pumo-uy   casik-ey tayhan   kekceng
    parents-GEN children-PART   worry 'Parents' worry about their children.'

Interestingly, if the event represented by the head noun has a culmination (or *telic*) or has no temporal duration (b, c), then the logical object of the predicative noun cannot be marked by *-ey tayhan*. Further considerations might be needed to generalize this observation.

(5) *Minswu-uy heli-ey tayhan wuntong Minsu-GEN waist-PART exercise 'Minsu's waist exerise.'

The data in (5) may not be the problem, as some activity events can be modified by temporal

adverbials (e.g., *in an hour*), indicating that they can have a logical culmination. Still, the difference between (4a) and (5) should be explained. This points out that the affectedness of the object (or a *theme*) is also affecting the realization of *-ey tayhan* (Tenny, 1987; Lee, 2005). Like *-ey tayhan*, *-ey uyhan* also shows alternating relationship with the structural case marker *-uy* (genitive marker in Korean), resembling transitivity alternation in English (for convenience, *-ey uyhan* is potentially assumed as an ergative marker). The following examples illustrate that the subject of the intransitive noun receives the same marking as the object of the transitive counterpart (6, 8). Conversely, if the object is elided, then the subject should be marked by *-ey uyhan*. This differs from the sentential subject being marked as the nominative case, implying causation of the event.

(6) kenmul-uy   kensul building-GEN construction
          'The construction of the building.'
(7) [%]Minswu-uy   kenmul-uy   kensul[1] Minsu-GEN   building-GEN construction 'Minsu's
          construction of the building.'
(8) Minswu-ey uyhan (kenmul-uy)   kensul Minsu-ERG   (building-GEN)
                       construction 'The construction of the building by Minsu.'

Selected References

Butt, M. 2006. *Theories of Case*. Cambridge University Press.

Chomsky, N. 1970. "Remarks on Nominalization," In *Readings in English Transformational Grammar*, R. A. Jacobs and P. S. Rosenbaum, eds., Ginn & Co., Waltham, Mass. Kagan, O. 2020. *The Semantics of Case*. Cambridge University Press.

Lee, S. 2005. *A Study on Argument Structure of Noun*. Worin Publishing Co. Tenny, C. 1987. "The Aspectual Interface Hypothesis," *NELS* 18: 490-508. Vendler, Z. 1957. "Verbs and Times," *The Philosophical Review* 66: 143- 160.

Woolford, E. 2006. "Lexical Case, Inherent Case, and Argument Structure," *Linguistic Inquiry* 37: 111- 130.

---

[1] There is some controversy over the grammaticality of (8). However, compared to the NPs where the *-ey tayhan* can be attached to the object noun, (8) is quite acceptable.

ChatGPT and Google Bard for Critical-PBLL in
Korean University English Education

Mi Kyong Kim
(Chodang University, Korea)

In alignment with the increasing importance of critical thinking skills and the transformative potential of text- based generative AI as educational tools within Project-Based Language Learning, this study aims to develop and explore a CriticalPBLL framework utilizing ChatGPT and Google Bard. This encompasses a comprehensive understanding of students' English learning experiences by implementing this framework over a six-week period, embedded into a college English class. The study involved a total of 20 students divided into four groups. Data from five sources were analyzed, including 16 post-questionnaires, 16 reflection notes, 4 sets of group journals, 4 sets of group Social Network Service (SNS) communications, and 13 interviews. The findings of the study revealed that students had positive English learning experiences in various aspects: Constructing knowledge related to project topics and the English language; promoting reflective thinking; engaging in playful learning; and recognizing the teacher's role as a facilitator. On the other hand, students also reported negative learning experiences: Overreliance on ChatGPT and Google Bard, leading to reduced discussions; initial learning anxiety due to unfamiliarity with prompt literacy and critical thinking-based tasks; and technical limitations associated with ChatGPT. Some pedagogical implications include the collaborative development of prompt literacy guidelines with learners; the active utilization of English outputs generated by ChatGPT and Google Bard as valuable language learning resources within language learning models; and the collaborative establishment of guidelines with learners for the appropriate use of ChatGPT and Google Bard.

Positioning Analysis of AI-Based Medicine in Published Expert Interviews: A Preliminary Analysis

Ki-tae Kim

(Keimyung University, Korea)

Health communication involving artificial intelligence (AI) involves an emerging multidisciplinary discourse, on which little has been explored. Attempting to fill this gap in research, the present study investigates published interviews with medical and technical experts involved in IBM's discontinued AI-based healthcare service, Watson Health. Using Positioning Theory (Davies & Harré 1990/1999) as a framework, while building upon the researcher's research on conceptual metaphors on AI-based healthcare service (Kim 2022), the study identifies different positions and their modes. Due to the lack of ethnographic data after the unfortunate closure of Watson Health in early 2022, the study qualitatively analyzes approximately 80 articles, press releases, and other sources containing interviews with experts who were (then) currently or formerly affiliated with Watson Health.

The findings highlight how experts positioned themselves and Watson Health simultaneously, within the same discursive space, across multiple discourse levels (with a focus on the institutional level), and fluidly over time. Experts also exhibited agency in selectively adopting positions albeit within the confines of their moral obligations (especially, institutional and professional ones).

The study also explores three predominant modes (van Langenhove & Harré 1999) through which positioning was implemented in the data: third-order positioning in interviews (that is, outside of unfolding consultations with patients), accountive positioning in interviews (as opposed to performative positioning in some parts of unfolding consultations), and moral positioning in institutional and societal settings (as opposed to personal positioning, which may emerge in unfolding consultations).

The implications for the future development and utilization of generative AI-based healthcare are also discussed.

Selected References

Davies, Bronwyn and Rom Harré. 1990(1999). "Positioning and Personhood," in Harré, Rom and Luk van Langenhove (eds.), *Positioning Theory: Moral Contexts of Intentional Action*, 32-52. Oxford, UK: Blackwell.

Kim, Ki-tae. 2022. "What Are You, Watson?: Conceptual Metaphors on Artificial Intelligence-based Medicine," *Ene 'Korean Journal of Linguistics'* 47.2, 229-260.

Van Langenhove, Luk and Rom Harré. 1999. "Introducing Positioning Theory," in Harré, Rom and Luk van Lagenhove (eds.), *Positioning Theory: Moral Contexts of Intentional Action*, 14-31. Oxford, UK: Blackwell.

Word-order Reflexes of Information Structure under the Influence of Bilingualism: A Pilot Study on Mandarin-English Bilingual Speakers

Xiaotian Wang

(Harvard University, USA)

Many previous studies have contributed to the discussion of how information structure, more specifically givenness, is formally expressed. Different hypotheses, such as the prosodic hypothesis and the word-order hypothesis, have been proposed regarding the expression of givenness. The main research question this paper aims to tackle is: Does bilingualism shift the word-order reflexes of information structure for Mandarin speakers? This question is especially meaningful because even though previous works have compared different hypotheses, modeled distinctive factors, and analyzed languages from a wide range of language families, the impact of bilingualism has not yet been taken into account. By looking at the acceptability of the word-order hypothesis, namely the given-before- new ordering, of bilingual speakers of Mandarin and English, this study aims to shed light on the cross- linguistic influence on word-order reflexes. The two languages were chosen because Mandarin has a more flexible and free word order compared to English.

Previous literature has focused on the comparison between these two hypotheses. Most & Saltz (1979) conclude that prosody, more specifically stress, is a better predictor of focused/new information than word order. Other studies, which look at one of the hypotheses, managed to achieve corresponding results. Nooteboom & Kruyt (1987) focus on the prosodic hypothesis in Dutch, and their results correspond well with the prosodic hypothesis. However, the cross-linguistic study conducted by Skopeteas and Fanselow (2010) claims that the givenness of an argument will lead to its clause-initial placement. Skopeteas and Fanselow (2010) certainly observed support for the connection between word order and givenness, but prosody was not controlled in their experiment, thus creating a possible limitation to their study (Šimík & Wierzba, 2017). The interaction between givenness and prosody (the prosodic hypothesis) received more support than the word order hypothesis in previous literature.

20 monolingual speakers of Mandarin and 15 bilingual speakers of Mandarin and English were recruited for the pilot trial. They were asked to read through four dialogues written in Mandarin, each consisting of two sentences (a question and an answer to the question). The dialogues can be divided into two groups (without the participants knowing). The first group involves the first two dialogues, and the second group contains Dialogue 3 and 4. Dialogues in the same group provide two different answers with distinguishable word order to the same question. The contexts of each dialogue are labeled below.

Dialogue 1: given, -fronted Dialogue 2: given, +fronted Dialogue 3: new, +fronted Dialogue 4: new, -fronted

According to the results from T-TEST, both the judgments of monolingual Mandarin speakers (Fig.1) and bilingual speakers of Mandarin and English (Fig.2) are significantly influenced by different word orders and givenness independently (for both factors, $p<0.05$). However, it is concluded from a closer look at the two-way interaction between word order and givenness that no significant effect was spotted in any of the groups. Moreover, there was no significant difference between monolingual and bilingual speakers in any of the dialogues since all p-values were smaller than 0.05.

The effects of givenness and different word orders suggest that Mandarin speakers are sensitive to topicalization (fronting) and givenness. However, the lack of interaction effect suggests that the

word- order hypothesis may not be the primary predictor of givenness in Mandarin, aligning with the conclusions of many previous works. Future studies could further explore the expression of givenness through prosody in Mandarin. Meanwhile, it is quite surprising that bilingualism did not necessarily play a role since Mandarin and English have different degrees of freedom in terms of word order. However, the limited number of trials could be a plausible explanation for this observation. Moving forward, more trials involving more diverse lexical items should be employed for this research question.
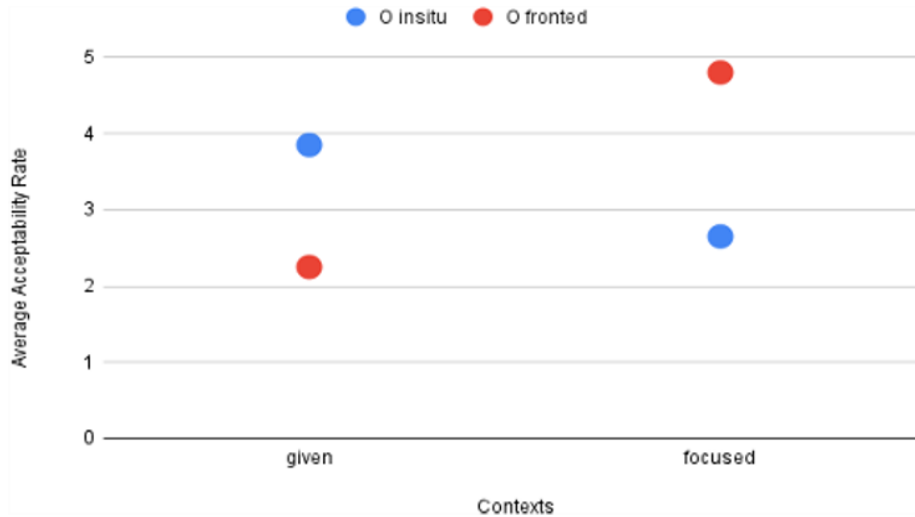
Figure 1: *Mean Acceptability Rate of Each Dialogue given by Monolingual Speakers*
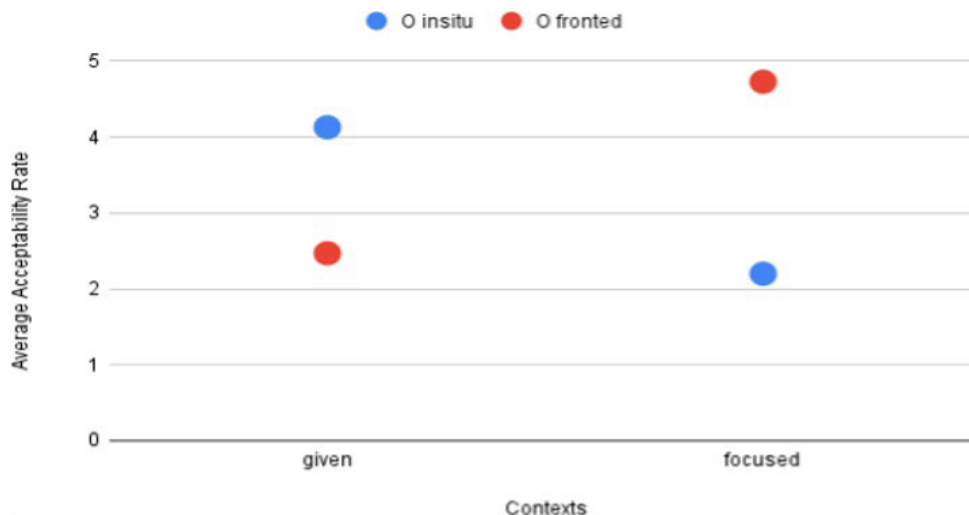
Figure 2: *Mean Acceptability Rate of Each Dialogue given by Bilingual Speakers*

Selected Reference:

[1] Most, R. B., & Saltz, E. (1979). Information structure in sentences: New information. *Language and speech*, *22*(1), 89-95.

[2] Nooteboom, S. G., & Kruyt, J. G. (1987). Accents, focus distribution, and the perceived distribution of given and new information: An experiment. *The Journal of the Acoustical Society of America*, *82*(5), 1512-1524.

[3] Šimík, R., & Wierzba, M. (2017). Expression of information structure in West Slavic: Modeling the impact of prosodic and word-order factors. *Language*, *93*(3), 671-709.

[4] Skopeteas, S., & Fanselow, G. (2010). Focus types and argument asymmetries: A cross-linguistic study in language production. *Contrastive information structure*, 169- 197.

Translating a Culture: Evidence from the Japanese, English and Finnish Comparable Corpora

Elena V. Beloglazova

(Herzen State Pedagogical University, Russia)

Problem statement

Thinning borders in the modern world result in intense translingual/ transcultural communication. Yet, a language is originally fit to describe its own culture, while cultural reorientation requires considerable adaptation.

Research aim

The paper aims at researching the mechanisms of languages' cultural reorientation as exemplified by the Japanese-, English- and Finnish-language Russia-centred discourse. The cultural reorientation of a language is seen as a mode of translation – translation of culture.

Research hypothesis

Basing on the preliminary investigations, the following research hypothesis has taken shape: translation of culture, just as translation in general, depends on the distance between the cultures in contact and the dominant cultural ideology. As a result, notwithstanding similar or comparable linguistic mechanisms of a language adaptation to an exoculture, the source culture translates differently into the different target-language ideologically conditioned discourses.

Research questions

- What is the linguistic strategy underlying translation of culture and how does it compare to the traditional translation?
- What elements of the source culture are selected for constructing the exoculture-oriented discourse in the focus languages?

Materials and methods

The research is based on data in three unrelated languages. The data is organized in comparable corpora of culturally loaded books' titles:

- Rossica-T-En of around 450 book titles covering 17-21 centuries;
- Rossica-T-Fi of around 280 book titles covering 19-21 centuries;
- Rossica-T-Jap of around 200 book titles covering 19-21 centuries.

The corpora have been subject to manual annotation according to 3 groups of parameters, encompassing structural and stylistic features, as well as dominant motifs.

Theoretical framework

The research rests on the following theoretical cornerstones:

- theory of interlinguaculturology, focused on the peculiarities of language of secondary cultural orientation (Kabakchi & Beloglazova 2016);
- titology, being the study of structural and functional aspects of titles (Hollander 1975);
- theory of foregrounding, originating from the works of V. Shklovsky (Shklovsky 1925);
- theory of discourse and ideology (Van Dijk 1988).

Study results
The analysis of the linguistic techniques employed by different languages in the process of cultural reorientation allows for the following generalizations.

Firstly, we have observed the specificity of the strategy of translating culture as contrasted to that of traditional translation, consisting in (a) greater consideration given to preserving the unique culture being the object of translation; (b) less consideration given to language smoothness and readability. The smoothness is normally disrupted by exocode and exoculture elements, as illustrated in (1).

   (1) a. En: White terror : *Cossack warlords of the Trans-Siberian*;
        b. Fi: *Stalinismin lyhyt kurssi. Suomalaiset Moskovan Lenin-koulussa 1926-1938*;
        c. Jap: コサックのロシア―

Secondly, it has been discovered that the features of the exoculture-centred discourse are affected by linguaculture-specific intra- and extra-linguistic factors. The study of the culturonyms, selected to represent the source culture, shows a considerable variation in the choice of representatives, while the lexical types are largely comparable – they are proper names of people and places chosen as prototypes of the cultural category, as well as terms nominating historical and cultural realia (2).

   (2) a. En: *czar, bolshevik*, *Ivan, Muskovy;*
        b.   Fi: *bolshevismi, duma, Venäjä, Pietari*;
        c.   Jap: [*kosakku*], [*rosia*], [*Dousutoefusuki*]).

Thirdly, the preferred translation techniques appear to be to some extent language specific. Thus, Japanese sometimes resorts to a kind of relay translation mediated by English (3).

   (3) .チンの FSB 体制まで(lit., *from Soviet KGB to Putin's FSB*).

The Finnish discourse, foreignizing as it is, avoids transplants, apparently viewing this degree of foreignization as extreme, while for the Japanese it appears a rather common practice to add a decorative transplant in Cyrillic, especially as a paratextual element.
Fourthly, the choice of dominant themes and representatives is largely dependent on the current ideology and on the stereotypes associated with the source culture in the target culture: the discourse has its established canon, guiding what and how can be said about Russia in this particular linguaculture.

Implications of the research
The cross-linguistic research of exoculture-oriented discourse aims at understanding the mechanisms, limits and limitations of bringing cultures over their original linguistic borders, which is crucial for intercultural dialogue and exchange.

Selected References
Kabakchi, Viktor, Beloglazova, Elena. 2016 "Russian English linguaculture," in *Russian English: History, functions, and features*, 35–80. Cambridge: Cambridge University Press.
Hollander, John. 1975 ""Haddocks' Eyes": A Note on the Theory of Titles," in *Vision and Resonance*, 216–
       226. New York: Oxford University Press.
Shklovsky, Viktor. 1925. *On the Theory of Prose*. M.: Krug. (in Russ.)
Van Dijk, Teun. 1988. *News as discourse.* New Jersey: Lawrence Erlbaum Associates Publishers.

A Comparative Bibliometric Analysis of Gender Equality Research Trends in the Maritime Industry and Global Domains

Cuilin Liu, Se-Eun Jhang

(Korea Maritime and Ocean University, Korea)

A recent comprehensive bibliometric study by Liu and Jhang (2024) on gender equality within the maritime industry has set the stage for further explorative comparisons across various global industries. This study aims to expand upon this foundation by employing a comparative bibliometric analysis to scrutinize gender equality research trends from 2003 to 2024. This methodological approach, rooted in the analytical traditions of bibliometrics first conceptualized by Pritchard (1969), has proven effective in diverse fields such as congenital heart disease and cancer research (Farhat et al., 2013) and even in evaluating linguistic variations in COVID-19 research outputs (Fan et al., 2020), thus providing a robust framework for this study.

The current research seeks to address pivotal questions regarding the evolution of gender equality discourse within the maritime sector compared to other industries. These questions include the analysis of research volume changes over time, identification of influential works and authors, contributions from countries with emerging maritime markets, and the evolution of research themes post the adoption of Sustainable Development Goal (SDG) 5 in 2015, which emphasized gender equality and women's empowerment.

Utilizing data from the Web of Science, this study employs both quantitative and qualitative bibliometric methods to analyze research volume, key topics, and the most-cited articles, alongside examining the interconnections among topic clusters. The methodology involves converting raw data from the Web of Science into a plain text format for comprehensive analysis, supported by the use of bibliometric tools like Python, R's bibliometrix package, VOSviewer, and Zotero.

Keywords: gender equality, SDG5, bibliometric analysis, research trends, maritime industry, global domains, comparative study

Selected References

Fan J, Y. Gao, N. Zhao, R. Dai, H. Zhang, X. Feng, G. Shi, J. Tian, C. Chen, B. D. Hambly and S. Bao. 2020. "Bibliometric Analysis on COVID-19: A Comparison of Research Between English and Chinese Studies," *Frontiers in Public Health* 8(477), 1-10. doi: 10.3389/fpubh.2020.00477

Farhat, Theresa, Zahi Abdul-Sater, Mounir Obeid, Mariam Arabi, Karim Diab, Samer Masri, Zohair Al Haless, Georges Nemer, and Fadi Bitar. 2013. "Research in Congenital Heart Disease: A Comparative Bibliometric Analysis Between Developing and Developed Countries," *Pediatric Cardiology* 34, 375-382.

Liu, Cuilin and Se-Eun Jhang. 2024. "A Bibliometric Analysis of Research on Gender Equality in the Maritime Industry," a paper presented at the 2024 Spring Conference of The New Korean Association of English Language and Literature held in Korea Maritime and Ocean University, Busan, South Korea on May 11, 2024.

Pritchard, Allan. 1969. "Statistical Bibliography or Bibliometrics?" *Journal of Documentation* 25(4), 348-349.

Integrating Cultural Contexts in AI:
A Study on Thai Boxing Lexicon and Language Comprehension

Jinawat Kaenmuang

(Chulalongkorn University, Thailand)

The lexicon of *Muay Thai*, or Thai boxing, is not just a collection of words but a unique cultural phenomenon that transcends everyday conversation. Unlike languages that draw on sports metaphors from various games (Segrave 2000), *Muay Thai* is a singular all-encompassing sport deeply interwoven into Thai society, covering leisure activities, local sports, martial arts, and combat skills. This integration has endowed Thai boxing with a vast reservoir of figurative language, enabling the conceptualization of diverse activities, events, and situations. This study addresses a significant scholarly gap by examining how this unique *Muay Thai* terminology permeates Thai daily discourse, influencing metaphorical expressions and cultural conceptualizations.

To explore this phenomenon, linguistic manifestations of Thai boxing were analyzed across various cultural contexts. Multilexemic units and idioms derived from boxing were scrutinized to evaluate their impact on semantic meaning and discourse. The dataset, a comprehensive collection of 1,500 examples, was sourced from the Thai National Corpus (TNC), X (formerly Twitter), and Pantip (a popular Thai discussion forum), with 500 examples from each source. Cognitive semantics, ethnosemantics, and cultural semantics were employed to construct metaphorical models that uncover the intricate relationship between sport, language, and culture.

Findings underscore the distinctive lexicon of Thai boxing in discourse, emphasizing its role in shaping event conceptualization within Thai culture. The study explores the profound influence of sports on the conceptualization of daily life, from pre-fight preparations such as "to get started" (lit. "open the ring," *pɤ̀ːt săŋ.wīan*), "to join the competition" (lit. "get on the squared circle," *kʰûn săŋ.wian*), and "to quit doing" (lit. "say farewell to the ring," *ʔām.lāː săŋ.wīan*), to in-ring combat actions like "to attack with secret information" (lit. "throw a punch," *plɔ̀ːj màt*), "to take turns" (lit. "cross fists with," *lêːk màt*), and "to fight without choices due to being oppressed for a long time" (*lûat kʰâw tāː*), and post-fight outcomes such as "severely or life-threatening injured" (lit. "yellow chin," *kʰāːŋ lŭaŋ*), "to hang something up" (lit. "hang boxing gloves," *kʰwɛ̌ːn nūam*), and "to agree to lose dishonestly" (lit. "pretend to lose in a boxing match in order to receive agreed-upon benefits," *lóm mūaj*).

To advance the field in alignment with recent developments in artificial intelligence (AI), particularly regarding the development of large language models (LLMs), it is crucial to understand the nuanced cultural contexts embedded within language use (Chowdhary 2020, Li et al. 2024), as demonstrated by the terminology of Thai boxing. The ability of AI to effectively process and interpret natural language relies heavily on comprehensive datasets that capture the cultural and metaphorical nuances specific to languages. The intricate lexicon of *Muay Thai* exemplifies this challenge, as its terms frequently embody indirect and figurative meanings deeply rooted in cultural practices. This underscores the complexity of the task at hand and highlights the importance of culturally informed data for enhancing the interpretative capabilities of AI in processing such specialized language.

To achieve this goal, the study adopted a comprehensive interdisciplinary approach integrating cognitive semantics, ethnosemantics, and cultural semantics. Cognitive semantics illuminates how language users conceptualize their experiences, while ethnosemantics provides insights into the cultural-specific meanings of terms. Cultural semantics bridges these disciplines, revealing how metaphorical language reflects and shapes cultural norms (Croft 2009). This multifaceted analysis highlights the necessity for AI to integrate linguistic and cultural insights to process languages like

Thai effectively, underscoring the interdisciplinary nature of this study.

Ultimately, this research underlines the need for a balanced integration of linguistic and cultural insights to advance AI. AI technologies such as LLMs can leverage extensive corpora to identify patterns in the use of Thai boxing terminology. Through meticulous analysis of vast datasets, AI can uncover the widespread use and metaphorical implications of these terms across diverse contexts. This approach enhances AI's precision in interpreting natural language and underscores the essential role of cultural knowledge in language comprehension. By incorporating culturally nuanced elements such as the *Muay Thai* lexicon into AI training datasets, AI's ability to process language is significantly enhanced, demonstrating the practical implications of this research for AI technology.

This study makes valuable and significant contributions to the field by establishing a robust correlation between cultural understanding and AI technology in processing natural language, particularly in languages rich in cultural and figurative nuances like Thai. This integration not only enhances AI's interpretative abilities but also serves as a blueprint for future research in other culturally rich languages, thereby broadening the scope and applicability of AI in linguistics.

Selected References

Chowdhary, K. R. 2020. *Fundamentals of Artificial Intelligence*. New Delhi: Springer India. Croft, William. 2009. "Towards a Social Cognitive Linguistics," in Vyvyan Evans and Stéphanie Pourcel, eds., *New Directions in Cognitive Linguistics*, 395–420. Amsterdam/Philadelphia: John Benjamins.

Li, Cheng, Chen, Mengzhou, Wang, Jindong, Sitaram, Sunayana, and Xie, Xing. 2024. "CultureLLM: Incorporating Cultural Differences into Large Language Models", *arXiv:2402.10946v1*. Online first, doi: 10.48550/arXiv.2402.10946

Segrave, Jeffrey O. 2000. "The Sports Metaphor in American Cultural Discourse," *Culture, Sport, Society* 3(1), 48–60.

General Sessions
(Day 2)

Hands that Speak: Metaphorical and Metonymical Insights from Korean 손 'Hand'- Related Idioms

You-Yun Yu

(National Taiwan University Graduate Institute of Linguistics, Taiwan)

Body-part idioms are common in everyday language but often pose challenges for non-native speakers. Despite significant research on body-part idioms, studies focusing on metaphors and metonymies in Korean 손 'hand' related idioms are limited. Most existing studies focused on comparisons between Korean and other languages. For instance, Kim (2014) analyzed bodily idioms in Korean and Chinese. Many researchers, like Kwon and Sung (2009), have concentrated only on ontological metaphors, identifying "BODY-PARTS ARE CONTAINERS" as a frequent metaphor. However, these studies typically analyzed metaphors and metonymies related to all body parts rather than focusing on one specifically, limiting their depth of analysis. To address this gap, the current study aims to comprehensively investigate conceptual metaphors and metonymies in Korean hand-related idioms. Additionally, as a pilot analysis for future studies, it compares our findings with Teng's (2017) study on Mandarin hand-related idioms.

Applying Lakoff and Johnson's (2008) conceptual metaphor theory and metonymy framework, we analyzed 94 idioms, identifying orientational metaphors, ontological metaphors, and four types of metonymies such as PART STANDS FOR WHOLE, WHOLE STANDS FOR PART, INSTRUMENT STANDS FOR ACTION, and INSTRUMENT FOR ABILITY TO PERFORM THE ACTION. The examples with their literal and idiomatic meanings for each type of metaphor and metonymy are illustrated in (1).

(1)   a. 손이 비다 (Orientational metaphor)
              Hand is empty. (Having no work on hand)
         b. 손을 들다 (Ontological metaphor)
              Raise one's hand. (Give up; Surrender)
         c. 손 달리다 (PART STANDS FOR WHOLE)
              Not enough hands. (short-handed)
         d. 손을 넘기다 (WHOLE STANDS FOR PART)
              Missed one's hand. (Miss an opportunity)
         e. 손을 멈추다 (INSTRUMENT STANDS FOR ACTION)
              Stop one's hand. (Stop doing things)
         f. 손에 오르다 (INSTRUMENT FOR ABILITY TO PERFORM THE ACTION)
              Get on one's hand. (Become familiar with tasks)

The study found that shared human experiences lead to common conceptual metaphors and metonymies in Korean and Chinese idioms related to hands. For example, the metaphor HOLDING IN THE HAND IS CONTROL, exists in both languages, reflecting how hands are typically used to grab objects. The metonymies, HAND STANDS FOR PERSON and HAND STANDS FOR ACTION, are also prevalent in both languages. However, certain expressions are unique to each culture. For instance, Korean idioms include HAND IS ROPE, HAND STANDS FOR EMOTION, and THOUGHTS ARE HANDS, while Mandarin features PRESERVATION IS BACK and HANDS STAND FOR PROCESS. These differences reflect varying cultural perspectives. For example, in Chinese, HANDS STAND FOR PROCESS highlights the notion of being under a second process, as seen in the term 二手車 ('second-

hand car'), whereas the Korean term 중고차('second-hand car') emphasizes the car's age and condition. The appearance for the unique conceptual metaphors and metonymies in Korean and Mandarin is summarized in (2).

(10) Differences of Hand-Related Idioms between Korean and Mandarin.

Types of Conceptual Metaphor	Korean	Mandarin
HAND IS ROPE	○	×
THOUGHTS ARE HANDS	○	×
PRESERVATION IS BACK	×	○
Types of Metonymy	Korean	Mandarin
HANDS STAND FOR PROCESS	×	○
HAND STANDS FOR EMOTION	○	×

In conclusion, this study provides insights into the types of conceptual metaphors and metonymies that appear in hand-related idioms in Korean and offers a brief comparative analysis with Mandarin, highlighting both similarities and distinctions.

Selected References

Lakoff, G., & Johnson, M. (2008). *Metaphors we live by*. University of Chicago press.

Kim, Hyejin (2014). *A Cognitive-Linguistic Study on Conceptualization of Korean and Chinese Bodily Idioms: Conceptual Metaphor and Chinese Teaching Application*. https://hdl.handle.net/11296/xafnfj.

Kwon Ik-Ho, & Sung Min-Hee (2009). On Conceptualization of Korean and Japanese Bodily Idioms in Terms of Ontological Metaphors. *JOURNAL OF JAPANESE LANGUAGE AND LITERATURE*, 41, 19-36. 10.18631/jalali.2009..41.002

鄧孟倫. (2017). 華語與閩南語手部三字格認知隱喻研究. *輔大中研所學刊*, (35&36), 85-109.

Understanding Embodiment in Phraseological Expressions –
Body Part Nouns in German and Korean

Dongeun Lee

(Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany)

In embodiment hypothesis, body parts are seen as the primary means through which individuals engage with their physical environment. This interaction between human body and the extralinguistic reality is mapped onto abstract concepts, fundamentally shaping our conceptualization of the world. Accordingly, in language, nouns denoting body parts are one of the most significant components of phraseologisms across various languages, displaying varying degrees of syntactic fixedness and semantic compositionality. Such units express a range of concepts from physical activity and spatial orientation to abstract notions such as emotional states, control, hierarchy and social interaction, and have been the subject of extensive research (Palm 1995, Fleischer 1997, Staffeld & Ziem 2008, Krohn 1994).

The present study is driven by the question of whether two linguistically distinct languages, Korean and German, exhibit similar patterns of embodiment that underscore cognitive regularities. The investigation focuses on phraseologisms that incorporate body part nouns, aiming to capture both similarities and differences. Previous research, such as Kang (2011) and Kim (2003), has provided contrastive analysis of somatic phraseologisms centered on specific body parts in these languages, while others like Byun (2003) have focused on in-depth syntactic and morphological categorizations. Building on these studies, this research extends the scope to include a broader array of body part nouns and analyze their mechanisms of embodiment.

A dictionary-based method was employed to select somatic phraseologisms for examining conventionalized and entrenched linguistic units. Expressions registered under each lexical item were extracted from 2019 edition of Langenscheidt's *Großwörterbuch Deutsch als Fremdsprache* and the online version of the *Standard Korean Dictionary* by the National Institute of the Korean Language. A qualitative analysis was conducted on 437 items in German and 482 items in Korean, focusing on keywords such as *face, eye, nose, mouth, chest, heart, hand, neck, leg, foot* and *ear*. The analysis involved an examination of each body part as well as in relation with its physically adjacent body parts. The key findings of this study reveal both divergent and convergent patterns in the use of somatic phraseologisms between Korean and German. Divergences likely arise from differences in syntactic structures and cultural contexts, which influence how experiences are conceptually encoded. These divergences can be categorized into three types: (i) seemingly equivalent phraseologisms with distinct meanings in each language (false friends), (ii) same body part nouns denoting different target domains, (iii) similar concepts conveyed by different, yet locally adjacent body part nouns. For instance, the concept of losing one's face is expressed using *face* in German but *nose* in Korean. Similarly, expressions of sincerity and a broad range of emotions are often conveyed with *heart* in German and *chest* in Korean.

Nevertheless, the study also identified convergences in the embodiment of non-linguistic actions, driven by conceptual metaphor and metonymy, which revealed two primary patterns across the languages. To begin with, both languages demonstrated shared mechanisms of metaphoric extension for certain body parts. A notable example is the metaphorical conceptualization of *eyes* as windows or counters, implying they provide insight into one's emotional state, desires, or thoughts. They can be selectively opened or closed, reflecting selective interest or attention. This supports the hypothesis that bodily actions underpin a significant foundation of our conceptualization, generally consistent across languages. Furthermore, there were prominent metonymic patterns observed in both languages. These patterns indicate cognitive regularities in

how physical actions or stimuli are extended and entrenched as relevant linguistic expressions.

Selected References

Fleischer, Wolfgang (1997): *Phraseologie der deutschen Gegenwartssprache*. 2. Aufl. Tübingen: Niemeyer.

Jackendoff, Ray (2002): *Foundations of Language. Brain, Meaning, Grammar, Evolution*. Oxford/New York: Oxford University Press.

Krohn, Krain (1994): *Hand und Fuß. Eine kontrastive Analyse von Phraseologismen im Deutschen und Schwedischen*. Göteborg: Acta Universitatis Gothoburensis.

Lakoff, George & Johnson, Mark (1980): *Metaphors We Live By*. Chicago: The University of Chicago Press.

Palm, Christiane (1995): *Phraseologie. Eine Einführung*. Tübingen: Gunther Narr.

Staffeldt, Sven & Ziem, Alexander (2008): Körper-Sprache: Zur Motiviertheit von Körperteilbezeichnungen in Phraseologismen. *Sprachwissenschaft* 33(4), 455-499.

강병창 (2011): 독일어와 한국어 손 관용구의 비유적 의미구성 비교 – 은유, 환유, 은환유를 중심으로. 『독어교육』 제 50 집, 81-104.

김수남 (2003): 독일어와 한국어의 관용표현 비교연구 – 신체어휘 관련 관용구를 중심으로. 『독일어문학』 제 20 집, 331-357.

변명선 (2003): 독일어와 한국어 신체어휘 관용표현 비교 연구. 『독일언어문학』 제 19 집, 59-80.

Pragmatic Analysis of Testimonial Strategies in Perjury Court Rulings

Sue Young Chung, Eugene Chung
(Korea University, Korea)

This study aims to examine the language of perjury in Korean court rulings. Perjury cases offer rich material for linguistic analysis as they are crimes executed through language. Grice (1975)'s Cooperative Principle (CP) posits that participants in a conversation intend to be cooperative. The CP consists of four maxims expected of a cooperative speaker: the maxim of Quantity (provide the right amount of information), Quality (say only what you believe to be true), Relation (be relevant), and Manner (be clear and unambiguous) (Birner, 2013). A lie, which involves knowingly making a false statement, violates the maxim of Quality. When such a lie is made under oath during court testimony, it constitutes perjury (Shuy, 2011).

Previous research has explored the relationship between the language of perjury and principles of lying, along with other forms of deception (Green, 2018; Skoczen, 2022; Solan, 2022), and the connection between perjury and the CP (Tiersma, 1990). Shuy (2011) proposed a framework encompassing broad contextual and detailed interactional perspectives, which this study adopts. While prior studies have primarily focused on classifying statements as lies that mainly violate the maxim of Quality, this research delves into how testimonies strategically navigate between clear-cut lying and truth. Specifically, it investigates 1) how violations of the maxims relate to testimonial strategies and 2) the roles of power and cooperative dynamics in these strategies.

The framework adapted from Shuy (2011) is applied to analyze perjury court rulings, followed by an examination of power and cooperation dynamics through Critical Discourse Analysis (van Dijk, 2005). Perjury cases were collected from district court websites and online repositories of court cases, which were screened to include a minimum number of question-and-testimony pairs. The framework employs contextual, interactional, and featural-level tools. The contextual perspective examines speech events, participant schemas, and interlocutor agendas as suggested by topic usage. This is complemented by the interactional perspective, including the roles of the CP (Grice, 1975) and speech acts (Searle, 1965). The featural level analyzes language ambiguity and word meanings to provide additional insights into perjury case conversations.

The following are excerpts of sample court rulings.

(1) *Case of perjury in selling alcohol to minors, Seoul Northern District Court, 2016kodan3567*

Q1. (Judge) "What were you doing then? Were you drinking with them?" A1. (Witness) "I was just sitting there."

Q2. (Judge) "You didn't have any alcohol?"

A2. (Witness) "No, I just sat there and watched."

Q3. (Judge) "So, you did not consume any alcohol there, correct?"

A3. (Witness) "No, I had already had alcohol elsewhere before I arrived there."

Q4. (Attorney) "When you prepared your written statement, you reported being at the [anonymized] bar without consuming alcohol. You are stating there has been an amendment to this particular detail. Is that right?"

A4. (Witness) "Yes."

Q5. (Judge) "The police conducted a raid while you were there; were you still not drinking any alcohol at [anonymized], not even a glass?"

A5. (Witness) "I was drinking something, like water, to sober up; like water. I think I had water or cola, just eating the stew and other snacks."

(2) *Case of perjury in false car insurance claim, Ulsan District Court, 2015kojung1862-1*

    Q1. (Attorney) "Did you ride in the defendant's car, and while driving around, engage in kissing several times, and was it during one of those times that the accident occurred?"

    A1. (Witness) "Yes."

    Q2. (Attorney) "At the time of the accident, were you too shocked to remember exactly what happened, but upon regaining your senses, found the car had plunged into a reservoir?"

    A2. (Witness) "Yes."

Examples (1) and (2) demonstrate how questioners exert control over the topic and the overall interaction. In (1), the witness repeatedly denies consuming alcohol, which contradicts the judge's schema, or expectations of events. This leads the judge to persist with the same line of questioning. In (2), the lawyer employs their control over the topic to advance the order of events smoothly, leaving minimal opportunity for the witness to contribute.

The two examples also highlight how witnesses provide contrasting responses based on the level of cooperation between participants. In (1), the questioners and the witness have a neutral-to-adversarial relationship, as evidenced by A1 where the witness compromises the maxim of relevance and slightly avoids a direct answer, and in Q2, where the judge redirects the question to clarify if the witness has consumed alcohol at that point of time. In contrast, in (2), the questioner and witness have a cooperative relationship with a shared agenda, which facilitates a rapid progression of topics.

The examples illustrate the strategies the participants employ to achieve their objectives. In (1), the judge establishes a compelling narrative in Q5 by prompting the witness to warrant their testimony. This is achieved by using the phrases "still" and "not even a glass". In (2), the attorney violates the maxim of Quantity, gradually shaping the events for the witness in a cooperative manner. Conversely, despite the witness's simple "yes" responses, the maxim of Quantity is not violated due to the abundance of information in the questions.

The analysis suggests that power dynamics, cooperativity, and strategic goals play a significant role in how witnesses formulate their testimonies, and how judges and lawyers place their questions. Power asymmetry within the courtroom affects the testimony, as legal professionals control the dialogue and questioning sequence. Witnesses are obliged to answer as requested with the knowledge that what they say may be used against them, thereby adding to the burden on their speech (Shuy, 2011; van Dijk, 2008). Cooperativity also influences participant speech: while courtroom exchanges require observation of the CP for effectiveness, the inherently adversarial nature of the proceedings fosters mistrust and competition toward winning. Although adherence to the Quality maxim of the CP is expected, the other three maxims may be strategically compromised under this setting (Solan, 2022; Tiersma, 1990). Finally, questioners and witnesses strategically craft their speech, often violating conversational maxims to a certain degree in order to balance the competing goals of conveying sufficient information while minimizing the potential risk (McCornack, 1992). This complex interplay of factors highlights the need for further investigation of conversation strategies in legal testimonies.

Selected References

Birner, B. J. (2013). Introduction to pragmatics. Wiley-Blackwell, A John Wiley & Sons, Ltd., Publication.

Green, S. P. (2018). Lying and the Law. In J. Meibauer (Ed.), The Oxford Handbook of Lying (pp. 483- 494). Oxford University Press.

Grice, H. P. (1975). Logic and Conversation. Syntax and Semantics, 3, 41-58.

McCornack, S. A. (1992). Information manipulation theory. Communication Monographs, 59(1), 1-

16. Shuy, R. W. (2011). The Language of Perjury Cases. Oxford University Press.

Skoczen, I. (2022). Modelling Perjury: Between Trust and Blame. International Journal for the Semiotics of Law, 35(Issue 2), 771-806.

Solan, L. M. (2022). Lies, Deception, and Bullshit in Law. In L. R. Horn (Ed.), From Lying to Perjury: Linguistic and Legal Perspectives on Lies and Other Falsehoods (pp. 381-407). De Gruyter Mouton.

Tiersma, P. M. (1990). The Language of Perjury: Literal Truth, Ambiguity, and the False Statement Requirement. Southern California Law Review, 63(Issue 2), 373-432.

van Dijk, T. A. (2005). Critical Discourse Analysis. In D. Schiffrin, D. Tannen, & H. E. Hamilton (Eds.), The Handbook of Discourse Analysis (pp. 349-371). John Wiley & Sons Incorporated.

van Dijk, T. A. (2008). Discourse and power. Palgrave Macmillan.

Investigating Mandarin Clear Speech:
Acoustic Cues and Their Applications in Technology and Therapy

Chun Ting Chien, François-Xavier Brajot

(National Taipei University of Nursing and Health Sciences, Taiwan)

Modifying speech for effective communication under challenging conditions is a key issue in speech science and rehabilitation. While "clear speech" is well-studied in many languages (Maniwa et al., 2008; Smiljanić & Bradlow, 2005), it remains underexplored in Mandarin, particularly in terms of segmental changes and lexical tone effects. Contemporary AI can potentially help clinicians extract relevant parameters in real-time, aiding in dynamic assessment, iterative analysis, and procedural assistance. However, the effectiveness of these applications depends on automatic speech recognition systems, which struggle with disordered speech. Understanding how clear speech enhances acoustic cues in Mandarin can improve automatic speech recognition models and clinical practices in speech-language pathology.

Thirty female speakers (mean age 25) with healthy hearing and voice produced twelve target words embedded in carrier sentences using two different speaking styles: clear and casual. Acoustic measurements included vowel space area, vowel space dispersion, tone space area, pitch range, and several tonal cues (see Table I.; Chen & Gussenhoven, 2008; Leung & Wang, 2020; Tupper et al., 2020). Statistical analysis employed linear mixed-effects models to compare measurements across speech registers, a principal component analysis and a logistic regression on the tonal cues to break down psychoacoustic percepts into acoustic correlate.

First, vowel analysis revealed an expansion and dispersion of vowel space in clear speech, driven by higher /a/ F1 and lower /u/ F2, unlike findings reported for English. Second, tonal analysis showed an expansion in tone space area and pitch range, driven by an increase in the higher frequency range.

Clear speech was characterized by specific pitch and temporal patterns, but not equally across all tones. For example, Tones 1 and 2 were associated with a significant increase in mean f_o, while Tones 3 and 4 were mainly associated with steeper f_o slopes. In general, clear speech tones exhibit higher f_o mean, steeper f_o slope, wider tonal range, longer duration, an earlier relative turning point, and less creakiness. Principal component analysis identified pitch change cues (f_o slope, tonal range, relative turning point) as primary psychoacoustic parameters, while pitch height cues (f_o mean, HNR) are secondary. More importantly, logistic regression revealed an apparent coupling between duration and pitch change in casual speech that was not apparent in clear speech. In other words, there seems to be a trade-off between duration and pitch change in casual speech, while pitch change remains a distinctive attribute in clear speech. This finding implies potential capability of acoustic measurements to identify changes in psychomotor strategies, and it supports the autosegmental phonology framework and Soft TEMplate Mark-up model (Kochanski & Shih, 2003).

This study underscores the importance of vowel and tonal cues in enhancing speech intelligibility, which can aid speech therapists using automatic speech recognition models. Models identifying changes in acoustic characteristics (eg. vowel space area, tone space area, relationship between pitch change and duration, etc.) could provide information on treatment techniques, or psychomotor strategies changes. Further research will refine ASR models narrow phonetic transcription and detailed suprasegmental data to better accommodate disordered speech and inform effective treatment strategies in speech-language pathology.

TABLE I. List of tonal measurements used in the present study and their brief description.

Tonal Cues	Description
f_o mean	average f_o values of target tones, which represent pitch height
f_o slope	the slope magnitude of the f_o rise or fall between max f_o and min f_o
Tonal range	range between f_o maximum and f_o minimum of target tones Duration the length of time of the target tones
Relative turning point	the proportional time at which the f_o fall or rise begins relative to the tonal duration
Creakiness	quantified by harmonics-to-noise ratio (HNR), and occurs when vocal folds compress tightly and vibrate irregularly

Selected References

Chen, Y., & Gussenhoven, C. (2008). Emphasis and tonal implementation in Standard Chinese. *Journal of Phonetics*, *36*(4), 724–746.

Smiljanić, R., & Bradlow, A. R. (2005). Production and perception of clear speech in Croatian and English. *The Journal of the Acoustical Society of America*, *118*(3), 1677–1688.

Prosodic Tones and Boundaries: Disambiguating
Non-canonical *Wh*-questions and *Wh*-exclamatives in Korean

Geonhee Lee

(Stony Brook University, USA)

Korean *wh*-words are lexically ambiguous in that their interpretation can be *wh*-pronouns as in *wh*-questions or indefinite pronouns as in *yes-no* questions. For instance, *mwues* can mean 'what' or 'something/anything' depending on its context. It has been argued that the reading of *wh*-words can be disambiguated by phonological phrasing (Richards 2010, Cho 1990): a *wh*-word is focused when it is used as a *wh*-pronoun, so a phrase boundary does not exist in a phonological phrase of *wh*-questions. Sentence-final intonation is also known to interact with the sentence types (see Lee & Ramsey 2000 and others). When the sentence is a *wh*-question or a statement, sentence-final intonation is a falling contour, and it is a rising contour when it functions as a *yes-no* question.

Interestingly, the Korean *wh*-word *mwues* 'what' allows non-canonical readings when it is followed by the adverb *kulehkey* 'so', as the sequence can yield a *why*-like *what*-question reading and a *wh*- exclamative reading, in addition to the canonical *what*-question reading (1). Kim (2022) notes that such readings are possible when *mwues* 'what' and *-ul* is concatenated as a single word *mwel* 'what-ACC'.

(1)   *Yuna-ka   mwel   kulehkey   manh-i   mek-ess-e*
        Yuna-NOM   who-ACC   so   much-CONN   eat-PST-NEUT
        a.   'What did Yuna eat so much of?'   [*what* question]
        b.   'Why did Yuna eat so much?'   [*why* question]
        c.   'How much Yuna ate!'   [*wh*-exclamation]

Although the sequence has been studied from syntactic and semantic perspectives, no systematic examination has been provided on its prosodic properties. In this study, an online speech corpus was examined to observe whether prosody can differentiate the three possible readings of this sequence. Since all three readings involve *wh*-pronouns, indefinite pronoun interpretations were excluded in this study. The data was gathered from YouGlish, a website that compiles videos from YouTube for

language learning and allows searches for linguistic expressions. The search string '뭘 그렇게' yielded

a total of 204 instances. With the manual exclusion of irrelevant data, 57 audio files were extracted, categorized by various factors. Firstly, the interpretation of each data point was primarily based on the author's native intuition of Korean. Then, the annotation of the audio files were done using Praat. Tonal transcription was based on the K-ToBI (Korean Tones and Break Indices) system, which describes intonation as tonal sequences at Accentual Phrase (AP) and Intonational Phrase (IP) levels (Jun 2000). AP tones were marked as H (High) or L (Low), and two additional markers were used for indicating boundary tones: 'Ta' for IP-medial AP final tones and 'T%' for IP-final AP final tones. The voicing of the onset of the first syllable of *kulehkey* helped in detecting the presence of a prosodic boundary, with a voiced [k] indicating no boundary and a devoiced [k] implying a boundary.

Figure 1: General distribution of three interpretations of the sequence *mwel kulehk*ey based on the presence of a prosodic boundary
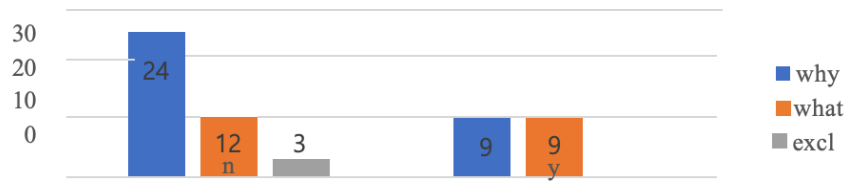
As shown in Figure 1, the *mwel kulehkey* sequence more frequently appears without a prosodic boundary (39) than with one (18). Regardless of the presence of boundary, the interpretation favored a *why*-question (33), followed by a *what*-question (21), with a *wh*-exclamative being the least frequent reading (3). In terms of tonal patterns, *why*-questions typically had the 'L +H La' tonal contour without a boundary (24) than with one (9). The frequency of *what*-questions without a boundary (12) slightly outnumbered the instances with a boundary (9), favoring 'L +H La' contour. Thus, it suggests that tonal contour without a phrase boundary generally triggers a *wh*-question reading, which is in line with the dephrasing effect (Richards 2010). In terms of the IP-boundary, as shown in Table 1, LH% appeared most frequently in that the pattern occurred in *wh*-questions. For *wh*-exclamative readings, however, they tended to associate with more complex boundary tones (e.g., LHL%, HLH%) that are known to convey a speaker's emotions like frustration or surprise.

		IP-boundary						Total
		LH%	L%	H%	HL%	LHL%	HLH%	
type	*why*-q	**16**	7	7	1	1	1	33
	what-q	**8**	7	3	3	0	0	21
	wh-excl	0	0	1	0	**1**	**1**	3
Total		24	14	11	4	2	2	57

Table 1: Distribution of IP-boundary tones based on the sentence types

Another novel finding was a disambiguation cue in the word *kulehkey,* as it has two different meanings: 'so' and 'like that'. The 'like that' interpretation was only found in *why*-questions, not in *what*-questions. Accordingly, the type of the predicate helps disambiguate the readings. For instance, a 'like that' reading appeared primarily when the syntactic type of the main predicate of the sentence was a verb, whereas a 'so' reading was found in various syntactic contexts.

	why-q	*what*-q	*wh*-excl	Total
so	17	21	3	41
like that	**16**	**0**	**0**	16
Total	33	21	3	57

Table 2: Distribution of the meanings of *kulehkey* based on the sentence types

In sum, while there were no clear AP-tonal distinctions of the *mwel kulehkey* sequence between three readings, there was a difference in terms of the presence of AP-tonal boundary in the sequence: the absence of the boundary mainly occurred in *wh*-question readings, detected by AP-internal intervocalic voicing cues, aligning with the dephrasing effect of *wh*-questions. It also exhibited a difference in terms of IP-boundary tones of each reading. In particular, LH% occurred frequently in *wh*-question readings. Overall, there seem to be both IP-internal and external disambiguation cues of the sequence to a weaker degree, in addition to the contextual cue from the predicate type. The findings suggest that the interplay between prosody and syntax is crucial for understanding the meaning of *wh*-questions in Korean, providing insights of how tonal patterns and prosodic boundaries contribute

to disambiguation of Korean *wh*-words.

References

Cho, Young-mee Yu. 1990. Syntax and Phrasing in Korean. *The Phonology-Syntax Connection*. In S. Inkelas and D. Zec (eds.), 47-62. Chicago: University of Chicago Press.

Jun, Sun-Ah. 2000. K-ToBI Labelling Conventions. https://linguistics.ucla.edu/people/jun/ktobi/k-tobi.html.

Kim, Okgi. 2022. *The Syntax and Semantics of Non-standard Wh-constructions in Korean*. Ph.D. dissertation, Wisconsin: University of Wisconsin-Milwaukee.

Lee, Iksop, and S. Robert Ramsey. 2000. *The Korean Language*. Albany: State University of New York Press.

Richards, Norvin. 2010. *Uttering Trees*. Cambridge, MA: MIT Press.

Perceptual Adaptation to Foreign Accented Speech by L2 Listeners

Riley Kent, Eunhae Oh (Konkuk University, Korea)

In their first language (L1), listeners employ normalization and adaptation processes to recognize and comprehend new talkers within a few seconds of exposure (Clarke & Garrett, 2004). They can also generalize that learning to new talkers with the same L2 accent, resulting in faster reaction times and higher accuracy rates than listeners without previous exposure to the accent (Xie et al., 2018). Millions of successful L2 speakers of English as a lingua franca globally are accomplishing the same feats, but their processes of adaptation and generalization to foreign-accented speech are less studied.

The current study examined the patterns of L2 listeners adaptation to foreign accented speech within brief exposures, following the methodology of two previous studies on rapid accent adaptation and generalization in L1 listeners: Clarke and Garrett (2004) and Xie et al. (2018). In the current study, thirty-six Korean university students performed an online bimodal matching task that measured processing efficiency through reaction time (RT) and accurate recognition of final monosyllabic words in English sentences spoken by different accented female speakers. After 18 sentences of training (3 blocks of 6 sentences) that varied by condition, all participants were tested on 6 sentences (test block) spoken by the same Spanish-accented speaker. The Control group heard blocks 1-3 spoken by an American (accent adaptation without exposure), the Speaker group heard the same Spanish talker as the test block (accent adaptation with exposure), and the Accent group heard a different Spanish talker (accent generalization).

The current study accounted factors known to influence accent adaptation in L1 listeners through careful talker and stimuli selection using accentedness ratings by L1 and L2 listeners, as well as instructions and short breaks that prepared listeners for talker changes. Stimuli were also chosen from a different speech-in-noise test list (HINT list) than previous studies because the have greater phonemic. Considering the additional difficulties of reading in an L2, target and probe words were also controlled for spelling transparency and L2 familiarity.

L1 listeners in previous studies showed initial processing delays when exposed to foreign-accented speech, but then rapidly improved in processing efficiency within very brief exposures (within 4 sentences for RT, 16 sentences for error rates). Additionally, some L1 listeners' processing of L2- accented speech adapted to match RTs in L1-accented trials (Clarke & Garrett, 2004). Lastly, when a new talker of the same L2-accent was introduced, L1 listeners trained on that L2 accent showed generalization effects with RTs faster than the control group (Xie et al., 2018). Control groups in both studies showed processing difficulties with slower RTs and higher error rates when they encountered foreign accented speech for the first time in the test block.

Results from the L2 listeners in the current study showed some similarities but also some marked differences from L1 patterns of adaptation. Although the Accent group showed steady improvement across all four blocks, overall accuracy rates were so high as to induce ceiling effects that made further statistical analysis unfeasible. Analysis of RT (log-transformed) was sometimes similar to L1 listener results: the Accent and Speaker groups showed a decrease across the entire test, both in exposure blocks and then further decreasing in RT in the test block (see Figure 1), and the Control group exhibited quicker decreases in RT during training trial blocks. A mixed-effects linear regression model confirmed significant effects of group as well as group and progressive block interactions, with random intercepts for individual speakers and individual stimuli.

In general, the groups got progressively faster, most likely indicative of improvements based on practice with the task. The previous studies used differences in RT slope between groups as measures of comparative adaptation, but the Control group in the current study seemed to hit a

lower threshold in the middle of the second block, indicating that continued decreases by the other two groups did show adaptive processing efficiency improvements.

Figure 1. Mean Log-transformed Reaction Time (LogRT) by block and group

However, in stark contrast to previous studies, L2 listeners in the Control group showed very little comparative processing delay when encountering Spanish-accented speech in the test block. While their RTs did increase slightly from the previous block of American-accented speech, they remained faster overall than the groups previously exposed to Spanish-accented speech (see Figure 1). Additionally, examining the initial training block wherein all three groups encountered a change of speaker, accent, or both from the practice block speaker (see Figure 2), none of the three groups show evidence of processing delays comparable to delays seen in L1 listeners after changes in accent or speaker.

Figure 2. Mean Log-transformed Reaction Time (LogRT) by trial
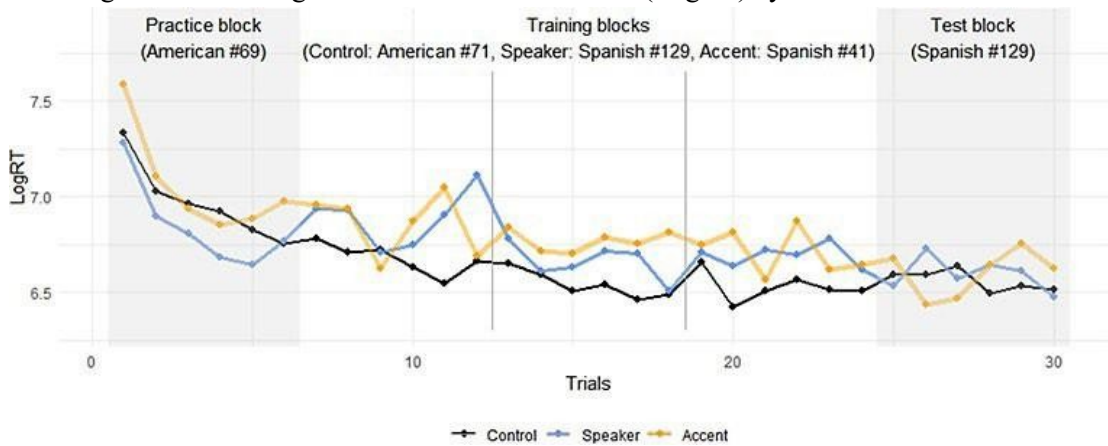
These results suggest that L2 listeners may have much lower sensitivity to changes in accent or speaker, and that overall patterns and processes of talker normalization, adaptation, and generalization may be different in an L2.

Selected References

Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *Journal of the Acoustical Society of America, 116*(6). DOI: 10.1121/1.1815131

Xie, X., Weatherholtz, K., Bainton, L., Rowe, E., Burchill, Z., Liu, L., & Jaeger, T. F. (2018). Rapid adaptation to foreign-accented speech and its transfer to an unfamiliar talker. *Journal of the Acoustical Society of America, 143*(4), 2013. DOI: 10.1121/1.5027410

Second Language Processing of Errors in
Korean-to-English Machine-translated Output

Eun Seon Chung
(Hankyong National University, Korea)

While previous investigations on online machine translation (MT) in language learning have analyzed how second language (L2) learners use and post-edit MT output (Chung & Ahn, 2022; S.-M. Lee, 2020; Lee & Briggs, 2021; Tsai, 2019), no study as of yet has investigated how the learners process MT errors and what factors affect this process using response and reading times. The present study investigates L2 processing of MT errors that are caused by syntactic, morphological, and semantic differences between the source and target language. Notable differences between English and Korean in the use of case markers, tense/aspect, pro-drop, adnominal clauses, and scope of negation were used as morphological and syntactic categories, and semantic ambiguities arising from homonyms and proverbial or idiomatic expressions were used in the semantic category to examine how Korean learners of English analyze MT errors. The present study additionally addresses the following questions: (a) Does L2 proficiency have a moderating effect on how L2 learners process these errors? (b) Does having constant visual access to the source text facilitate or interfere with processing MT output?

Forty-seven Korean L2 learners of English (M age = 23.5; range = 20–29; 33 female and 14 male) who are undergraduate or graduate students at a university in Seoul participated in the study. The main experimental task was an Acceptability Judgment Task (AJT) in which participants read a Korean sentence and then its translated counterpart in English and had to judge the accuracy of the translated sentence on a four-point scale (Very accurate '4,' Accurate '3,' Inaccurate '2,' Very inaccurate '1'). The task consisted of 36 Korean sentences and their machine translated English counterparts, half of which were accurate and the other half inaccurate. The Korean sentences were controlled for length (30–35 characters) and were all embedded complex sentences that contained an independent clause and a dependent clause. To additionally investigate the effect of visual display, the sentences were either shown together on the same screen (synchronous display) or shown separately on consecutive screens (asynchronous display). The accuracy of the translated English sentence (accurate vs. inaccurate) and the display type (synchronous vs. asynchronous) were counterbalanced across participants and items resulting in four counterbalanced presentations lists with 11–12 participants randomly assigned to each list

Statistical analyses of the three dependent variables (accuracy judgment, response latencies, and total reading times) revealed that MT error categories, learners' L2 proficiency, and visual presentation all interact to determine how MT output is processed. Analysis of response latencies—the time taken for the learners to make an accuracy judgment after reading the source and target sentences—indicated that the ways in which learners process MT errors of different categories (morphological, syntactic, and semantic) largely depend on their L2 proficiency level. Advanced learners took longer to respond to items in the semantic category than the other two categories, whereas the less proficient learners spent the least amount of time to respond to items in the semantic category and longest for those in the syntactic category. The effect of proficiency level was also evident in accuracy judgments. Analyses of total reading times—the time taken for learners to read and process the source and target sentences— revealed a significant effect of Display. Overall, the results revealed that (a) learners generally find it harder to reject mistranslations than to accept correct translations, (b) high and low proficiency learners focus on different aspects of language when processing translated output, and (c) constant visual access to

the source text does not facilitate but rather interferes with processing MT errors.

As one of the first studies that have examined L2 processing of MT errors using response and reading times, the present work delineates the manner in which learners employ their linguistic knowledge in L1-to-L2 translation processing and contributes to our understanding of the factors involved in L2 learners' real-time processing of MT errors. With MT being increasingly viewed as a CALL tool that can facilitate language learning, the present findings underline the importance of L2 proficiency, crosslinguistic differences between SL vs. TL, and the visual presentation of texts as significant factors that affect L2 processing of errors in the MT output.

Selected References

Chung, E. S., & Ahn, S. (2022). The effect of using machine translation on linguistic features in L2 writing across proficiency levels and text genres. Computer Assisted Language Learning, 35(9), 2239–2264.

Lee, S.-M. (2020). The impact of using machine translation on EFL students' writing. Computer Assisted Language Learning, 33(3), 157–175.

Lee, S.-M., & Briggs, N. (2021). Effects of using machine translation to mediate the revision process of Korean university students' academic writing. ReCALL, 33(1), 18–33.

Tsai, S.-C. (2019). Using google translate in EFL drafts: A preliminary investigation. Computer Assisted Language Learning, 32(5–6), 510–526.

The Mechanism behind Understanding Anaphoric *One* and Noun Phrase Ellipsis: How Language odels Work

Nayoun Kim, Ziying Li, Sanskriti, Kwangsu Kim
(Sungkyunkwan University, Korea)

Anaphora refers to the anaphoric element that gets its interpretation from a previously mentioned word or phrase (i.e., antecedent) in a sentence or discourse (Aone & Bennett, 1995; Poesio et al., 2010; Mitkov, 2014; Poesio, 2016). Identifying the correct antecedent for an anaphoric element is crucial for understanding the sentence, but this task can be challenging due to various factors such as distance, number agreement, and syntactic structure between the antecedent and the anaphoric element (Linzen et al., 2016; Kim et al., 2019; Linzen & Baroni, 2021). For some types of anaphora (e.g., pronouns and Anaphoric *one*), the antecedents may not always be located within the same sentence as the anaphoric element, further complicating the resolution process (Hankamer & Sag, 1976; Sag & Hankamer, 1984). This linguistic complexity is well illustrated by the Anaphoric *one* and Noun Phrase Ellipsis (NPE) (Kim et al., 2019). Understanding the Anaphoric *one* can be difficult as its antecedent may not occur in the same sentence and may not agree in number (e.g., John has many tables. He decided to sell the red *one*.) This task can be more challenging for NPE, which involves the omission of a noun phrase (NP) following certain determiners (e.g., "Many students like syntax but some [$_{NP}$] do not") (Matthews, 2014; Saito et al., 2008). To understand NPE, one must first recognize that an NP has been elided and then recover the antecedent in the elided NP site (Kim et al., 2019).

Linguistic research has noted that humans show difficulty in understanding Anaphoric *one* and NPE due to these complexities (Kim et al., 2019). In this study, we investigate how well language models can accurately identify the correct antecedents and understand the factors that can potentially affect this process. We employed pre-trained BERT (Devlin et al., 2018) and RoBERTa (Yinhan et al., 2019) models, fine-tuning them for our anaphora resolution tasks. We created five datasets to train the models, each including manually annotated texts containing Anaphoric *one* and NPE from English Wikipedia. While annotating the texts, we specifically marked three elements: the anaphoric elements, the correct antecedent, and another NP option that the models might mistakenly consider as the antecedent (i.e., incorrect antecedent).

- **The General A(naphoric)O(ne) corpus**: Includes 457 texts where either the correct or incorrect antecedent can be close to or far from the anaphoric *one*, or agree or disagree with the anaphoric *one* in the number feature.
  o **The Distance factor A(naphoric)O(ne) sub-corpus**: Includes 143 texts where the correct antecedent is always far from the Anaphoric *one*, intervened by the incorrect one.
  o **The Number factor A(naphoric)O(ne) sub-corpus**: Includes 154 texts where the incorrect antecedent always matches the Anaphoric *one* in number, while the correct one does not.
- **The General NPE corpus**: Includes 391 texts where either the correct or incorrect antecedent can be close to or far from the NPE licensor.
  o **The Distance factor NPE sub-corpus**: Includes 164 texts where the correct antecedent is always far from the NPE licensor, intervened by the incorrect one.

Among the datasets, the General AO corpus and General NPE corpus were created to investigate the models' performance in correctly finding the antecedent of the Anaphoric *one* and NPE. The Distance factor and Number factor AO sub-corpus, and the Distance factor NPE sub-

corpus were created to investigate the influence of the distance and number agreement between the anaphoric element and antecedent on the models' ability to find the correct antecedent.

In addition to our datasets, we also adopted the GAP dataset (Webster et al., 2018) and WSC dataset (Levesque et al., 2012) for the task of Gendered, and Gender-neutral coreference resolution, respectively. The model we created for coreference resolution serves as a baseline, which we then fine-tuned for the Anaphoric *one* and NPE resolution tasks. Figure 1 shows the training architecture across all four tasks.
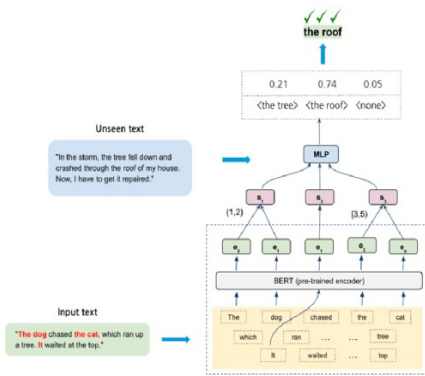
Figure 1. The detailed architecture of the tasks   **Figure 2**. Accuracy of 4 anaphoric elements
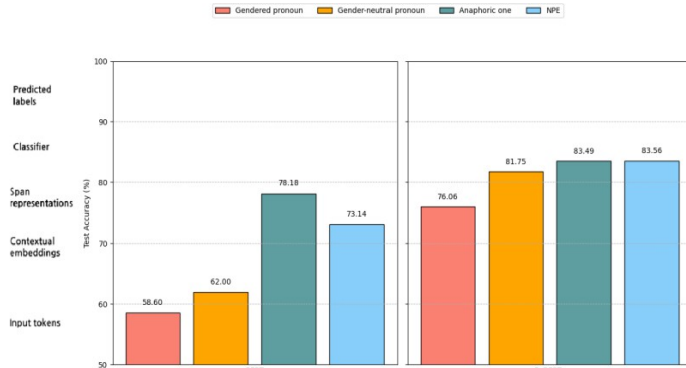
The BERT tokenizer preprocesses raw text, marking possible antecedents (A, B) and the pronoun (P). This text is converted into numerical identifiers and processed by the BERT model, which captures contextual information and converts tokens into vectors. A multi-layer perceptron (MLP) then processes these vectors, with the final layer predicting whether A, B, or neither is the correct antecedent.

Our experimental results demonstrated the effectiveness of our approach in resolving different types of anaphoric elements. As shown in Figure 2, the BERT model achieved over 50% accuracy in all four anaphora resolution tasks, and exceeded 70% accuracy in the Anaphoric *one* and NPE resolution tasks. RoBERTa performed even better, achieving over 70% accuracy in all four tasks, with over 83% in the Anaphoric *one* and NPE resolution tasks. This suggests that the models may perform differently with different training data (our manually annotated corpus included more detailed information than the GAP and WSC datasets). Furthermore, our experiments also revealed that the models' performance was affected by the distance and number agreement between the antecedent and the anaphoric element. As shown in Figures 3 and 4, the models achieved lower accuracy when trained with the distance and number sub-corpora compared to when trained with the general corpora.
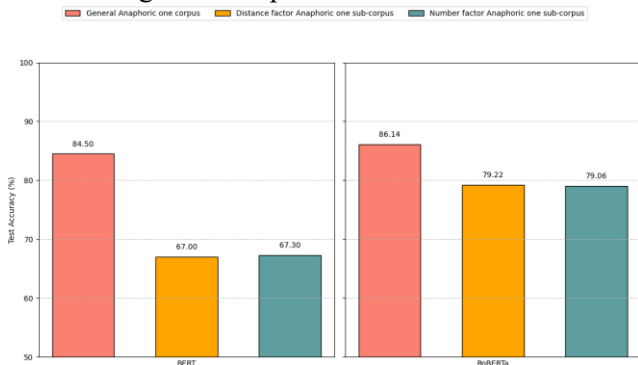
[Figure 3] Accuracy of Anaphoric *one* across 3 (sub-)corpora   **[Figure 4]** Accuracy of NPE across 2 (sub-)corpora
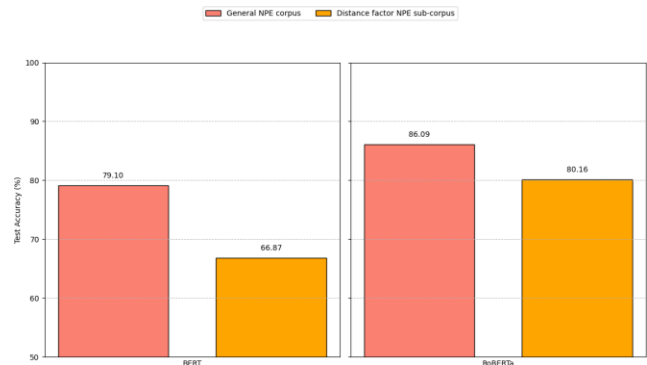
Overall, our study reveals that while LLMs like BERT and RoBERTa can effectively resolve anaphoric dependencies, their performance is influenced by the training data and linguistic

complexity of the anaphora. These findings contribute to the development of more advanced NLP applications that can better understand the complexity of human language.

Selected References

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Kim, N., Brehm, L., & Yoshida, M. (2019). The online processing of noun phrase ellipsis and mechanisms of antecedent retrieval. *Language, Cognition and Neuroscience*, *34*(2), 190-213.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*.

(Withdrawn)

Korean Expletive Negation in a Subordinate Clause:
Focusing on Its Variants

Mija Kim
(Kyung Hee University, Korea)

The major goal of the present study is to gain more insight into Korean expletive negation (KEN) in terms of collostructional analysis within the framework of Construction Grammar. Expletive negation, in essence, refers to a construction wherein the presence of a grammatically acceptable negation marker does not affect the overall meaning of the utterance (or proposition) in which it appears, as illustrated in (1).

(1) kunun   noan-i   oci-**anh**-ul-kka   kekcengha-koiss-ta.
     He-Top   presbyopia-Nom   develop-**Neg**-Fut-NFcomp   fear-Asp-Decl 'He is worried that he might develop presbyopia.' (2021 NEWS)

The phenomenon of expletive negation has been widely examined extensively within Romance languages. Previous research has primarily focused especially on intuitive insights into KEN (Yoon 2009; Choi & Lee 2017; Jin & Koenig 2021). The key issues focus on whether the negator in this construction is expletive and the function of this negator. The literature delves into two viewpoints on this, depending on the meaning of the negator. Either the negator in EN construction does not contribute to its meaning of the total sentence, or it contributes to the negative stylistic features, supporting the claim that it is used as the complement of only predicates denoting a negative meaning. Comprehensively speaking, the literature assumes that this negator serves functions distinct from the regular negator of typical negative sentences.

   Therefore, this study endeavors to scrutinize the syntactic distribution of semantically vacuous negative markers in Korean through corpus-based analysis, thereby providing both quantitative and qualitative analyses supported by theoretical and statistical data. In pursuing this aim, this study addresses three variants of KEN such as *anh-ulci*, *anh-ulkka* and *anh-na* that are used exclusively in an embedded complement clause. This study focuses on two research targets. The first is on whether these three variants function as a construction of their own or if they present variations composed of different morphemes that collectively serve as a singular expletive negation marker. The second is to provide the syntactic properties of KNE as a type (form) of syntactic intrusions, as well as their corresponding pragmatic functions tied to these syntactic properties. Its importance stems from the fundamental perspective of CxG, which views a construction as pairings of form and meaning. Syntactic intrusion in CxG refers to a phenomenon where elements that are not originally part of a certain construction intrude or are inserted into it, often leading to unconventional structures. For example, *If I had've eaten it, I would have died.* This sentence includes a redundant *have*, thereby creating a syntactic intrusion (Fillmore 1985). Thus, this may not fit the traditional rules of syntax but nonetheless convey specific meanings or serve particular communicative functions. Syntactic intrusion introduces a new syntactic pattern that may denote its own construction meaning or, at the very least, convey a specific pragmatic force.
Based on the results, this study suggests that KEN construction is a type of negative construction preferred strategically for avoiding the ambiguity arising from the rhetorical function of its positive clause counterpart. The KEN clause serves as a complement of both nouns and verbs. The analysis of predicates in a main clause, classified into eight types in terms of meaning, highlights the

difference in their semantic functions across three variants of KEN construction. The study identifies the constructional meanings assignable to each variant using the collocation strength imposed on each predicate. Specifically, the EN *anhulkka* construction prototypically denotes an emotional stance of avoidance, the EN *anhulci* construction expresses a neutral stance between avoidance and anticipation, while the EN *anhna* construction conveys an emotional stance of anticipation or acceptance. The study concludes that there is a likely correlation between the semantic functions of main predicates and their embedded clauses.

Selected References

Baauw, S., Drijkoningen, F., & Meroni, L. (Eds.). (2021). *Romance languages and linguistic theory 2018: Selected papers from'Going Romance'32, Utrecht* (Vol. 357). John Benjamins Publishing Company.

Choi, Y., & Lee, C. (2017). Expletive negation and polarity alternatives. In *Contrastiveness in information structure, alternatives and scalar implicatures* (pp. 175-201). Cham: Springer International Publishing.

Delfitto, D. (2020). 'Expletive Negation', in Viviane Déprez, and M. Teresa Espinal (eds), *The Oxford Handbook of Negation*, Oxford Handbooks.
https://doi.org/10.1093/oxfordhb/9780198830528.013.9

Espinal, M. T. (1992). Expletive negation and logical absorption. *The Linguistic Review*, *9*(4), 333-358. Espinal, M. T. (2000). Expletive Negation, Negative Concord and Feature Checking. *Catalan Working Papers in Linguistics*, V. 8 (2000), p. 47-69, ISSN 1132-256X

Gries, S. T, and Stefanowitsch, A. (2004). Extending collostructional analysis: A corpus-based perspective on alternations. *International journal of corpus linguistics* 9(1): 97-129.

Gries, S. T., & Stefanowitsch, A. (2014). Co-varying collexemes in the *into*-causative. In M. Achard & S. Kemmer (Eds.), *Language, culture, and mind. Stanford,* CA: CSLI Publications.

Jin, Y., & Koenig, J. P. (2021). A cross-linguistic study of expletive negation. *Linguistic Typology*, *25*(1), 39-78.

Van der Wouden, T. (1994). Polarity and 'Illogical Negation', In Makoto Kanazawa & Christopher J. Piñón (eds.), *Dynamics, polarity, and quantification*, 17–48. Stanford: CSLI Publications.

Yoon, S. (2013). Parametric variation in subordinate evaluative negation: Korean/Japanese versus others. *Journal of East Asian Linguistics*, *22*, 133-166.
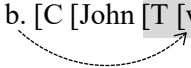
Revisiting Phase-internal Movement

Manabu Mizuguchi

(Kokugakuin University, Japan)

Issue: I-language is a thought-generating system and produces two categories of thought (*propositional* and *clausal*), which is known as Duality of Semantics (DoS). According to Chomsky (2023), DoS naturally follows from two modes of Merge (that is, External Merge, EM and Internal Merge, IM). He further argues that two categories of thought must be segregated in the derivation, saying that segregation is imposed by IM: an element, once moved to a phase edge, is segregated from the ongoing derivation, becoming inaccessible to further operations in the moved position. To see this, consider (1):

(1) What did John read?

What receives information-related interpretation and is in the clausal domain, which is due to IM. In the derivation, *what* is IMed to Spec-v, a phase edge (=(2a)), segregated at the Spec position and subject to consultation by the phase head for interpretation and externalization (=(2b)) (shading indicates segregation):

(2) a. [what$_2$ [v [V read what$_1$]]]
    b. [C [John [T [what$_2$ [v [read what$_1$]]]]]]

This enables *what* to take scope over the clause and spelled out in the vicinity of C as in (1).

    Now consider phase-internal movement or A-movement. Chomsky says that segregation does not occur with IM when it is to a non-phase edge. For instance, consider (3), where *John* moves to Spec-T, a non-phase edge:

(3) [$_\alpha$ John$_2$ [$_\beta$ T [John$_1$ [v [read novels]]]]] (John will read novels.)

In (3), Spec-T must not be segregated: it must be accessible for the purposes of Agree, labeling and anaphora (Chomsky 2023:15). T is too weak to function as a label and requires an accessible Spec-T for labeling (Chomsky 2015). If Spec-T (*John$_2$*) were segregated and became inaccessible, the objects [$_\alpha$] and [$_\beta$] would not be labeled, with Full Interpretation violated.

Problems: I point out that Chomsky's approach has two major problems. The biggest problem is that it virtually assumes two types of IM: the IM that segregates an element and the IM that does not. Given that there's just Merge with two modes of its application (external and internal), a natural assumption will be that IM is IM; that is, there should not be two types of IM.

    Another problem is that the distinction between the two types of IM rests on whether or not IM is to a phase edge. Chomsky tacitly assumes that the Spec of a phase head is a phase edge. This assumption, however, is questionable. For instance, Mizuguchi (2024) argues that given structure building in the Minimalist Program, a phase edge is not a pre-determined position and that it is determined by labeling, proposing (4) (Mizuguchi 2024):

(4) The phase edge is a position immediately contained in a set identified as having a phase property (informally, "CP" and "vP"); the phase complement is a sister to a phase head.

Given (4), when a *wh*-phrase moves to the Spec of a phase head as in (5), it is not immediately contained in *v*P since the whole object is yet to be labeled (as indicated by ?):

(5) [? what$_2$ [v [V read what$_1$]]]

Assuming (4), IM is not to a phase edge and *what* cannot be segregated from the ongoing derivation when IMed to the Spec of *v*. Chomsky says that EM and IM must be segregated and that the segregation is due to IM to a phase edge. However, with (4) in place, IM cannot separate an element from the ongoing derivation when it is IMed to the Spec of a phase head, with the result that there will be no segregation at all.

Proposal: I argue, following Chomsky (2023), that IM establishes segregation of an element: IM carries the derivation from the propositional to the clausal domain, creating an element that has no further interactions with the propositional domain or the ongoing derivation. Given this function of IM, I claim that segregation for clausal thought always occurs upon IM, whether an element is IMed to a phase edge or not; the phase edge is irrelevant to segregation.

This proposal can solve the issues with Chomsky's approach but raises a question on phase-internal movement. Recall that Spec-T must not be segregated; otherwise, labeling failure will result. With IM having a domain-changing function, I further propose that phase-internal movement is structured by EM. In the derivation of (3), *John$_2$* and *John$_1$* are EMed to Spec-T and Spec-*v*, respectively. Spec-T has a secondary semantic role (Chomsky 2023) or a secondary theta role, which I argue makes EM to Spec- T possible under DoS. These two *John*'s are interpreted as one and the same element. The copy relation is ensured by (6):

(6) Structurally identical inscriptions in a c-command configuration must share the same interpretation. (cf. Chomsky 2021 on Form Copy (FC); Friedin 2021 on Staiblity)

John$_2$ and *John$_1$* are structurally identical (X ≡ X), interpreted as a copy pair at the phase level.

Given the proposals here, there is no raising to Spec-T or there is no A-movement. It follows that A and A' correspond completely to EM and IM, respectively, in a one-to-one manner; they can be separated through two modes of Merge. The distinction between A- and A'-positions and between A- and A'- movement is eliminable in favor of ineliminable Merge.

Consequences: The proposed analysis has three consequences. The first consequence is that it can satisfy Minimal Yield (MY), which requires Merge to create only one new accessible object (Chomsky 2021). I will show that MY will be violated if the subject is IMed to Spec-T and then argue that this problem does not arise if Spec-T is structured by EM: in the case of EM, each application of Merge adds only one new accessible element, satisfying MY.

The second consequence is that a Subject Condition violation can be explained, which I argue is due to two structurally non-identical subjects being created in Spec-T and in *v*P/VP by Merge. This has the effect that a copy relation cannot be established between them, and ill-formed interpretation (and wrong externalization) will follow, explaining the ill-formedness of a Subject Condition violation.

The third consequence is that raising and control can be analyzed uniformly. Chomsky (2023) argues that for DoS reasons, raising is due to IM and control is due to EM while Hornstein (1999) argues that IM is involved in both raising and control. Under the proposal in this paper, Spec-T must be structured by EM, which is the only option for labeling or Full Interpretation. In both raising and control, Spec-T is created and it follows from the proposal in this paper that EM is relevant to the creation of the Spec in the two constructions. Unlike Chomsky (2023), raising and control can be given a unified analysis under EM.

Selected References

Chomsky, Noam. 2015. Problems of projection: Extensions. In *Structures, strategies and beyond*, ed. by Elisa Di Domenico et al., 3-16. Amsterdam: John Benjamins.

Chomsky, Noam. 2021. Minimalism: Where are we now, and where can we hope to go. *Gengo Kenkyu* 160, 1-41.

Chomsky, Noam. 2023. The miracle creed and SMT. Ms., University of Arizona. Freidin, Robert. 2021. The strong minimalist thesis. *Philosophies* 2021, 6. 97.

Hornstein, Norbert. 1999. Movement and control. *Linguistic Inquiry* 30, 69-96.

Mizuguchi, Manabu. 2024. A'-movement: A single shift or multiple shifts? Paper presented at the 2024 Winter Conference of the Korean Association of Language Sciences (KALS), Dong-A University.

On the Syntactic Category of *Rúguǒ* 'if' and
Its Implications for Clausal Structure in Mandarin Chinese

Pin-Wei Li, Jen Ting

(National Taiwan Normal University, Taiwan)

This paper investigates the syntactic roles of *rúguǒ* 'if' in Mandarin Chinese (henceforth Chinese), focusing on whether it can function as a prepositional head (P) in addition to serving as a complementizer (C). Notably, *yīnwèi* 'because' can serve either as a P or as a C (Lu 1980, Pan & Zhu 2023). While *rúguǒ* 'if' typically introduces a clause, it intriguingly appears to introduce a nominal object *dìzhèn* 'earthquake' in (1), similar to how *yīnwèi* 'because' functions in (2).

(1)	Rúguǒ	dìzhèn,	tā	huì		wújiākěguī.
	if	earthquake	he	will		homeless

'In case of an earthquake, he will become homeless.'

(2)	Yīnwèi	dìzhèn,	yǒu	rén	wújiākěguī.
	because	earthquake	have	person	homeless

'Because of an earthquake, someone may be homeless.'

We argue that unlike *yīnwèi* 'because,' *rúguǒ* 'if' cannot function as a P. Instead, we propose that *rúguǒ* 'if' in (1) remains a C, introducing a clause with a null expletive subject. This analysis has significant implications for the clausal structure of Chinese, which we will also address. We consider two possible analyses of the construction *rúguǒ dìzhèn* in (1): *rúguǒ* as a P or as a C. If *rúguǒ* 'if' were a preposition taking the event noun *dìzhèn* 'earthquake' as its object, it would be problematic why it cannot take regular nouns as objects, unlike *yīnwèi* 'because,' as shown by the contrast in (3).

(3)	*Rúguǒ / Yīnwèi/	zhège	měilì	de	huánjìng,	tā	měi-tiān dōu
	if / because/	this	beautiful	DE	environment	he	every-day
	hěn	kāixīn.					
	DOU very	happy					

'Because of / In case of this beautiful environment, he is happy every day.'

Thus, it is not plausible to treat *rúguǒ* 'if' as a preposition in (1). If *rúguǒ* 'if' is a complementizer, we need to examine the structure of this CP. Given that Chinese allows for 'single-word sentences' (Liu et. al 1983) (cf. 'fragment as fragment' in Li & Wei 2023), it might be possible that *dìzhèn* 'earthquake' introduced by *rúguǒ* 'if' in (1) is a fragment.

However, we argue against this proposal. 'Single-word sentences' are licensed under certain conditions, such as special intonations (Chen 1989), which are not required in (1). Furthermore, complementizers typically select a TP complement (Benmamoun 2000), so we reject the structure where *rúguǒ* 'if' takes a fragment as its complement.

We propose that *rúguǒ* 'if' in (1) selects a full-fledged clause, where *dìzhèn* 'earthquake' functions as a verb and the subject position is occupied by a null expletive, similar to the English expletive *it*. That *dìzhèn* 'earthquake' can serve as a verb is supported by sentences like (4).

(4) Zuótiān   dìzhèn-le   sān   cì. yesterday   earthquake-LE three
     times There were three earthquakes yesterday.'

The underlying structure of (1) is (5). To account for the interpretation of an earthquake occurring, we posit the presence of a light verb OCCUR (Lin 2001) in the structure, to which the event noun *dìzhèn* 'earthquake' raises, resulting in the verb *dìzhèn* 'earthquake.'

(5) *Rúguǒ*   [TP pro [v OCCUR] *dìzhèn*]
     if   earthquake
     'If there is an earthquake'

In summary, we argue that unlike *yīnwèi* 'because,' which can function either as a C or as a P, *rúguǒ* 'if' remains a C in Chinese. In constructions where *rúguǒ* 'if' seems to take a nominal object, the clausal structure involves a null expletive subject and a verb derived from head movement into a light verb. This study contributes to our understanding of the syntax of light verbs and null expletives in Chinese.

Selected References

Benmamoun, Elabbas. (2000). *The feature structure of functional categories.* Oxford: Oxford University Press.

Li, A., & Wei, T. (2023). Sentence fragment ellipsis in Chinese. *Oxford Research Encyclopedia of Linguistics.*   Advance   online   publication. https://doi.org/10.1093/acrefore/9780199384655.013.1059

Lin, T.-H. J. (2001). Light verb syntax and the theory of phrase structure. Doctoral dissertation, University of California, Irvine.

Liu, Y., Pan, W., & Gu, W. (1983). *Shiyong xiandai Hanyu yufa* [Practical modern Chinese grammar].
     Beijing: The Commercial Press.

Pan, V. J., & Zhu, B. (2023). On the syntax of causal clauses in Mandarin Chinese. In Ł. Jędrzejowski & C. Fleczoreck (Eds.), *Micro- and macro-variation of causal clauses: Synchronic and ciachronic insights* (pp. 221-249). John Benjamins. https://doi.org/10.1075/slcs.231.08pan

Function of the Diminutive in Narrative of Tunis Arabic

Taku Kumakiri

(Shukutoku University, Japan)

1. Introduction

The diminutive is "any morphological device that means at least 'small'" (Jurafsky 1996: 534). This presentation focuses on its function in stories written in Tunis Arabic.

2. Previous research on the diminutive of Tunis Arabic

Tunis Arabic is one of the modern Arabic dialects. It is spoken in Tunis, the capital of Tunisia in North Africa (Singer 1984, Gibson 2009).

Morphological and semantic studies have been conducted on the diminutive in Tunis Arabic (Singer 1984: 480-496; Procházka 2017). However, little attention has been paid to its function in storytelling.

This situation is due to the conversational nature of the diminutive. Singer (1984: 481) notes that the diminutive is used in friendly or polite conversation. Therefore, it is not surprising that some might conclude that the diminutive in storytelling is limited to the conversational parts rather than the narrative (descriptive) parts. As a result, the use of diminutives in storytelling has yet to be a major research focus. However, diminutive forms appear in the narrative parts of stories in Tunis Arabic. For example, in a story called *it-taːʒir l-akħal* ('the black merchant'), the diminutive *sliːʕa* ('small merchandise,' the diminutive of *silʕa* 'merchandise') is used three times exclusively in the narrative part. In comparison, *silʕa* is used both in the narrative part (six times) and the conversational part (five times). This suggests that the diminutive in Tunis Arabic may have a specific function in the narrative parts of stories.

3. Formation of the diminutive and its meaning

According to Singer (1984: 482) and Procházka (2017), Tunis Arabic has three main diminutive patterns: CCiːCa, CCajjiC (CCajjCa), and CCiːCiC (CCiːCCa). For example, *ħofra* ('hole'), *kalb* ('dog'), and *tiskra* ('ticket') become *ħfiːra*, *klajjib*, and *tsiːkra*, respectively.

The diminutive primarily denotes smallness, as seen in (1) (DIM for diminutive). Additionally, it is pragmatically used to express affection (i.e., caritative, as seen in (2)) or to attenuate the speaker's request (as shown in (3)).

(1) ʃnuːwa   haːk-il-ɣdiːr, *ħfiːra*   sˤɣiːra   fiː-qaːʕ-haː   blaːtˤa what   that-the-
     cave hole.DIM   small   in-bottom-its stone-plate "What is this cave?
     It is a small hole, and its bottom is a stone plate."

(2) *wxajj*-naː   wsˤul   l-haːk-il-blaːd brother.DIM-our he.arrived   to-that-the-
     country
     "Our brother (our main character) had arrived in that country."

(3) luːkaːn   taʕtˤiː-niː   *ʒmajjil*   aːxur
     if   you.give-me camel.DIM   another "What about giving me another
     camel?"

4. The diminutive in the narrative

The diminutive has several functions in the narrative. It can describe smallness (as in "small hole" in (1)) or express affection for a character from the perspective of the speaker and the audience (as in "our brother" in (2)).

However, one of the most important uses of the diminutive in the narrative part is to introduce an object that is somehow related to the main character and plays a specific role in the plot. See example (4) below.

(4) ***fħajjil***   fiː-maktuːb-uː.   juxruʒ   l-il-ɣaːba,
                                                                             juqʕu
                                                                       d
     flute.DIM   in-pocket-his   he.goes.ou   to-the-forest he.sits
                                     t

     taħt-ʃaʒra   w-jiʒbid   haːk-il-***fħal***   w-jibdaː   jtˤawwiħ. under-tree   and-
     he.pulls.out   that-the-flute and-he.starts he.plays
     "A flute was in his (the prince's) pocket. He used to go out to the forest. He used to sit under a tree, pull out that flute, and start playing."

In (4) above, a flute is introduced into the story by the diminutive form ***fħajjil***. The flute belongs to a prince, the main character. Because of his enthusiasm for the flute, his father will expel him from his country.

It is also noteworthy that in the following sentence, the normal form ***fħal*** is used to indicate the same flute of the prince. Thus, the prince's flute is not small. Therefore, the diminutive form in (4) does not describe the smallness of the flute; rather, it signals that the newly introduced flute is related to the prince and plays an important role in his story.

The diminutive's introductory function in the narrative stems from its caritative use. Since the focus of affection naturally attracts the audience's attention, the diminutive introduces an object as a key element in the narrative.

5. Conclusion

This presentation studies the use of diminutives in the narrative in Tunis Arabic. The conclusion is that the diminutive in the narrative introduces a new object that is somehow related to the main character and has a specific role in the plot.

References

Gibson, Maik 2009. "Tunis Arabic." In: Versteegh, Kees, and Mushira Eid, Alaa Elgibali, Manfred Woidich, and Andrzej Zaborski (eds.) 2009, 563-571.

Jurafsky, Daniel 1996. "Universal Tendencies in the Semantics of the Diminutive." *Language*. Vol. 72, No. 3, pp. 533-578. Linguistic Society of America (https://www.jstor.org/stable/416278)

Procházka, Stephan. 2017. "The Diminutive in the Arabic Dialect of Tunis." (Studies on Arabic Dialectology and Sociolinguistics. Proceedings of the 12th International Conference of AIDA held in Marseille from May 30th to June 2nd, 2017) (https://books.openedition.org/iremam/4017)

Singer, H-R. 1984. *Grammatik der Arabischen Mundart der Medina von Tunis*. Berlin/New York: Walter de Gruyter.

Versteegh, Kees, Mushira Eid, Alaa Elgibali, Manfred Woidich, and Andrzej Zaborski. (eds.) 2009. *Encyclopedia of Arabic language and linguistics*. Vol. IV. Leiden/Boston: Brill.

Japanese *Yabai* with Empathy

Ahbi Koh

(Kobe College, Japan)

This pilot study focuses on the neutral meaning. To elucidate that, the relationship between empathy and the psychological distance is discussed.

In this pilot study, Japanese adjective *yabai* is taken up. Japanese adjective *yabai* has the original negative meaning of *bad*. However, Japanese native speakers use not only the negative meaning but also positive meaning of good and neutral meaning.

(1) A: Kyono eigono tesuto yabai. Mattaku benkyo shitenai.
     " Today's English exam will be bad/ good. I didn't study at all.
     B: Watashimo yabai. Benkyo shitenai.
     " I will be bad/ good. I didn't study.
(2) A: Kono Kukki oishii. *Yabai.*
     " This cookie is yummy. ~~Bad~~/ good."
     B: Oishisugiru. *Yabai.*
     " Too yummy. ~~Bad~~/ good."

Yabai is used with the original meaning in (1).

However, if *yabai* is used with the original meaning, the sentence is not accepted. It is clear that *yabai* can be used with the meaning of the positive meaning in (2).

(3) A: Ano hito mite. *Yabai.*
     " Look at that person. Bad/ good."
     B: Sugoi. *Yabaine.*
     " Wow, Bad/ good."

The conversation (3) demonstrates that *yabai* has the neutral meaning.

According to examples (1)-(3), speakers can use *yabai* with any meanings. The author argues that there are no words which have the negative, positive and neutral meanings in all languages.

However, these neutral meanings in (3) needs the psychological distance and empathy.

The survey has been done to affirm to clear the neural meaning. The informants, all Japanese native speakers, were asked about the frequency of its use in the sentence. Then, the other question is to choose one from the positive or negative meaning.

As a result, *yabai* with the neutral meaning is acceptable for Japanese native speakers.

The author finds that when Japanese native speakers use *yabai* which is used with the neutral meaning in the sentence, the speaker share ideas and feelings with the hearer at the utterance time. This phenomenon shows that the speaker thinks the psychological distance is close with the hearer. Then, the speaker unconsciously believes that if the speaker uses *yabai* with the positive meaning

or neutral meaning, the hearer can understand it.

When the speakers use *yabai* which is used with the original negative meaning, they do not emphasize with the hearers.

However, when the speakers use *yabai* with the nonoriginal meanings, they want to emphasize with the hearers very much. Then, the psychological distance is close between the speaker and the hearer.

The psychological distance and empathy affect *yabai* with the nonoriginal meaning.

This study explores into the semantic change of *yabai* concerning the pragmatic notion psychological distance and empathy.

Selected References
Kamio, Akio. 1990. "Jyohouno Nawabari Riron (Territory of Information)" Tokyo: Taishukan
        Syoten. Kamio, Akio. 1997. *Territory of Information*. Amsterdam: John Benjamins .
Koh, Ahbi. 2016. *"Psychological Distance" as a Linguistic Primitive: Case Studies in Korean in
        Contrast with Japanese and English.* Doctoral Dissertation, Kobe College.
Koh, Ahbi. 2021. The mechanism of the semantic shift in Japanese yabai and English awesome.
        The 25th PAAL International Online Conference Proceedings, 18.
Kuno, Susumu. 1978. "Danwa no bunpo (Grammar of Conversation)." Tokyo: Taishukan Syoten.
Kuno, Susumu. 1987. Functional Syntax: Anaphora, Discourse, and Empathy.
Chicago: University of Chicago Press.
Levinson, S.C. 1983. *Pragmatics*, Cambridge University Press, Cambridge

Human Processing of Animal Metaphors in English and Chinese

Ben Yue Shi Yang
(National Chung Cheng University, Taiwan)
Jenny Yichun Kuo
(National Chiayi University,Taiwan)

This research is a research based on Yang and Kuo's study in 2024. They conducted a research on animal metaphors in English and Chinese, but lacked to find out the reason of how participants rate those animal metaphors as positive or negative. In this research, we tried to build up the gap left by

them. Two kinds of animal metaphors, which are ox (*niu*, "牛) and dragon (*long*, "龍"), are used as the main metaphors to test people's perception. We selected the two animals as one being a prototypical

animal in both cultures, and dragon is an imaginary creature in both cultures. Data were collected from the Academia Sinica Balanced Corpus of Modern Chinese (ASBC) and Corpus of Contemporary American English (COCA). We selected the concordance lines from the two corpora and turned it into a questionnaire form. Fifteen participants have to rate from 1 to 5 on a Likert scale, where 1 means the most negative, and 5 means the most positive. After that, participants have to give reasons of why they mark this score. Table one shows the concordance lines of ox, bull, buffalo, cow, dragon, niu (牛), and long (龍). Table two shows the mean and standard deviation of the animal metaphors.

Table 1: Concordance lines of ox, bull, buffalo, cow, dragon, niu (牛), and long (龍)

(1) 一片經濟復甦聲中，亞洲四小龍在新的一年將孰盛孰衰。

(Chinese Pinyin: Yipian jingji fususheng zhong, yazhou sixiaolong zai xinde yinian jiang shushengshuai.)

(English translation: Amidst the voice of economic recovery, the four little Asian dragons will thrive in the economy and will also decline.)

(2) 末天良來到茶館，瞧瞧他的模樣，嘻嘻。真是一朵鮮花插在牛糞上。

(Chinese pinyin: Motianliang laidao chaguan, qiaoqiao tade moyang, zhenshi yiduo xianhua chazai niufen shang.)

(English translation: Motianliang came to the teahouse. Look at his appearance! Ha ha! Isn't he like an ordinary man married to a beautiful woman?)

(3) Holy *cow*, I need a life.

(4) He's as stubborn as an *ox.*

(5) This *Buffalo* woman told McCain she had breast cancer.

(6) They think they were doing the best for me, but that's a *bull* crap.

(7) I am worried for my younger brother. My mom's getting angrier and angrier like a fire- breathing *dragon.*

Table 2.: The mean and standard deviation of the animal metaphors

Animal metaphors	Niu (牛)	Long (龍)	ox	cow	buffalo	bull	dragon
Mean score	2.92	3.46	3.3	2.4	3.42	2.83	3.56
Standard deviation	0.65	0.63	0.9	0.83	0.69	0.89	0.28

Both Chinese and English speakers recognize positive traits like strength and reliability in ox metaphors (e.g., built like an ox, cash cow, "*jiu niu er hu de li chi*" 九牛二虎的力氣 for immense strength) and negative traits like stubbornness and filth (e.g., stubborn as an ox, bull in a China shop, "*niu fen*" 牛糞 cow dung). However, language is a vehicle to carry nuance interpretations, particularly with negative expressions like having a cow (overreaction) and paper cows (deception). For dragon metaphors, cultural distinctions are evident: Chinese dragons "*long*" 龍 symbolize power and prosperity (e.g., "*long ma jing shen*" 龍馬精神 - the spirit of the dragon and horse), while in English, dragons typically symbolize danger and evil (e.g., slay the dragon). Humans emphasize cultural differences, which demonstrates deeper cultural understanding and contextual awareness.

The findings reveal that humans recognize themes of power and grandeur in metaphors, but their evaluations differ significantly. Humans provide interpretations grounded in cultural contexts and personal experiences. Humans differentiate metaphors by drawing on their cultural backgrounds, situational contexts, and personal experiences. They interpret metaphors through a combination of cultural knowledge and context-specific understanding, which allows them to discern subtle differences and meanings. This approach enables humans to capture the emotional and connotative layers of language that vary with cultural and situational contexts, resulting in a more nuanced and comprehensive understanding of metaphors.

Keywords: animal metaphors, human perception, English, Chinese, cultural dimensions

Selected References
Chinese Knowledge and Information Processing. (1994). *Academia Sinica Balanced Corpus of Modern Chinese.* Taipei City, Taiwan.

Davies, Mark. (2008-) *The Corpus of Contemporary American English (COCA)*. Available online at https://www.english-corpora.org/coca/.

Yang, Ben Yue Shi and Kuo, Jenny Yichun (2024). *Conceptual Differences of Animal Metaphors Between English and Chinese*. Paper presented in The 11th Conference on Language, Discourse and Cognition. National Taiwan University, Taipei, Taiwan.

A Study on the Output Type of Korean Particle '-의[ɰi]'
in Automatic Speech Recognition (ASR)

Kyung Eun Lee

(Seoul National University, Korea)

In many cases, Korean particle '-의[ɰi]' is phonetically realized as [ㅔ(e)].

Based on the fact that in the prior discussion, the Korean particle '-의[ɰi]' can be pronounced as either [ㅢ(ɰi)] or [ㅔ(e)], it is necessary to determine which proportion of [ㅢ(ɰi)] and [ㅔ(e)] are pronounced in actual speech. For this purpose, a total of 6,613 recited speech data (excluding 427 missing files) from 40 men and women ages of 20′s, for 176 word segment containing the particle '-의[ɰi]' were collected from the *National Institute of the Korean Language*. 2003. "Recited Speech Corpus of Seoul speakers". Each was listened and transcribed three times, and the results are as follows.

(1) Pronunciation aspects of the particle '-의[ɰi]'

Pronunciation aspects of the particle '-의[ɰi]' total of 6,613 voice datas, 40 men and women ages of 20′s		
[ㅢ(ɰi)]	845	About 12.8%
[ㅔ(e)]	5256	About 79.5%
Miscellaneous and errors	512	About 7.7%

As we can see, 79.5% of total pronounce the particle '-의[ɰi]' as [ㅔ(e)]. This rate is likely to increase further in spontaneous speech. We will discuss about output types of the Korean particle '-의[ɰi]' when it realized as [ㅔ(e)] in ASR, and we will also consider about problems of those output types.

In this study, we will use ASR that reflects the language model. That means Even if the voice input type does not match the grammar, the output type is derived through uniform correction according to the language model.

(2) **STT (Sound to Text) programs used for experiments**
① NAVER's CLOVER Speech
voice file input → End Point Result
② SAMSUNG ELECTRONIC's Bixby
voice detection → Partial Result → End Point Result

The output types of sentences containing the particle '-의[ɰi]' in ASR are as follows.

(3) **Output Types of sentences containing the particle '-의[ɰi]' in ASR**

① NAVER's CLOVER Speech

Output Types in CLOVER Speech: Voice input [ㅔ(e)] for the particle '-의[ɰi]'

   a. Voice input [ㅔ(e)] / End Point Result '-의[ɰi]'
   b. Voice input [ㅔ(e)] / End Point Result '-에[e]'

② SAMSUNG ELECTRONIC's Bixby

Output Types in Bixby: Voice input [ㅔ(e)] for the particle '-의[ɰi]'

   a. Voice input [ㅔ(e)] / Partial Result '-의[ɰi]' / End Point Result '-의[ɰi]'
   b. Voice input [ㅔ(e)] / Partial Result '-에[e]' / End Point Result '-의[ɰi]'
   c. Voice input [ㅔ(e)] / Partial Result '-에[e]' / End Point Result '-에[e]'

In the STT program (NAVER CLOVER Speech and SAMSUNG Bixby), when the voice type is [ㅔ(e)] and End Point Result appears as '-에[e]' without correction according to the ASR language model(3-①-b, 3-②-c) there are syntactic structure differences, and also there are meaning differences. This problem is even more serious in sentences which either the particle '-의[ɰi]' or '-에[e]' is allowed.

Selected References

Areum Kim. 2022. "A Phonological Study on the Error Types of Korean Automatic Speech Recognition", Korea: Ph.D. Thesis of Seoul National University.

Sangheun Song. 2022. "Using the Deep Learning Models and Surprisal for Language Analasis", *Journal of AI Humanities* 12. 169-191.

Evaluating Unsupervised Parsing: DIORA and the Quest
for Structural Understanding in Natural Languages

Jieun Kim

(University of Ulsan, Korea / University of California, Santa Cruz)

This study emerges from a curiosity about the extent to which an unsupervised parsing algorithm comprehends the structural nature of natural languages. Traditionally, parsing algorithms trained with a gold-labeled tree structure acquire a direct understanding of language structure, benefiting from explicit patterns which neural network models excel at recognizing. In contrast, unsupervised training necessitates the derivation of structural information solely from the textual arrangement and semantic interrelations among words. This exploratory study posits a critical view, rooted in linguistic theory, that true language structure embodies a deep hierarchical organization and recursiveness not readily apparent in surface-level patterns. Thus, it questions whether unsupervised models, devoid of explicit structural annotations, can genuinely apprehend the essence of linguistic structure or merely resort to semantic chunking predicated on word meanings.

The heart of this research unfolds in the experiment section, where we present a novel experiment designed to assess the structural discernment capabilities of DIORA (drozdov et al. 2019) parsing. By comparing the parsing outcomes of DIORA with those derived from a manually constructed Context- Free Grammar (CFG) and analyses performed by a linguist, this study endeavors to distinguish between genuine structural understanding and mere semantic chunking. Through this comparative lens, we aim to shed light on the true extent of structural comprehension achievable by unsupervised parsing algorithms.

The main experiment investigates the hypothesis that DIORA performs "semantic chunking" rather than true "syntactic structural parsing." This leads to a pivotal question: what happens when DIORA is presented with sentences that are structurally impeccable but semantically nonsensical? Will it maintain a similar level of performance? To address this, I devised an experiment. I have created synthetic data of adequate length involved prompting ChatGPT to generate sentences of similar types. I ensured that the structural patterns were repetitive and simple, aiming for the model to learn straightforward patterns rather than complex ones. The grammar rules employed were limited in number. In the examples, each sentence follows a consistent pattern where the verb phrase (VP) comprises a single verb accompanied by one adverbial phrase and two prepositional phrases (PPs). In the correct syntactic parsing, the primary focus is to determine the level at which these two prepositional phrases are attached.

Upon reviewing DIORA's parsing results, I have identified three general observations: First, recursiveness is a defining characteristic of natural language structure, yet DIORA's parsing mechanism fails to capture this aspect of recursiveness. Second, there is a noticeable inconsistency in parsing "meaningless" elements, such as determiners and prepositions (e.g., 'of,' 'a,' 'the,' 'to'), indicating a lack of understanding of their roles. Third, there is a variation in combinatory patterns based on the length of constituents. For example, a verb may combine with a direct object on its right or a subject on its left, contingent upon their lengths. This suggests that the verb does not consistently recognize its inherent structural relationships. However, assessing the parsing quality through verbal descriptions has its limitations. Only linguists, particularly syntacticians, can fully appreciate the extent of the parsing inaccuracies. This level of critique is not readily accessible or apparent to a general audience. So, I have measured the precision, recall, and f-1 score for the parsing results of DIORA.

Here, I have evaluated the precision, recall, and F-1 score of the DIORA parsing results in comparison to the gold standard tree structure crafted by a linguist.

Figure1. Evaluation F1 Result

```
gold_tree_file_path = "GOLD_LABEL_Parse.txt"
parsed_tree_file_path = "DIORA_chatgpt_text_parse_tree.txt"

# 파일에서 트리 읽기
gold_tree = read_tree_from_file(gold_tree_file_path)
parsed_tree = read_tree_from_file(parsed_tree_file_path)

# 평가 실행
precision, recall, f1 = evaluate(gold_tree, parsed_tree)
print(f"Precision: {precision:.4f}, Recall: {recall:.4f}, F1 Score: {f1:.4f}")
Precision: 0.1818, Recall: 0.0645, F1 Score: 0.0952
```

The precision is measured at 18%, recall at merely 6\% and the F-1 score is calculated to be 9.5. These statistics are significantly low, indicating that the parsing results from DIORA do not align with the gold-standard labels. Nonetheless, these figures, in isolation, do not provide a comprehensive understanding of the impact that 'removing semantic content from a sentence' has on the parsing process. The evaluation results presented here are derived from synthetic data generated by ChatGPT. These results are juxtaposed with the evaluation outcomes using the standard validation dataset referenced in the original DIORA paper. I established the sentence structures in the DIORA validation set as gold- labels, employing the same methodology described

This difference suggests that DIORA's parsing might lean more towards semantic chunking rather than syntactic parsing, although further investigation with a larger dataset and statistical validation is required to support this assertion. The findings hint at the feasibility of initiating research projects of this nature. However, whether syntactic parsing can attribute to language model development is a separate inquiry.

Selected References

Andrew Drozdov, Subendhu Rongali, Yi-Pei Chen, Tim O'Gorman, Mohit Iyyer, and Andrew McCallum. 2020. "Unsupervised Parsing with S-DIORA: Single Tree Encoding for Deep Inside-Outside Recursive Autoencoders" In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 4832–4845, Online. Association for Computational Linguistics.

Nilay Patel, Jeffrey Flanigan. 2023. "Forming Trees with Treeformers". *RANLP 2023*, arXiv:2207.06960.

A Construction-Based Approach to
the English *Is-it-just-me-or-X?* Construction

Ji-in Kang

(Hankuk University of Foreign Studies, Korea)

This study aims to provide a Construction Grammar account of the English *Is-it-just-me-or-X?* construction and analyzes it from a usage-based perspective (Goldberg 1995, Hilpert 2014, Hoffmann 2022). By critically reviewing the limitations that a previous study of the *Is-it-just-me-or-X?* construction has, this paper focuses on its constructional constraints generalized from the attested examples from the Now Corpus (Davies 2016-).

  (1) Is it just me, or are headphones the most popular accessory people are wearing nowadays?
  (2) Is it just me or Leo looks like he has a face inside his face. He still looks like a kid!
  (3) Is it just me or why is Khalid Latif not been selected for the national team?

In example (1), the speaker does not seek for an answer to the question whether it is s/he or a headphone that is the most popular accessory people wear; rather, the speaker intends to ask if an addressee agrees with his or her thought that headphones are the most popular accessory (hence, it could be paraphrased as, 'Am I the only one who thinks that headphones are the most popular accessory people are wearing nowadays'). The speaker ultimately conveys her judgment that people treat their headphones not as equipment for listening to music but as an accessory. Example (2) conveys the speaker's assessment that Leo still has his childhood look on his face. Example (2) differs from (1) in that the clause following the conjunction *or* is in a declarative form (*Leo looks like he has a face inside his face*) whereas the preceding clause is in a polar interrogative (*Is it just me*), making the two conjuncts incompatible as alternatives. The examples (1) and (2) include the polar interrogative *is it just me*, but the speakers do not actually seek answers from the addresses; rather they express their subjective judgments on the situations. In this regard, the first clause of each example is a rhetorical question whose the primary purpose is to convey the speaker's strong belief. Example (3) is different from examples (1) and (2) in that a *wh*-question follows the *is it just me or* sequence and the implied predicate is not *think*; it can be paraphrased as 'Am I the only one who does not understand why Khalid Latif has not been selected for the national team or do you not understand either?' In (3), the clause following the conjunction *or* does not directly concern the speaker's thoughts that Khalid Latif should have been selected for the national team. By asking the reason for failing Latif as a member of the national team, the *wh*-question entails that Latif did not make it to the national team, which the speaker did not expect. The current study revisits Bai's (2014) work of the *just me* construction. First, the term *just me* construction is subject to criticisms because it fails to capture the construction's syntactic and functional properties. To reflect the interrogative and rhetorical nature of the construction, the current study names it as *Is-it-just-me-or-X?* construction.

Second, Bai (2014) does not fully explain the components of the construction. For example, Bai's explanation on the conjunction *or* focuses on the opposite situations of a hearer's agreement and disagreement with the speaker's thought. However, it is not the hearer's agreement/ disagreement that motivates the use of the conjunction *or*. The current study argues that the interrogative form of the construction indicates that a speech-act conjunction (Sweetser 1990: 95) is involved in the construal. Third, the amount of data used in the previous study is insufficient for generalizing the constraints of the construction. Bai's study extracted 92 tokens from the Corpus of Contemporary American English and the Corpus of Historical American English, but the size of the dataset seemed not enough to analyze the 12 types of the construction suggested in the study.

The current study collected a total of 4,405 attested instances from the Now Corpus, allowing for a detailed examination of the construction's elements and their modifiable slots. The dataset reveals that the most frequent construct of the *Is-it-just-me-or-X?* construction consists of the *Is it just me* conjunct coordinated with a polar interrogative by the conjunction *or*. This study argues that it is a case of a syntactic amalgam (Brenier and Michaelis 2005) combining the CLEFT and ALTERNATIVE QUESTION constructions. The ALTERNATIVE QUESTION construction typically takes the form of *A or B*, where the expected answer is to pick one of the offered alternatives (Huddleston and Pullum 2002: 868). However, the pairs of the conjuncts in (1)-(3) do not seem compatible as alternatives. The recovered form of the *Is-it-just-me-or-X?* construction would be *Is it just me who thinks that ... or do you also think that ...?* In this regard, the ALTERNATIVE QUESTION involved in the *Is-it-just-me-or-X?* construction is a case of SHARED COMPLETION (Hilpert 2014: 54), where a common part (*who thinks that ...*) and a parenthetical structure (*do you also think that*) are omitted. The IT-CLEFT construction often includes information in the non-focus position of the relative clause that is evident to the speaker but not to the hearer, ensuring that the information cannot be challenged ('pragmatic accommodation,' Lambrecht 2001: 485). In the paraphrase of example (1), 'Am I the only one who thinks that headphones are the most popular accessory people are wearing nowadays?' the speaker goes beyond merely asking whether he or she is the only one who believes the content of the relative clause; the speaker expresses his/her strong belief. This reflects the function of pragmatic accommodation of the IT-CLEFT construction and demonstrates how the IT-CLEFT, which typically involves non-disputable information, is used in an interrogative form to express the speaker's viewpoint to the hearer.

Data sources

(1)   : https://www.yahoo.com/lifestyle/amazon-just-secretly-dropped-70-203400381.html

(2)   : https://www.mcall.com/opinion/mc-ent-faith-beckwith-0710-20210708- wemgbdeu7zfstl57ac5z3o3nhu-story.html

(3)   : http://www.espncricinfo.com/icc-world-twenty20-2012/content/story/571945.html /

Selected References

Bai, Yinchun. 2014. "A Usage-based Study of the *Just Me* Construction," *Yearbook of the German cognitive linguistics association* 2(1), 127-146.

Brenier, Jason M., & Michaelis, Laura A. 2005. "Optimization via Syntactic Amalgam: Syntax-prosody Mismatch and Copula Doubling," *Corpus Linguistics and Linguistic Theory* 1, 45-88.

Davies, Mark. 2016-. "Corpus of News on the Web (NOW)," Available online at https://www.english- corpora.org/now/.

Goldberg, Adele E. 1995. "Constructions: A Construction Grammar Approach to Argument Structure," University of Chicago Press.

Hilpert, Martin. 2014. "Construction Grammar and its Application to English," Edinburgh University Press.

Hoffmann, Thomas. 2022. "Construction Grammar: The Structure of English," Cambridge University Press.

Huddleston, Rodney, & Pullum, Geoffrey. 2002. "The Cambridge Grammar of the English Language," Cambridge: Cambridge University Press.

Lambrecht, Knud. 2001. "A Framework for the Analysis of Cleft Constructions," *Linguistics* 39(3), 463-516.

Sweetser, Eve. 1990. "From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure," Cambridge University Press.

Identifying Sortal Classifiers in Korean

Ching-Syuan Shen
(National Chengchi University, Taiwan)
One-Soon Her
(National Chengchi University / Tunghai University, Taiwan)

Korean is a well-known numeral classifier language. Similar to the count/mass noun distinction, numeral classifiers consist of two subcategories: sortal classifiers (C) and mensural classifiers (M). When quantifying nouns, one must use either a C, as in (1), or an M, as in (2).

(1) *oleynci sey* 　　　　　*kay* orange three C 'three oranges'
(2) *oleynci sey* 　　　　　*sangca* orange　three　M.box　　　　　　　'three boxes of oranges'

While both Cs and Ms denote a quantificational unit, Cs also highlight certain inherent semantic features of the noun and constitute a closed word class. However, previous studies of Korean lack precise criteria to identify Cs. The methods we use are based on the view of Greenberg (1990[1972]) and Her (2012) that C/M serve as multiplicands, and if a C/M's mathematical value is necessarily *1*, it is a C; otherwise, it is an M. Specifically, in the tripartite phrase [N Num X], the numeral and X form a multiplicative relationship, i.e., [Num × X], and if X equals 1, then X is a C; otherwise, X is an M. This multiplicative theory has been successfully applied to Taiwan Mandarin (Her & Lin 2015), Taiwan Southern Min (Chen et al. 2020), Hakka (Liao 2014), and Japanese (Wang 2015).

　　This study first critically reviews previous studies on Korean numeral classifiers. We then develop a set of precise syntactic and semantic criteria to distinguish Cs from Ms as shown below, and produce a comprehensive inventory of Cs in Korean.

(1) Cs can be omitted in principle. [N Num C] = [N Num]
a. *oleynci　sey　kay*　=　*oleynci　seys* orange three　C　　　　　　orange　three
　　'three oranges'　　　=　'three oranges'
b. *oleynci　sey　sangca*　≠　*oleynci　seys* orange　　　　　three　M
　　　　orange　three 'three boxes of oranges' ≠'three oranges'
(2) Cs can be paired with countable nouns only. [NCount/*NMass Num C]
a. *mantwu　sey　kay*
　　dumpling three　C 'three dumplings'
b. **milkalwu*　sey　kay* flour　　　three　C '*three flour'

(3) Cs cannot be modified by an adjective. [(Adj) N Num (*Adj) C]

a.   *khun   oleynci   sey   kay* big   orange   three   C 'three big oranges'

b.   **oleynci sey*   *khun*   *kay* orange   three   big   C Intended: 'three big oranges'

Putative Korean sortal classifiers were collected from the lists compiled in Unterbeck (1994), Kuo (1996), and Oh (1994). The result of our examination identified 54 sortal classifiers in Korean, where 8 also serve as mensural classifiers. The findings can serve as an important resource for further research on Korean classifiers and for teaching Korean as a second language.

Keywords: Korean, numeral classifier, sortal classifier, mensural classifier, multiplication

Selected References

Chen, M.Y., Hsieh, M.L., & Her, O.S. (2020). 台灣閩南語分類詞之檢驗 [Distinguishing classifiers in Taiwan Southern Min]. *Language and Linguistics*, *21*(3), 375-407.

Greenberg, J. H. (1990). *Numeral classifiers and substantival number: Problems in the genesis of a linguistic type*. Stanford University, CA: Committee on Linguistics. (Original work published 1972)

Her, O.S. (2012). Distinguishing classifiers and measure words: A mathematical perspective and implications. *Lingua*, *122*(14), 1668-1691.

Her, O.S., & Lin, K.H. (2015). 分類詞語量詞的區分——以台灣華語為例 [Distinguishing classifiers and measure words: A case study of Taiwan Mandarin]. *Chinese Linguistics*, *4*, 56-68.

Kuo, C.W. (1996). 한국어 분류사 연구 [Research on Korean Classifiers] (Doctoral dissertation). Sungkyunkwan University, Seoul, Korea.

Liao, J.Y. (2014). 區分客語中之分類詞[Identifying classifiers in Hakka] (Master's thesis). National Chengchi University, Taipei.

Oh, S.R. (1994). *Korean Numeral Classifiers: Semantics and Universals* (Doctoral dissertation). State University of New York, Buffalo.

Unterbeck, B. (1994). Korean classifiers. In Y.-K. Kim-Renaud (Ed.), *Theoretical issues in Korean linguistics* (pp. 367-385). Stanford, CA: CSLI.

Wang, W. (2015). *Identifying classifiers in Japanese* (Master's thesis). National Chengchi University, Taipei.

Teaching Mandarin Tones:
How to Make it Comprehensible and Compelling

Anna S.C. Cheung

(The Hong Kong Polytechnic University, Hong Kong, China)

Tones, as exhibited in tone languages (e.g. Mandarin and Thai), is a challenging concept for most intontational language speakers (e.g. English and Korean) as their native languages do not have the rise and fall in pitches that would lead to a lexical and/or grammatical distinction in meaning. Even for tone language speakers, learning another tone language system is neither an easy task. This gives rise to the question of how tones are taught in second language (L2) settings. Over the years, a variety of strategies have been developed to help students distinguish the tones in the target language, such as using colours, gestures, graphs or symbols. It seems, however, that not a single strategy has found to be remarkably effective and in most cases learners are left with 'imperfect' tones.

Krashen (2003) advocates the importance of having compelling comprehensible input in learning an L2. The central idea of compelling comprehensible input is that the L2 teaching materials should only be set at a level a little higher than the learner's current level and should be interesting. Although his idea mainly focuses on acquiring vocabulary through extensive reading, the use of compelling comprehensible input can well be adapted into other realms of teaching, and in this context, tone articulation, an area that is often overlooked by tone language teachers.

Tones are often viewed as an analogy to musical notes as they resemble each other to a very large extent. This has already been noted in Chao's works on Cantonese (1947) and Mandarin (1948). However, most teachers and textbook authors (of these two languages/dialects) have often neglected this part even if they adopt Chao's five-point pitch scale (1930). Although there is still controversy regarding whether musical training would be beneficial to the learning of lexical tones (e.g. Alexander et al., 2005; Tao, et al., 2021), connecting solfège with the pitch values derived from the five-point pitch scale (i.e. Chao tone numerals) could help learners understand more about tones in an interesting way.

With reference to Chao's proposal (1930, 1947, 1948), below is an example of a more comprehensible and interesting illustration of the four basic tones in Mandarin.

Tone number	Tone 1	Tone 2	Tone 3	Tone 4
Chao tone numeral (pitch value)	55	35	214	51
Solfège	sol-sol	mi-sol	re-do-fa	sol-do
Example	ma1 *Mother*	ma2 *hemp*	ma3 *horse*	ma4 *scold*
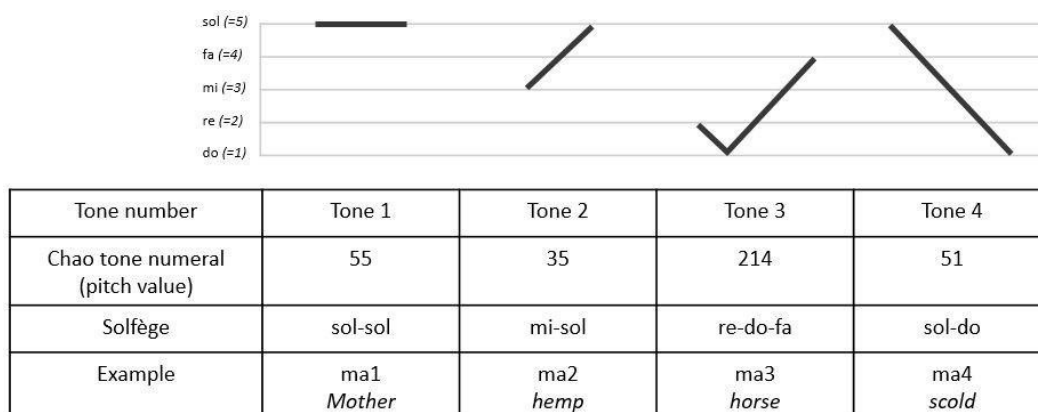
Figure 1. A comprehensible and interesting visual example showing the four basic tones in Mandarin, with reference to Chao's tonal notation (1930, 1947, 1948)

The first step to understand tones in Mandarin is to know the pitch value and contour of each of the four

tones, Figure 1 shows the 'appearance' of the four tones along Chao's tonal notation, mapping the five pitch values with five musical notes (from *do* to *sol*). This is similar to singing: reading the score is the first step to understand the melody of a song. Tones are introduced in a form similar to a prior experience (for most people) of singing in solfège, and such input could be counted as a level little higher than the learners' current knowledge. The analogy of tone and music itself is also interesting.

To bring learners again to a level little higher, the next step is to demonstrate (by teachers/native speakers) the contours of the four tones, either in real time or in videos. With technological advancement, mobile apps can serve this purpose, both for teachers/native speakers to demonstrate and for learners to have self-check. This is similar to checking the pitch in singing and this activity should be more interesting than previous proposals such as using Praat or other computer software to view pitches and tone contours. However, learners should be aware of the differences, though minor, in the shape of the contour and/or the pitch value in actual speech as Chao's notation is an ideal and simplified representation of a certain tone.

The above strategy for teaching and learning the four tones in Mandarin can well be extended to teaching other tone languages such as Cantonese or Thai.

References

Alexander, Jennifer A., Wong, Patrick C. M., and Bradlow, Ann R. 2005. "Lexical Tone Perception in Musicians and Non-musicians," Paper presented at the 9th European Conference on Speech Communication and Technology, Lisbon, 397-400.

Chao, Yuen Ren. 1930. "ə sistim əv "toun-letəz" [A System of "Tone-Letters"]," *Le Maître Phonétique*, 30, 24-27.

Chao, Yuen Ren. 1947. *Cantonese Primer*. Harvard University Press.

Chao, Yuen Ren. 1948. *Mandarin Primer: An Intensive Course in Spoken Chinese*. Harvard University Press.

Krashen, Stephen D. 2003. *Explorations in Language Acquisition and Use: The Taipei Lectures*. Heinemann.

Tao, Ran, Zhang, Kaile, and Peng, Gang. 2021. "Music Does Not Facilitate Lexical Tone Normalization: A Speech-Specific Perceptual Process," *Frontiers in Psychology*, 12, 1-14. DOI=10.3389/fpsyg.2021.717110

Revisiting Extrametricality in Isbukun Bunun

Zi-Xiang Lin

(Kang Chiao International School, Taiwan)

Unlike in *Optimality Theory* (OT; Prince & Smolensky, 1993/2004) with traditional constraints that evaluate candidates by the total number of violations, quantity-insensitive footing in directional evaluation (DE) is evaluated by directional constraints that harmonically order candidates by the location of violations (Lamont, 2022). As consequence, PARSE(σ) (Prince & Smolensky, 1993/2004) in DE eliminates the need for alignment constraints (McCarthy & Prince, 1993), because directional PARSE(σ) not only motivates iterative footing, but also, like alignment constraints, determines where feet surface. That is, for example, $(\sigma_1)\sigma_2\sigma_3\sigma_4$ (violation vector: $0_1 1_2 1_3 1_4$) is more optimal than $\sigma_1(\sigma_2\sigma_3)\sigma_4$ (violation vector: $1_1 0_2 0_3 1_4$) for PARSE(σ)$^\Rightarrow$, and $\sigma_1\sigma_2\sigma_3(\sigma_4)$ (violation vector: $0_4 1_3 1_2 1_1$) more optimal than $\sigma_1(\sigma_2\sigma_3)\sigma_4$ (violation vector: $1_4 0_3 0_2 1_1$) for PARSE(σ)$^\Leftarrow$. Moreover, TROCHEE (1) and IAMB (2) in DE are defined as such that both constraints penalize monosyllabic feet, wherein the stressed syllable is simultaneously the foot's leftmost and rightmost daughter. This renders FTBIN (Prince & Smolensky, 1993/2004) redundant in DE and as a consequence, OT with DE is argued to be more parsimonious than OT with traditional constraints. Following Lamont (2022), Lin (2024) uses Isbukun Bunun (Formosan), a right-aligned, quantity-sensitive trochee language wherein the single stress falls on the penults by default (Huang, 2008), for demonstration and shows that, with revised TROCHEE (3) and IAMB (4), the streamlined DE analysis can as well be extended over to quantity-sensitive footing.

(1) TROCHEE (Lamont, 2022): Assign one violation for every foot whose rightmost child is its head.

(2) IAMB (Lamont, 2022): Assign one violation for every foot whose leftmost child is its head.

(3) TROCHEE (revised; Lin, 2024): Assign one violation for every foot whose rightmost child is its head, or whose leftmost child is not its head.

(4) IAMB (revised; Lin, 2024): Assign one violation for every foot whose rightmost child is its head, or whose leftmost child is not its head.

However, footing in Isbukun Bunun adheres to morphological/positional extrametricality (Ouyang, 2012). In morphological extrametricality (6), inflectional suffixes never bears stress, which instead is shifted leftward onto the stem. Positional extrametricality (5), on the other hand, forbids word-initial stress in the case of a derivational suffix attaching to a monosyllabic, bimoraic root. A question arises as to whether Lin's directional evaluation can address extrametricality to increase its empirical adequacy while still keeping FTBIN out of the game. The answer is positive if NONINITIALITY$^\Rightarrow$ (Buckley, 1994), which forbids a prosodic word whose leftmost syllable is dominated by a foot, outranks TROCHEE$^\Leftarrow$. As (5) demonstrates, in the case of a derivational suffix attaching to a monosyllabic, bimoraic root, the dominance of NONINITIALITY$^\Rightarrow$ over TROCHEE$^\Leftarrow$ forces the second mora of the root to be footed with the following mora in the suffix. [1] Note that (5aii) and (5av) violate, respectively, MAX(μ)$^\Rightarrow$ and WEIGHT-TO-STRESS $^\Rightarrow$, both of which are dominated by TROCHEE $^\Leftarrow$. Regarding stress shift induced by morphological extrametricality, lexically indexing (Pater, 2010) are used to distinguish IAMB$_{INF}$$^\Rightarrow$ from the general IAMB$^\Rightarrow$. As (6) demonstrates, undominated IAMB$_{INF}$$^\Rightarrow$ and BAL-TROCH$^\Rightarrow$ drive the foot leftward onto the stem (6a) by penalizing, respectively, inflectional morphemes as part of the trochee (6b) and unbalance trochaic footing (6c). The proposed analysis not only increases the DE's empirical adequacy by adding into the picture morphophonological interactions in weight-sensitive footing in Isbukun Bunun, but also maintain DE's simplicity by keeping FTBIN and ALIGN out of the game.

---

[1] Coda consonants are nonmoraic in the language (Huang, 2008).

(5) *Positional extrametricality*: NONINITIALITY$^⇒$ >> TROCHEE$^∈$

    a. /mu:-an/ => mu.(ˈuan)   'your place (loc.)'

/mu:-an/ 'your place'	BAL-TROCH$^⇒$	NONINITIALITY$^⇒$	TROCHEE$^∈$
☞ i. mu_1(ˈuan_2)			$1_2 0_1$
ii. (ˈ$muan_1$)		$W1_1$	$L1_1$
iii. (ˈ$mu{:}_1 an_2$)	$W0_1 1_2$	$W0_1 1_2$	L
iv. (ˈ$mu{:}_1$)an_2		$W0_1 1_2$	$W0_2 1_1$
v. $mu{:}_1$(ˈan_2)		$W0_1 1_2$	$1_2 0_1$

    b. /suis + -un/ => su.(ˈi.sun)

/suis-un/ 'graved into threads (pat.)'	BAL-TROCH$^⇒$	NONINITIALITY$^⇒$	TROCHEE$^∈$
☞ i. su_1(ˈ$i_2 sun_3$)			
ii. (ˈ$sui_1 sun_2$)	$W0_1 1_2$	$W0_1 1_2$	
iii. (ˈsui_1)sun_2		$W0_1 1_2$	$W1_1 0_2$

(6) *Stress shift in morphological extrametricality*: IAMB$_{INF}$$^⇒$ >> TROCHEE$^∈$

/antalaun-in$_{INF}$-su$_{INF}$/ 'to have been accepted by you'	IAMB$_{INF}$$^⇒$	BAL-TROCH$^⇒$	NONINITIALITY$^⇒$	TROCHEE$^∈$
☞ a. $an_1 ta_2$(ˈlau_3)$nin_4 su_5$				$0_5 0_4 1_3 0_2 0_1$
b. $an_1 ta_2 lau_3$(ˈ$nin_4 su_5$)	$W0_1 0_2 0_3 1_4 1_5$			L
c. an_1(ˈ$ta_2 lau_3$)$nin_4 su_5$		$W0_1 0_2 1_3 0_4 0_5$		L
d. (ˈ$an_1 ta_2$)$lau_3 nin_4 su_5$			$W0_1 0_2 0_3 0_4 1_5$	L

Selected Reference:

Lamont, Andrew. 2022. "A Restrictive, Parsimonious Theory of Footing in Directional Harmonic Serialism," *Phonology* 39(1), 41-78.

The Inspiration of Seq2Seq and Its Error Analysis—
The Sound Changes of Chaoshan and Zhangzhou

Man-ni Chu
(Fu Jen Catholic University, Taiwan)
Yu-Chun Wang
(Dharma Drum Institute of Liberal Arts, Taiwan)

Past researches (Chu, 2009; Hsu, 2016) have demonstrated the closed relationship between Zhangzhou (ZZ) and Chaoshan(CS), while their different evolutionary pattern was also reported. CS has successfully changed its alveolar codas into velar codas., while such a change is not observed in ZZ. On the other hand, the attention-based seq2seq method was widely used in translating one language to another by calculating the probabilities of the sequence alphabets based on its learning algorithm. Thus, this study utilizes the attention- based seq2seq method to predict the changes of ZZ and CS over the past century. Based on the error data, and the parallel fieldwork done by linguists, a possible ongoing sound change of ZZ is predicted.

Table 1. demonstrates the scenario of the whole process. We chose one dictionary from each dialect as the input, while the other dictionary from the same dialect published 100 years later was the output. The same words in two dictionaries for the same dialect were cleaned manually based on its colloquial or literacy sounds, yielding 2233 words in ZZ, and 1995 words in CS. The characters from dictionaries in 19th for both dialects were used as the input token, and the tones were changed into numbers. The attention-based seq2seq predicts the same words orthography in the same dialect after 100 years. If the predicted orthography was identical as the ones reported in 20th for both dictionaries, it will be the correct answer for the prediction, yielding a 48% for ZZ and 68% for CS.

For ZZ, among the 1,150 words with errors, it was found that 296 were incorrectly predicted initials, 780 were rhymes, and only 302 were tones. For CS, among the 701 words with errors, it was found that 119 were incorrectly predicted initials, 306 were rhymes, and only 411 were tones. Then, we added Guangyun's information of classification on all words, and again clean the data to avoided any multiple correspondence. The information provided by Guangyun are the proto-vowel, in She (攝), the rhymes, in yun(韻), with glides or not, from the first to the fourth, in Dengdi (等第), the openness or closeness of mouth, in Kaihe(開合). Based on the middle Chinese information, we further classified the error data by the pivot function in excel,. Take the first three positions, around 35% of the whole errors. For ZZ, the errors occurred mostly in Shang, Gen and Dang she, where the rhymes are non-high vowels + alveolar/velar (nasal) stops. For CS, in Shang, Zhi, Xie and Yu she, where more than half rhymes are glide + diphthongs.

If the two dictionaries from the same dialect reported a regular sound change over the past century, the attention-based seq2seq method can learn it better. See the accuracy of CS. If the model cannot correctly predict the sound changes, it is possible that the dialect is experiencing an ongoing sound change. Our results demonstrate that ZZ has more errors predicted by the model than CS. More specific, the incorrectly predicted rhymes of ZZ mainly categorized in the non-high vowel+ alveolar/velar (nasal) stops, implying an ongoing sound change of ZZ. This phenomenon was observed in some fieldworks (Zheng, 2019; Lin,2018; Chen, 1999). Taken together, applying the attention-based seq2seq method being parallel to the ongoing sound change pattern observed by linguists seems feasible.

Table 1: The flow of applying attention-based seq2seq model

	input	attention-based seq2seq	output
ZZ (2233 words)	Medus' Fujian Dialect Dictionary (1837, reprinted in 1993)		Xia Zhang Quan Dictionary edited by Zhou Changji (2006)
CS (1995 words)	Goddard's Teochew-English Dictionary (1883)		Chaoshan Dictionary edited by Zhang Xiaoshan (2009, reprinted in 2015)

Keywords: attention-based seq2seq, Zhangzhou (ZZ), Chaoshan (CS). Sound change

References

Chen, Bijia(陳碧加). (1990). The evolution and application of nasal rhymes in Zhangzhou dialect. Journal of Zhangzhou Normal University, (1), 96-101

Chen, Ronghan(陳榮翰). (1999). Changes in Zhangzhou dialect in modern times. Journal of Zhangzhou Vocational and Technical College, (4), 80-87

Chu, Manni (朱曼妮). (2009). Exploring the changes in the final coda of Chaoshan from the perspective of language perception. Hsinchu: National Tsinghua University PhD thesis

Hsu, Yuhang(徐宇航). (2016). Research on the evolution of finals in Chaozhou dialect over the past hundred years. Linguistics Series, (1), 241-260

Zheng, Chaoqun(鄭超群). (2019). Research on the evolution of the rhyme ending of Yang tone in Wu and Min dialects in modern times. Fujian: Doctoral thesis of Fujian Normal University.

A Sociolinguistic Areal Analysis ofa Three-Village Micro-Linguistic Area (MLA) at the Haryana-Punjab Border in India

Ajay Malik

(Central University of Rajasthan, India)

Usha Udaar

(IIT Kanpur, India)

This study investigated the language dynamics and contact phenomena within a micro- linguistic area made up of three villages on the Haryana-Punjab border, namely Ujhana, Garhi, and Data Singh Wala. The study explains the correlation of patterns of language contact, convergence, and areal similarities in this specific three-village micro-linguistic area (henceforth, TV-MLA). The study utilizes both sociolinguistic and areal analyses.

The study of the 'linguistic area 'in India, initiated by Emeneau (1956) and advanced by Masica (1976), lays the foundation for conceptualizing India as a linguistic area. Abbi (1991, 2012) revisits this concept, exploring language contact. Khubchandani (1991) proposes the concept of "kshetra" as a collective reality to reflect India's linguistic diversity. Pattanayak (1990) addresses inequalities perpetuated by standard languages, while Annamalai (2006) contextualizes diversity within sociopolitical constructs, emphasizing functional multilingualism shaping its diverse linguistic areas. Dixon (2006, 2008) brought attention to the "small linguistic area," emphasizing its unique linguistic dynamics. Brito (2016) studies globalization's impact within an MLA. Romano (1999), Matheson (1996), and Korn (2022) examine MLAs beyond India. Kachru (2008) provides a comprehensive overview of the South Asian linguistic landscape. Mishra (2024) highlights the importance of linguistic diversity's role in fostering unity within 'India as one linguistic area'.

A mixed-method approach combines quantitative sociolinguistic surveys, qualitative ethnographic interviews, language attitude surveys, and linguistic analyses to study language variations.

This study addresses the research gap by conducting a comprehensive linguistic-area study on the TV-MLA, an aspect often overlooked in the existing literature that predominantly focuses on broader regional, national, or transnational contexts.

The primary argument is that the research conducted at the micro-linguistic level refines the basis of broader studies, such as 'India as One Linguistic Area. 'The secondary arguments address the role of language policy in shaping identity, the influence of socioeconomic factors on language shift, the effects of migration on language, intergenerational transmission and shared areal features. Findings show a multi-faceted correlation between sociopolitical factors, language usage, and identity; characterizing a dynamic linguistic landscape marked by shared features, language maintenance, language shift, and multilingualism.

This study contributes to the academic understanding of language dynamics and contact within MLAs; highlighting their implications for sociolinguistic theory and practice. It also informs language policy and planning (LPP), cultural exchange, identity formation, and practical on- ground community development initiatives; and supplements the broader regional- or national- level linguistic area studies.

Selected References

Abbi, A. (1991). *India as a Linguistic Area Revisited*. Pergamon Press.

Abbi, A. (2012). Languages of India and India as a Linguistic Area. Unpublished Manuscript. Centre of Linguistics, Jawaharlal Nehru University. https://www.academia.edu/10178488/Languages_of_India_and_India_as_a_Linguistic_Area.

Annamalai, E. (2006). India: Language Situation. In K. Brown (Ed.), Encyclopedia of Language &

Linguistics (2nd ed., pp. 610-613). Elsevier. ISBN 9780080448541. https:// doi.org/10.1016/B0-08-044854-2/04611-3.

Brito, A. (2016). Multilingualism and Mobility: A Linguistic Landscape Analysis of Three Neighbourhoods in Malmö, Sweden. MA Thesis, Lund University.

Dixon, R. M. W. (2006). Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics. Oxford University Press.

Dixon, R. M. W. (2008). Language Contact in the Cairns Rainforest Region. Anthropological Linguistics, 50(3/4), 223–248. http://www.jstor.org/stable/20639004

Emeneau, Murray B. (1956). India as a Linguistic Area. Language, 32(1), 3-16. https:// doi.org/10.2307/410649

Gumperz, J. J. (1982). Discourse Strategies. Cambridge University Press.

Kachru, B., Sridhar, S. N., & Kachru, Y. (Eds.). (2008). Language in South Asia. Cambridge University Press.

Khubchandani, L. M. (1991). India as a Sociolinguistic Area. Language Sciences, 13(2), 265-288. https://doi.org/10.1016/0388-0001(91)90018-V

Korn, A. (2022). Areal Features in the Languages of South Iran: Focus on Balochi and Baskhardi. Iranian Languages and Linguistics Lecture Series. Presented at Université de Toronto, Canada. Retrieved from https://shs.hal.science/halshs-03622909/

Labov, W. (1966). The Social Stratification of English in New York City. Center for Applied Linguistics.

LaPolla, R. J. (2001). The Role of Migration and Language Contact in the Development of the Sino-Tibetan Language Family. In Areal Diffusion and Genetic Inheritance: Case Studies in Language Change, ed. by RMW Dixon & AY Aikhenvald (pp. 225-254). Oxford University Press.

Masica, Colin P. (2005). Defining a Linguistic Area: South Asia. Chronicle Books.

Matheson, D. J. (1996). An Examination of Some of the Difficulties of Establishing an Open University in a Small Linguistic Area: The Case of Suisse Romande. International Journal of Lifelong Education, 15(2), 114–124. https://doi.org/10.1080/0260137960150206

Mishra, A. K. (Speaker). (2024). India As A Linguistic Area And Diversity In The Unity Of Indian Languages [Symposium Talk]. Bharatiya Bhasha Family, INDICA's Center for Bhasha Studies. https://www.youtube.com/watch?v=r_1q8MlYmqc

Pattanayak, D. P. (Ed.). (1990). Multilingualism in India (Multilingual Matters Series No. 61). Multilingual Matters. ISBN-13: 9781853590733

Rao, G. U. M. (2018). The Linguistic Archaeology of India: A Contribution to the Asian Languages. Moḻivādam: Collected Papers in Applied Linguistics, 1, 7.

Romano, A. (1999). A Phonetic Study of a Sallentinian Variety (Southern Italy). In Proceedings of the 14th International Congress of Phonetic Sciences (pp. 1051-1054). https:// www.lfsag.unito.it/antonio_romano/A_Phonetic_Study_of_a_Sallentinian_Varie.pdf

Satyanath, S. (2022). Standard-Ness, National Ideologies and their Embedding in 'Sociolinguistic Theory'. Asia-Pacific Language Variation, 8(2), 174-205.

Satyanath, S. (2015). Language Variation and Change: The Indian Experience. Globalising Sociolinguistics (pp. 107-122). Routledge.

Schmidt, A. (1985). Young People's Dyirbal: An Example of Language Death from Australia. Cambridge University Press.

Thomason, S. G., & Kaufman, T. (1988). Language Contact, Creolization, and Genetic Linguistics. University of California Press.

Weinreich, U. (1953). Languages in Contact: Findings and Problems. Linguistics, 40(6), 1-9.

"I" or "we" in a Pandemic? A Linguistic Case Study on the Personalization of Hungarian Political Communication during the Coronavirus[1]

Balázs Horváth

(Corvinus University of Budapest, Hungary)

As politicians increasingly share personal content and information about themselves, the importance of political collectives – such as political parties – is diminished (Aczél 2009, Holtz-Bacha et al. 2014, McAllister 2009, Plasser 2008, Rahat & Sheafer 2007, Szabó 2021, 2022). This phenomenon can also be observed in the Hungarian political discourse, including social media (Farkas & Bene 2022) and television (Horváth 2023). During the Covid-19 outbreak, the stakes of Hungarian political communication were even higher. This was due to the necessity of addressing the question of the existence of the rule of law (Szente 2020), the government's incitement of public opinion, for example through the creation of enemies (Schlett 2018), the identification of the virus with war metaphors (Szabó 2020), or the one-sided and limited government communication.

What were the tendencies in the personalization of political communication during the coronavirus pandemic? My research sought to answer this question through an analysis of transcripts of interviews (N=39) given by the Hungarian Prime Minister to Kossuth Radio between 2020 and 2021, during the first three "waves" of the coronavirus. Using corpus linguistics methods, I examined the use of first-person singular (1PS) and first-person plural (1PP) pronouns and other possessive markers. This method allowed me to quantify the degree of personalization versus collective representation in the Prime Minister's rhetoric. The database was examined using AntConc corpus linguistic analysis software (Anthony 2023).

In the case of 1PS, an upward-downward trend is apparent, while for 1PP an opposite curve is observed during the first three "waves" of the coronavirus. The use of first-person plural pronouns outnumbered the first-person singular pronouns, with the former representing 63.2% and the latter 36.8%. This indicates a predominant inclination towards collective representation. There was only one radio interview in which 1PS was used more frequently than 1PP. This can be explained by the Prime Minister's attempt to deflect responsibility from 1PS (himself) to 1PP (in this case the government).

With regard to first-person plural possessive markers, the statement can be made on behalf of several groups, which in my paper can refer to family, party, nation, nations, European Union and humanity. Szabó (2021) used a radial model to represent the relative distance of first-person plural denotations from the speaker, and I further developed this to create Figure 1, which serves as the methodological basis for the research.
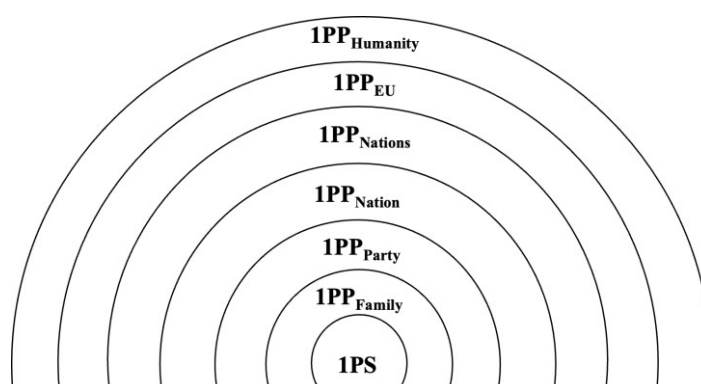

Figure 1: The relative distance of the referents of first-person plural pronouns to the deictic center (the authors' own work based on Szabó 2021, p. 65).[2]

The assumption of this model is that we see the world in an egocentric way, which also implies that the first-person singular (I) is located at the center, the first-person plural (we) is located further away (Szabó, 2021).[3] Among the 1PP possessive markers 1PP$_{Family}$ is the closest to the speaker, since this group includes the individuals with whom the politician has the closest private relationship. This group is followed by 1PP$_{Party}$, which is at least a community of people who are similar in their party identification. Next in line is 1PP$_{Nation}$, which includes people of the same nationality as the Prime Minister, i.e. Hungarians. Then come the supranational categories: 1PP$_{Nations}$ and 1PP$_{EU}$, and finally the broadest category 1PP$_{Humanity}$.

The frequency of references to family, party, and humanity subcategories decreased, while those to nation, nations and the European Union increased.

From the very beginning, families played a central role in their communication. It is interesting to note that the proportion of people speaking on behalf of their own family in the corpus during the coronavirus was negligible. This discrepancy can be explained by the fact that during the Covid-19 outbreak, the PM focused on his party and the nation rather than his own family. It may be counterproductive too to speak of the latter at a time when community distancing rules were in place. Furthermore, it may serve a "hiding" function to deliberately say little on their behalf. The number of references to the European Union has increased, presumably because Hungary required the support of the 27 Member States to join forces, both financially and politically.

Selected References

Aczél, Petra. 2009. "Új retorika: közélet, kommunikáció, kampány." Budapest: Kalligram Könyvkiadó. Anthony, Laurence. 2023. "AntConc (Version 4.2.0) [Computer Software]." Tokyo: Waseda University. Farkas, Xénia, & Bene, Márton. 2022. "Orbán Viktor vizuális és verbális populista stílusa a Facebookon," *Politikatudományi Szemle* 31(3), 82–108, doi: 10.30718/POLTUD.HU.2022.3.82.

Holtz-Bacha, Christina, Langer, Ana Ines, & Merkle, Susanne. 2014. "The personalization of politics in comparative perspective: Campaign coverage in Germany and the United Kingdom." in *European Journal of Communication* 29(2), 153–170, doi: 10.1177/0267323113516727

Horváth, Balázs. 2023. "Retúrjegy a „miénktől" az „enyémig": A politikai perszonalizáció vizsgálata Orbán Viktor televíziós interjúiban a személyes névmások elemzésén keresztül" Scientific Students' Association Thesis. Budapest: Corvinus University of Budapest, Institute of Marketing and Communication Sciences.

McAllister, Iam. 2009. "The Personalization of Politics," in J. Dalton, Russell – Klingemann, Hans-Dieter, eds., *The Oxford Handbook of Political Behavior*, 571–888. Oxford: Oxford Academic.

Plasser, Fritz. 2008. "Political consulting worldwide," in *Routledge handbook of political management*, 40–57. New York: Routledge.

Rahat, Gideon, & Sheafer, Tamir. 2007. "The personalization(s) of politics: Israel, 1949–2003," in *Political communication* 24(1), 65–80, doi: 10.1080/10584600601128739.

Schlett, István. 2018. "A politikai gondolkodás története Magyarországon." Budapest: Századvég. Szabó, Lilla Petronella. 2020. "Háború járvány idején: Hogyan beszél Orbán Viktor a koronavírusról?" in Körösényi, András, Szabó, Andrea, & Böcskei, Balázs, eds., *Vírusba oltott politika: Világjárvány és politikatudomány*, 123–133. Budapest: Napvilág Kiadó.

Szabó, Lilla Petronella. 2021. "Karnyújtásnyira a választóktól? A perszonalizáció vizsgálata a személyes névmásokon keresztül az Egyesült Államok politikai kommunikációjában," in *Politikatudományi Szemle* 31(1), 54–75, doi: 10.30718/POLTUD.HU.2022.1.54.

Szabó, Lilla Petronella. 2022. "It's nothing personal? – A linguistic account of the personalization of American political communication" Doctoral dissertation. Budapest: Corvinus University of

Budapest, Doctoral School of Sociology and Communication Science, doi: 10.14267/phd.2022049.

Szente, Zoltán. 2020. "A 2020. március 11-én kihirdetett veszélyhelyzet alkotmányossági problémái" in *Állam-és Jogtudomány* 61(3), 115–139.

Jobst, Ágnes. 2007. "A mi mint a hatalom és a szolidaritás névmása." in *Magyar Nyelvőr* 131(1), 29–47.

---

[1] Supported by the ÚNKP-23-1-I-CORVINUS-47. New National Excellence Program of the Ministry for Culture and Innovation from the source of the National Research, Development and Innovation Fund.

[2] The different categories of 1PP are indicated by a subindex based on Szabó (2021, 2022).

[3] The person markers are "located" at different distances from the speaker, and there is also a kind of ordering between them according to Rees' (1983 cited in Jobst 2007) model – in the current research I interpret proximity to the speaker in this way.

Language as Symbol in Urban Linguistic Landscapes: Comparative Analysis of China, Thailand, and Russia

Oleg Shcherbakov

(Lomonosov Moscow State University, Moscow, Russia)

The study of linguistic landscapes is perhaps the most widely recognized and utilized approach for analyzing visual multilingualism. It seeks to understand the public display of specific languages in urban environments as a reflection of and contributor to their ethnolinguistic vitality. This vitality encompasses the relative prominence of languages within their speech communities and their use beyond these contexts (Kelly-Holmes 2014). The linguistic landscape has been transformed into a social landscape, where its features can be interpreted through an examination of publicly displayed signage. This analysis reveals the ways in which languages are visually represented and how they contribute to the social fabric of these shared spaces (Maly 2019).

One of the most recent challenges for linguistic landscape researchers is using languages for symbolic (fetishised) rather than utility (instrumental-communicative) purposes (Kelly-Holmes 2014). This trend manifests in the labelling of certain languages, such as English and French, as appealing to international audiences and social media users. The perceived cosmopolitanism of these languages and their association with the influence and wealth of international trade further contribute to their symbolic appeal (Hult & Kelly-Holmes 2019).

Our previous research focused on the symbolic usage of language in linguistic landscapes of Thailand and China (Shcherbakov & Bagirova 2020, Shcherbakov 2021). We found that that the majority of signage studied use of English and other foreign languages instead of local languages as a symbolic tool for creating proper upscale ambiance for social media users and catering to tourists and expatriates. We hypothesized that with the majority of population in these countries either being not proficient in English or not speaking it at all, the tendency to use foreign languages presupposes limited access to the language of the city and, thus, contributes to inequality. Additionally, we concluded that the linguistic landscape of a place is now as important as its design and is paid much attention to by the founders.

In this research, we compare previously made results with the new dataset from Moscow, the capital of Russia. This city is also experiencing an increasing tendency to use language solely for symbolic purposes appealing to their target audiences. Of special interest in such cases are active social media users who use linguistic elements as objects for their photographs. Social media and the rise of Instagrammability have influenced consumer decision-making practices in various fields. It is now at the forefront of architectural studios' concerns when working on new projects as the clientele is requesting that they design with the Instagram feed in mind (Fiocco & Pistone 2019). We hypothesize that symbolic usage of language is a new marketing tool widely preferred by those targeting at younger audiences.

We analyzed 70 linguistic landscape elements in Moscow, which, in our opinion, serve a symbolic function. We concluded that foreign languages in Moscow are used less often than in the previously examined cities — these are mostly Asian languages in the thematic Asian eateries that gain momentum in Moscow. For example, a newly opened Korean-style coffeeshop named "Haru" designed a sign in Korean in the form of an online chat with a peer:
— What are you doing when you're done?
— You want a cup of coffee?
— Where?
— At the Haru café.
Similar cases were found at Chinese- and Thai-styled cafes and shops. We also registered the

pervasive use of the national Russian language, which may testify to the changing of its status within the society and growing concern of the Russian population regarding the extensive use of borrowed words and phrases. We argue that more in-depth research is needed to identify the reasons behind the ongoing transformations in the linguistic repertoire of Moscow.

Selected References

Fiocco, Fabiola, Pistone, G. 2019. "Good content vs good architecture: where does 'Instagrammability' take us?" *Strelka Mag*. Online first, https://strelkamag.com/en/article/good-content-vs-good- architecture.

Hult, Francis M., Kelly-Holmes, Helen. 2019. "Spectacular language and creative marketing in a Singapore tailor shop," *International Journal of Multilingualism* 16(1), 79-93.

Kelly-Holmes, Helen. 2014. "Linguistic fetish: The sociolinguistics of visual multilingualism," *Visual Communication*, 135-151.

Maly, Ico. 2019. "Hipsterification and capitalism: A digital ethnographic linguistic landscape analysis of Ghent," *Tilburg Papers in Culture Studies*. Online first, https://research.tilburguniversity.edu/en/publications/hipsterification-and-capitalism-a-digital- ethnographic-linguistic.

Shcherbakov, Oleg, Bagirova, Anna. 2020. "Linguistic Landscapes of Gentrified Localities: A Case Study of China," *2020 the 5th International Conference on Education Science and Development*, 98-103.

Shcherbakov, Oleg. 2021. "Hipsterification of English in Linguistic Landscapes of Bangkok's Cafe-Hopping Spots," *Proceedings of academic and methodological studies*, 240-255.

Themed Sessions
(Day 1)

ChatGPT-generated Modifications on Human-Generated TESOL Abstracts by Vietnamese Researchers

Sheng-Chieh Huang, Hoa Ming Truong, Issra Pramoolsook
(Suranaree University of Technology, Thailand)

Research on writing research article abstracts has highlighted both its importance and challenges, particularly for non-native writers of English, including Vietnamese academics. The rapid advancement of AI and natural language processing has led to the development of sophisticated and versatile language models, including generative AI (Ray, 2023). OpenAI's ChatGPT model, a powerful tool, has gained wide applications in various domains (Alharbi, 2023; Thorp, 2023). As a potential writing assistant, it has been explored for its potential in enhancing academic writing skills, especially through its ability to edit or polish human-generated versions. Its editing effort certainly brings about changes or modifications on both structural and linguistic levels of the original human-generated texts, necessitating further investigation on its decisions regarding the types and the amount of the modifications made. This study, therefore, aimed to investigate the differences between human-generated TESOL research article abstracts written by Vietnamese authors and their ChatGPT-edited versions in terms of move structures and linguistic features through understanding the textual modifications generated by ChatGPT. The findings from textual investigation involving the analyses of abstract types following Lorés (2004), rhetorical moves based on Hyland's (2000) framework, and key linguistic features of 50 original abstracts and 50 corresponding ChatGPT-edited ones indicated significant modifications in structural, lexical, sentential and tense dimensions of the two sets, while maintaining the type and the moves of the original versions. Examples include removal of ambiguity at the lexical level, simplification of sentence structure and clarification of complex ideas at the sentential level, and even modifications at the structural level that render changes in the content through making problem statements clearer, making methodology more detailed, and expanding results and discussion. Besides, we conducted interviews using four open-ended questions with two experienced Vietnamese abstract and manuscript reviewers in order to elicit their opinions on the quality of the modifications identified. Both our textual findings and interview insights complemented each other that ChatGPT-generated modifications improved clarity, readability and informativeness of the original abstracts. Concrete and authentic examples of those modifications extracted from our abstract data will be shared in this presentation. This paper ends with proposed pedagogical implications drawn from our findings to help not only Vietnamese researchers in TESOL but also those outside Vietnam and in other related fields understand and appreciate such modifications, which can support them to produce research article abstracts more effectively.

Selected References

Alharbi, W. (2023). AI in the foreign language classroom: A pedagogical overview of automated writing assistance tools. *Education Research International, 2023*, 4253331. https://doi.org/10.1155/2023/4253331

Hyland, K. (2000). *Disciplinary discourses: Social interactions in academic writing*. Longman. Lorés,

R. (2004). On RA abstracts: From rhetorical structure to thematic organization. *English for Specific Purposes, 23*(3), 280–302. https://doi.org/10.1016/j.esp.2003.06.001

Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems, 3*, 121–154. https://doi.org/10.1016/j.iotcps.2023.04.003

Thorp, H. H. (2023). ChatGPT is fun, but not an author. *Science, 379*(6630), 313–313. https://doi.org/10.1126/science.adg7879

Resolving RC Attachment Ambiguity in LLMs: Contrasting Head-Initial and Head-Final Language Processing

So Young Lee

(Miami University, USA)

Russell Scheinberg, Amber Shore, Ameeta Agrawal

(Portland State University, USA)

In computational linguistics, research on syntactic attachment ambiguity has received significant attention and been actively studied since its early days, but most of this work has focused on prepositional phrase attachment ambiguity (1) (Karanikolas et. al. 2023, among others).

(1)   Amber saw the boy with binoculars.

While relative clause (RC) attachment ambiguity (2) as one type of syntactic attachment ambiguity has also been consistently and continuously explored in psycholinguistics as a phenomenon to uncover the hidden mechanisms of human sentence processing, there has been little attention devoted to how large language models handle RC attachment ambiguity.

(2) Russell saw the <u>son</u> of the <u>doctor</u> who danced at the party.
    a. The doctor danced at the party. (LA: Low Attachment)
    b. The son danced at the party. (HA: High Attachment)

This study examines the syntactic processing of five Large Language Models (LLMs), GPT-3.5, GPT-4o, Gemini 1.5, Llama 3, and Claude Opus, compared to human processing, focusing on relative clause (RC) attachment ambiguity. We investigate two research questions: i) Do LLMs reflect language- specific preferences in RC attachment? ii) Is LLMs' processing similar to human processing, in that it is sensitive to linguistic factors such as the length of relative clauses and the positions of constituents? Our experiments consist of two language groups: head-initial languages (English, German, French, Spanish from Hemforth et.al. (2015)) vs. head-final languages (Japanese, Korean). To compare LLMs' processing results directly to human data, we replicated the experiments from Hemforth et.al. (2015) for the first group of languages. Despite the lack of human sentence processing results for Korean and Japanese, we included these languages to evaluate LLMs' performance on structurally distinct head- final languages, using translations of the English sentences. Two factors (length of RCs: short vs. long, position: subject vs. object) were crossed in each language stimuli. (3) shows an example set of the English stimuli.

(3)

length	position	Sentences
Short	Subject	The son of the colonel who died wrote five books on tropical disease.
Long	Subject	The son of the colonel who tragically died of a stroke wrote five books on tropical disease.
Short	Object	The doctor met the son of the colonel who died.
Long	Object	The doctor met the son of the colonel who tragically died of a stroke.

LLMs were prompted to complete two tasks: i) RC identification task, ii) comprehension task.
RC identification results are summarized in Figure 1. We found great differences in LLMs' performances depending on language groups. In the head-initial language groups, all models show high and consistent performance (above 110 out of 128). In head-final language groups, except Claude 3 Opus, all models show quite poor performance. We speculate that these differences can be attributed to models' architectures or training datasets. These results highlight the models' generalization capabilities across different linguistic contexts and identify areas needing improvement, especially in non- European languages. Regarding attachment preferences, humans show LA preference in English, HA preference in French, Japanese, and Korean, and both LA and HA preference in German and Spanish (Cuesto et. al. 1998, among others). The results in Figure 2 indicate that models process languages differently and do not always align with human sentence processing outcomes. Particularly in Korean, all models show LA preferences, diverging from the human HA preference. We speculate that these results occur because the models primarily process in English, as evidenced by Japanese and Korean inputs generating most responses in English, suggesting an underlying translation process for these languages. As for the influence of the length of RC and the constituent positions, longer RCs tend to increase HA preference across languages, suggesting additional context enhances HA inclination. In English and Korean, the models mainly show LA preferences, with slight HA in object positions, while in Spanish, French, and German, the models generally exhibit stronger HA preferences in object positions, except for Gemini-1.5 Pro, which often deviates.
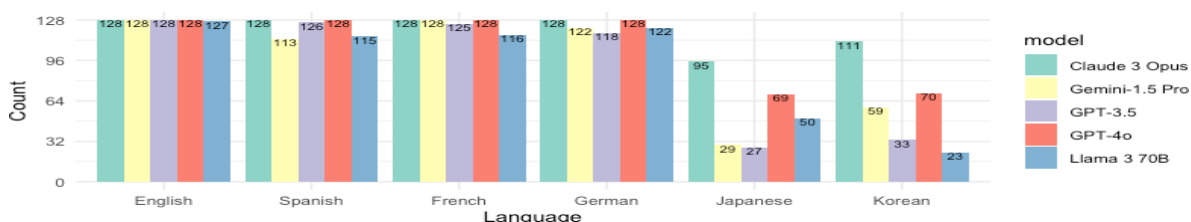
This study investigated the understudied issue of how LLMs process RC attachment ambiguity. Our research provides a deeper understanding of the characteristics of various models, illuminating the complexities of modeling human-like sentence processing.
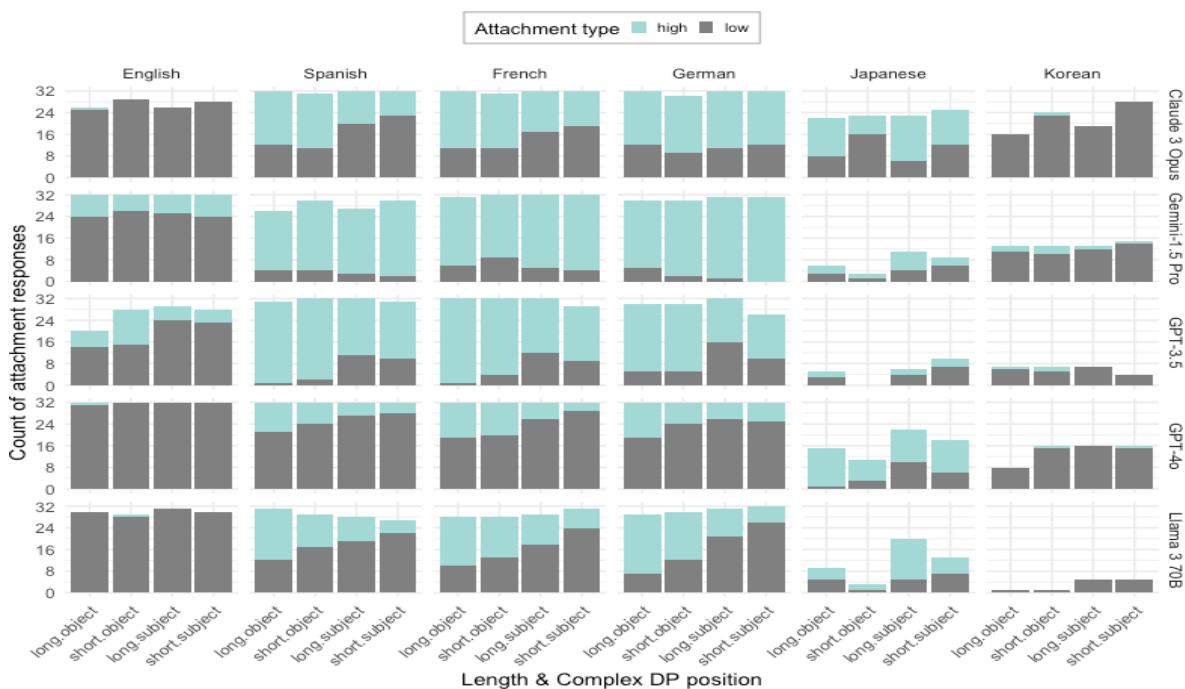
Selected References

Cuetos, Fernando, and Don C. Mitchell. 1988. "Cross-linguistic differences in parsing: Restrictions on the use of the Late Closure strategy in Spanish." *Cognition* 30.1, 73-105.

Hemforth, Barbara, et al. 2015. "Relative clause attachment in German, English, Spanish and French: Effects of position and length." *Lingua* 166 (2015): 43-64.

Karanikolas, Nikitas, Eirini Manga, Nikoletta Samaridi, Eleni Tousidou, and Michael Vassilakopoulos. 2023. "Large language models versus natural language understanding and generation." In *Proceedings of the 27th Pan-Hellenic Conference on Progress in Computing and Informatics*, 278-290.

An Ablation Study on the Automated Essay Scoring using ChatGPT

Ching-Yu Helen Yang
(National Chung Hsing University, Taiwan)
Jhih-Jie Chen
(National Tsing Hua University, Taiwan)

1. Introduction

Nowadays, English writing plays an integral part in language proficiency exams, such as TOEFL and IELTS. However, essay evaluation is a laborious undertaking that demands substantial time and resources. As a result, scholars have commenced investigating Automated Essay Scoring (AES), employing specialized computer programs to evaluate essays produced within educational contexts (Ke & Ng, 2019). One of the most popular methods at present involves leveraging ChatGPT for essay scoring (Mizumoto & Eguchi, 2023). Naismith et al. (2023) evaluate the feasibility of using ChatGPT for AES and explore the effectiveness of scoring with rationale in different forms. Previous works have demonstrated great potential in providing human-level assessment; however, the full extent of their capabilities and limitations remains a mystery.

In this study, we aim to maximize the effectiveness using different prompting techniques, which can guide the way of AES using ChatGPT in a zero-shot style. We focus on the following questions: Is GPT- 4 better than GPT-3.5? Can rating with rationales improve AES results? Is there any performance difference between prompts presented in different languages? We conduct an ablation study to examine the performance of these factors.

2. Methods

The **dataset** used in the study includes 200 English essays written by L2 learners with Chinese as L1, responding to composition test items. The test items were developed based on the guidelines (https://reurl.cc/9vqydO) from the College Entrance Exam Center (CEEC) in Taiwan. Three English teachers (A, B, C) are invited to provide the four aspect scores (0-5 points) and holistic scores (the sum of aspect scores, ranging from 0 to 20 points) of the responses according to the CEEC **rubric** (https://reurl.cc/Ke5vxM).

We use GPTs to predict holistic scores of the essays based on the test item prompts, rubric, and guidelines. Different GPT prompts are designed to explore how the following factors may impact GPTs' performance in the AES tasks.

- GPT versions: GPT-3.5 (**G3**; gpt-3.5-turbo-0613) vs. GPT-4 (**G4**; gpt-4-0125-preview)
- The language of prompts: English prompts (**EN**) vs. Chinese prompts (**CH**)
- Consideration of language aspects: predicting holistic scores directly (**Total**) vs. predicting holistic scores with the sum of four aspect scores (**Asp**)
- Rationales: rating without rationales (**NR**) vs. rating with rationales (**R**)

The product of the parameters of these factors results in 16 combinations (GPT models) . We evaluated their performances with how similar the AES results are to human ratings by quadratic weighted Kappa (QWK) and Spearman's rho (ρ).

3. Results

Among human ratings, the average QWK and ρ score are respectively 0.843 (0.832-0.854) and 0.861 (0.841-0.894). The QWK and ρ scores of the GPT models are the average QWK and ρ scores of each model and the three human raters (see Table 1). The GPT models obtain scores ranging from 0.39 to 0.84 in QWK and 0.621 to 0.864 in ρ. One GPT QWK and eight GPT ρ scores achieve the human

standard (QWK 0.832 and ρ 0.841). We compare GPTs that only differ in a factor to reveal the factor's contribution to the overall performance (e.g., EN4_Total_R vs. CH4_Total_R for the impact of prompt languages), as in Table 2. For instance, 5 out of 8 pairs show that GPTs fed with EN prompts perform better than those with CH prompts, and 3 out of 8 indicate the opposite.

	EN4_Total_R	EN4_Total_NR	CH4_Total_NR	EN4_Asp_NR	EN3_Asp_R	EN4_Asp_R	CH3_Asp_R	CH4_Asp_NR
QWK	**0.838**	0.827	0.767	0.761	0.748	0.707	0.679	0.666
P	**0.864**	**0.842**	**0.848**	**0.859**	0.808	**0.858**	0.8	**0.849**
	CH4_Total_R	CH4_Asp_R	CH3_Asp_NR	EN3_Asp_NR	CH3_Total_NR	CH3_Total_R	EN3_Total_NR	EN3_Total_R
QWK	0.605	0.6	0.554	0.539	0.44	0.414	0.405	0 .388
P	**0.844**	**0.851**	0.703	0.756	0.732	0.621	0.724	0.705

Table 1: The QWK and ρ scores of GPTs

	EN >CH	CH >EN	G4>G3	G3 >G4	Asp >Total	Total >Asp	R >NR	NR >R
QWK	5/8	3/8	8/8	0/8	4/8	4/8	3/8	5/8
ρ	5/8	3/8	6/8	2/8	6/8	2/8	4/8	4/8

Table 2: The contribution of the four factors to the AES results

4. Discussion and Conclusion

The results reveal the GPTs' performance with different components, providing suggestions for improvement and application. First, GPT-4 with English prompts instructing the model to predict holistic scores with rationales achieves human standards (QWK, ρ). It can be a trustworthy copilot in grading essays or a trainer for improving learners' writing skills. Second, the GPT version and the language of prompts are crucial factors in AES tasks. Language teachers and learners are suggested to use GPT-4 with English prompts to obtain more accurate scores. Third, the contrasts of the other two factors (Asp vs. Total; R vs. NR) are unclear. As in Table 2, GPTs with Total and Asp predict equally good results in QWK (4 vs. 4); Rating only (NR) and rating with rationales (R) also perform similarly (4 vs. 4). Interestingly, if pondering the average of the differences of each pair, Asp performs significantly better than Total (0.218 vs. 0.076), and yet the contrast of ρ differences between R and NR is less noticeable (0.043 vs. 0.034). Considering the quality of the four aspects only slightly improves the AES results. Contrary to the mainstream that chain of thoughts may enhance the performance of GPT, generating rationales does not significantly improve AES results. Fourth, many GPT results achieve the human standards in ρ but not QWK. The contrast suggests that GPTs can predict scores in a high correlation with human ratings, but the accuracy still needs to be improved.

Selected References

Ke, Z., & Ng, V. (2019). Automated Essay Scoring: A Survey of the State of the Art. In *IJCAI* (19), 6300-6308.

Mizumoto, A., & Eguchi, M. (2023). Exploring the potential of using an AI language model for automated essay scoring. *Research Methods in Applied Linguistics*, 2(2), 100050.

Naismith, B., Mulcaire, P., & Burstein, J. (2023). Automated evaluation of written discourse coherence using GPT-4. In Proceedings of BEA 2023 (pp. 394-403).

Intuitive Minds vs. AI Algorithms:
Anaphoric Marker Choices in Korean Bridging Contexts

Sarah Park, Sang-Geun Lee

(Korea University, Korea)

Anaphoric markers are essential in maintaining coherence and clarity in discourse, and their choice can be influenced by various linguistic and pragmatic factors. The choice between bare nominals without 'ku' and demonstrative nominals with 'ku' in the Korean nominal system offers a unique opportunity to explore these influences. This study aims to explore the distinct strategies employed by naive Korean speakers and AI models, specifically ChatGPT, in selecting Korean anaphoric markers, especially, in different bridging contexts. The focus is on two key linguistic conditions: types of bridging references (whole-part vs. product-producer) and types of subject markers (topic marker -(n)un vs. nominative marker -i/-ka).

Cond1 (Semantic relations)	Cond2 (Subject markers)	Sentence Items
Whole-Part	Topic	진수가 집을 사려고 한 아파트를 방문하였다. 그 거실은 놀라울 정도로 넓었다.
	Nom	진수가 집을 사려고 한 아파트를 방문하였다. 그 거실이 놀라울 정도로 넓었다.
Product-Producer	Topic	진수가 길거리에서 풍경화 한 점을 치는 샀다. 그 화가는 폭풍우 바다를 그렸다.
	Nom	진수가 길거리에서 풍경화 한 점을 치는 샀다. 그 화가가 폭풍우 바다를 그렸다.

Although OpenAI models have demonstrated remarkable proficiency across a broad spectrum of linguistic tasks, their ability to replicate human-like judgment mechanisms remains uncertain. Building upon the study (Lee, Ma, & Kim 2023) that collected Korean speakers' acceptability judgments on 24 sets (1set with 4 items) of bridging contexts, we further conducted a comparative study on the strategies employed by naïve Korean speakers and AI models (specifically ChatGPT) in selecting Korean anaphoric markers – bare nominals without 'ku' and demonstrative nominals with 'ku'. It will help us to understand how efficiency and preferences in anaphoric marker choices vary between the two groups, guided by the economy principle that language users tend to communicate efficiently by choosing fewer words and simpler structures.

This study employs a mixed-method approach comprising two main tasks: a Cloze task and a rating task. In the Cloze task, ChatGPT is presented with the same 24 sets of Korean complex sentences featuring blanks for the demonstrative 'ku'. These sentences are structured to test both whole-part and product-producer bridging references, combined with topic and nominative subject markers. ChatGPT fills in the blanks with the demonstrative 'ku' if necessary, and its choices are analyzed to determine its strategy in anaphoric resolution. In the rating task, ChatGPT rates the naturalness of pre-filled sentences with 'ku' on a Likert scale from 1 (least natural) to 5 (most natural). The sentences used are the same as those in the Cloze task, allowing for a direct comparison between the two tasks. Statistical analyses, including ANOVA and paired t-tests, are conducted to compare the choices and ratings between humans

and ChatGPT.

The previous findings on Korean speakers' judgment (Lee, Ma, & Kim 2023) suggest that the relation of whole-part is often straightforward and hence easily predictable to Korean speakers, with a clear inclusion of the part in the whole. This simplicity allows them to choose a simpler anaphoric reference of bare nominals without 'ku'. On the other hand, the relation of product-producer is more intricate and less straightforward and hence less predictable to Korean speakers. This obscurity leads them to choose an additional demonstrative marker like 'ku' to explicitly indicate the anaphoric link between the product and its producer. This pattern detected in naïve Korean speakers' judgment will be compared against ChatGPT's choices to identify any discrepancies in anaphoric resolution strategies. This study will measure the convergence rates between the judgments of Korean speakers and ChatGPT.

The statistical analyses are expected to reveal significant differences in the anaphoric marker choices between humans and AI, particularly in how each group navigates the efficiency and clarity trade-off in different linguistic conditions. This comparative analysis provides several critical insights. By examining the strategies humans and AI use in anaphoric resolution, this study contributes to our understanding of how the economy principle is distinctively perceived by the two groups, especially, in article-less languages like Korean.

The findings will highlight the current limitations of AI models in replicating human-like language processing, guiding the development of more sophisticated AI language models that better mimic human linguistic behavior. Additionally, this research bridges the gap between human cognitive processes and AI algorithms, offering valuable data on how different entities approach the task of maintaining coherence in discourse. Insights from this study can be applied to improve AI-driven language learning tools, translation services, and other applications where natural language processing is crucial.

Selected References

Ahn, D. 2019. THAT thesis: A Competition Mechanism for Anaphoric. Doctoral dissertation, Harvard University.

Dayal, V. & J. Jiang. 2022. The Puzzle of Anaphoric Bare Nouns in Mandarin: A Counterpoint to Index!
    Linguistic Inquiry 54: 147–167.

Devitt, M. 2006. Intuitions in linguistics. *British Journal for the Philosophy of Science* 57: 481–513.

Jenks, P. 2018. Articulated definiteness without articles. *Linguistic inquiry* 49: 501-536.

Kang, A. 2021. Marking definiteness in an articleless language: The role of the domain restrictor KU in Korean. *Language and Linguistics* 22: 301-336.

Kim, M-J. 2023. Bare vs. demonstrative anaphoric definites in Korean and their cross-linguistic implications. *Japanese/Korean Linguistics* 30: 127-141.

Lee, S, J. Ma, & H. Kim 2023. An experimental study on the definiteness of two types of Korean nominals in bridging references. *Korean Journal of Linguistics* 48: 785 – 811.

Schwarz, F. 2013. Two kinds of definites cross-linguistically. *Language and Linguistics Compass* 7: 534-559.

Prompting Strategies of Generative AI for Korean Pragmatic Inference

Ye-eun Cho

(Sungkyunkwan University, Korea)

Background: Understanding an utterance requires not just knowing the meaning of words and grammar but also understanding the communicative features of language. Grice (1975) distinguished between 'what is said', which refers to truth-conditional, and 'what is implicated', which refers to non-truth-conditional (Grice, 1975). Between two types of 'what is implicated', such as conventional and conversational implicature, we will focus on conversational implicature, which depends on the idea that the meaning of utterance is context-dependent as well as that the speaker is following Cooperative principle. This principle suggests that conversation can work because the interlocutors are trying to contribute to the conversation. This principle is categorized into four maxims as follows:

- **The Maxim of Quantity:** Do not make your contribution more informative than is required.
- **The Maxim of Quality:** Do not say what you believe to be false.
- **The Maxim of Relation:** Make sure your contributions are relevant to the conversation.
- **The Maxim of Manner:** Be brief; Be orderly; Avoid obscurity and ambiguity.

When using generative AI, it is crucial to effectively modify input samples through prompting strategies in order to enhance output quality. With standard prompting as a baseline, the prompts of the key prompting strategies commonly used are provided as follows:

- **Standard prompt** (E) involves the question (Q) and the answer (A) ⬚ i.e., $E_j = (Q_j, A_j)$.
- **Chain-of-thought (CoT) prompt** (E) is composed of the question (Q), the reasoning step (T), and the final answer (A), as an in-context learning. The intermediate reasoning steps enable the models to solve tasks in a step-by-step manner (Wei et al., 2022) ⬚ i.e., $E_j = (Q_j, T_j, A_j)$
- **Contrastive-CoT prompt** consists of one question along with sets of the reasoning step and the answer in positive (+) and negative (-) versions (Chia et al., 2023) ⬚ i.e., $E_j = (Q_j, T_{j,+}, A_{j,+}, T_{j,-}, A_{j,-})$.
- **Zero-shot-CoT prompt (**Kojima et al., 2022) is composed of the question and one consistent reasoning step T_{zero} without providing the answer ⬚ i.e., $E_j = (Q_i, T_{zero})$.

Previous studies on prompting strategies often focused on how generative AIs are engaged in truth-conditional tasks, such as arithmetic or logical tasks. While a few studies have explored prompting strategies focusing on conversational implicature, these have been restricted to either English or a single prompting strategy. Therefore, this study will investigate prompting strategies appropriate for understanding the conversational implicature, which is non-truth-conditional and context-dependent, particularly within the Korean context. Additionally, we will see the output patterns depending on four maxims of Grice's Cooperative principle and generative AI models.

Experiment: The dataset used for the experiment is 'Korean pragmatic test set (Park et al., 2024)', consisting of total 120 test units, with 30 units allocated to each maxim. Between two types of datasets, multiple-choice questions (MCQs) and open-ended questions (OEQs), the MCQs are only used for the experiment. Each unit contains the given context, the statement, and four multiple-choice options. The selected models include GPT-3.5, GPT-4o, Gemini-1.0, Gemini-1.5, Llama3-8B, and Clova-X. These models were provided with four types of prompts, including standard prompt, zero-shot-CoT prompt, CoT prompt, and contrastive-CoT prompt, with the task of selecting one of the four MCQ options as the answer.

Result: As an evaluation metric, **'accuracy'** was defined as the proportion of correct answers from the total number of MCQs. That is, prompting strategies that achieved higher accuracy than the standard

prompt (baseline) were considered effective. Notably, the accuracy of contrastive-CoT prompting should be higher than that of the CoT prompting as well as that of the standard prompting. The accuracy of prompting strategies by each model and maxim is presented in the tables below.

  Regarding the prompting strategies, CoT prompting showed the most effective results, while zero-shot-CoT prompting and contrastive-CoT showed either minimal effectiveness or occasionally performed worse than CoT prompting. Regarding the maxims, the Maxim of Quality showed the highest accuracy since the examples include clear context showing irony or sarcasm. In contrast, the maxim of Manner exhibited the least consistency in results because the meaning of conversational implicature is not clearly revealed even when context is considered. In terms of models, advanced models like GPT-4o and Gemini-1.5 consistently showed higher accuracy compared to base versions like GPT-3.5 and Gemini-1.0 within the same model line. Since these advanced models have sophisticated architectures capable of handling complex language tasks, CoT prompting aligns well with these capabilities. On the other hand, Clova-X has a robust inherent capability for understanding Korean pragmatics, reducing the relative impact of different prompting strategies. Furthermore, despite some improvements with CoT prompting, the accuracy of GPT-3.5, Gemini-1.0, and Llama3 remained lower than that of the other models. It is also essential to highlight the overall enhancement in capabilities of a model itself, regardless of prompting strategies. Advanced models like GPT-4o, Gemini-1.5, and the Korean-specialized model Clova-X accurately interpreted some contexts of conversational implicature better than the previous study (Park et al., 2024) even without prompting strategies.

Model	Standard	Zero-shot	CoT	Contrastive				
	Overall maxims							
GPT-3.5	45.8	42.5	**53.3**	52.5				
GPT-4o	85.0	78.3	**90.0**	87.5				
Gemini-1.0	68.3	65.0	67.5	68.3				
Gemini-1.5	82.5	**84.2**	**95.0**	92.5				
Llama3	35.8	**38.3**	**40.0**	35.0				
Clova-X	81.7	**82.5**	**83.3**	82.5				
	The maxim of Quantity				**The maxim of Quality**			
GPT-3.5	43.3	36.7	**46.7**	43.3	53.3	50.0	**66.7**	**70.0**
GPT-4o	80.0	76.7	**90.0**	86.7	96.7	90.0	**100.0**	96.7
Gemini-1.0	73.3	66.7	70.0	66.7	80.0	**86.7**	**83.3**	83.3
Gemini-1.5	80.0	80.0	**90.0**	**93.3**	96.7	96.7	96.7	96.7
Llama3	33.3	**40.0**	33.3	36.7	43.3	43.3	**53.3**	53.3
Clova-X	83.3	83.3	**90.0**	83.3	96.7	96.7	96.7	96.7
	The maxim of Relation				**The maxim of Manner**			
GPT-3.5	43.3	43.3	**53.3**	50.0	43.3	40.0	**46.7**	46.7
GPT-4o	83.3	73.3	**93.3**	90.0	80.0	73.3	76.7	76.7
Gemini-1.0	66.7	56.7	66.7	63.3	53.3	50.0	50.0	**60.0**
Gemini-1.5	80.0	**83.3**	**96.7**	96.7	73.3	**76.7**	**96.7**	83.3

| Llama3 | 33.3 | 33.3 | **36.7** | 23.3 | 33.3 | **36.7** | **36.7** | 26.7 |
| Clova-X | 83.3 | 76.7 | 80.0 | 80.0 | 63.3 | **73.3** | **66.7** | 70.0 |

Selected References

Grice, H. P. 1975. "Logic and conversation". In Speech acts (pp. 41-58). Brill.

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., V. Le Q. & Zhou, D. 2022. "Chain-of-thought prompting elicits reasoning in large language models," *Advances in neural information processing systems*, 35, 24824-24837.

Chia, Y. K., Chen, G., Tuan, L. A., Poria, S., & Bing, L. 2023. "Contrastive chain-of-thought prompting," *arXiv preprint* arXiv:2311.09277.

Kojima, T., Gu, S. S., Reid, M., Matsuo, Y., & Iwasawa, Y. 2022. "Large language models are zero-shot reasoners." *Advances in neural information processing systems*, 35, 22199-22213.

Park, D., Lee, J., Jeong, H., Park, S., & Lee, S. 2024. "Pragmatic Competence Evaluation of Large Language Models for Korean," *arXiv preprint* arXiv:2403.12675.

Why do humans feel unsatisfied with conversations with unsupervised chatbots?: The comparative study on the topical depth using T-R nexus and FSP

Kaoru Amino

(ShanghaiTech University, China/Japan)

Chatbots referred to as "supervised AI" serve specific interactional purposes in fields such as medicine and service-oriented industries. However, chatbots that engage in aimless small talk fall under the category of "unsupervised AI" lacking a defined measure of success. If they do have a purpose, it is often for entertainment, providing comfort to lonely users or entertaining human interlocutors. These unsupervised chatbots exhibit distinct conversational characteristics that differ from human conversations, including an unnatural flow compared to small talk between humans.

In a study analyzing feedback within the IRF exchange structure framework (Sinclair and Coulthard, 1975) in a chatbot called "Replica" (created by QUARTZ), it was observed that the chatbot frequently employed highly versatile words like "Great" and "Good" in its feedback acts, while there were relatively few extended feedback acts in the form of comments (Amino et al., 2023). This suggests that chatbots are unable to serve as a substitute for a meaningful human interlocutor.

Additionally, it is widely recognized that effective communication entails deep conversation, where discussion points can be clearly articulated, and the conversation can progress smoothly by accurately understanding the conveyed message (Well-being workers®, 2023). From the perspective of discourse analysis, Weil (1987) stated that conversation begins on the common ground upon which two intelligences meet and progress toward another part of discourse, which forms the statement and that has a high level of possibility to be related to the concepts of Theme (known and shared information) and Rheme (unknown information). Fishman (1977) also argued that maintaining topic continuity is a characteristic of supportive conversation, avoiding sudden topic shifts. Hence, it can be inferred that the persistence and development of topics, as well as the depth of discussion, play a crucial role in the enjoyment of a conversation.

Regarding topic shifts, Reichman (1978) proposed the concept of "context space" and argued that a topic is not formed solely through a linear sequence but instead encompasses a hierarchical context space with an inner structure. These context spaces progress gradually to the next renewed space.

Regarding the progression of Theme and Rheme, a framework known as the T-R nexus was introduced by Danes (1974) to analyze this "context space" from the perspective of a topic tree. This model was further developed into Functional Sentence Perspective (FSP), which categorizes thematic progression patterns into themes (known information) and rhema (unknown information). It further distinguishes an additional three categories according to the topical progress pattern: 1) Simple Linear Thematic Progression (SLP); 2) Thematic Progression with a Continuous Theme (TPC); and 3) Thematic Progression with Derived Themes (TPD), as illustrated in Figure 1.

FSP Type1: SLP   FSP Type 2: TPC   FSP Type 3: TPD
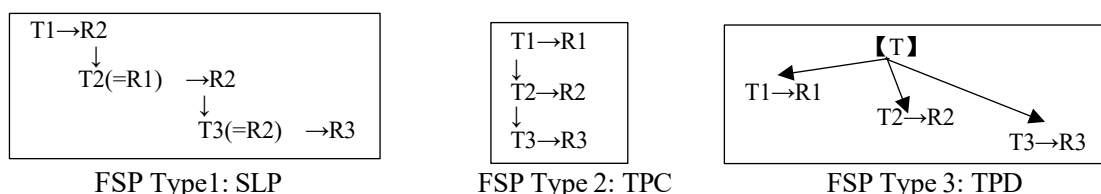
Figure 1. Three models of FSP

Meanwhile, Murakami and Kumatoridani (1995) examined the typology of topical shift and put forward sequential categories of new, derivation, and regeneration.

This research used the T-R nexus to measure the shallowness of conversation and topical communication depth as a scale of satisfying and deep conversation by applying three categories of topical type, in addition to three topical indicators provided by Murakami and Kumatoridani (1995). The purpose of this study was to assess the depth of topical flow in conversation, examining the progression of conversation by utilizing the T-R nexus framework.

Based on conversational data from conversation of "Replica" with a human interlocutor and scripts from "Male language in the workplace," which consist of 500 turns each, qualitative and quantitative observations were conducted to analyze the topic progression using FSP and the framework of new, derivation, and regeneration. (Regarding the latter dataset, it is important to note that nearly half of the participants were women, and approximately one-third of the conversations involved small talk during breaks.)

The results of the study show that the "Replica" chatbot predominantly introduced new topics, while derivative and regenerated topics were relatively rare. This indicates that the topical flow in the chatbot's conversations primarily relies on the introduction of completely new topics. Even when derivative types of topics (SLP, TPD) are present, most of them exhibit TPC, in which the themes are not relevant to each other, as shown in Figure 2.

Conversely, human conversations display a more complex transition between themes and rhema, with fewer instances of entirely new topics. Among the derivative-type topics, SLP and TPD were frequently observed, intricately intertwined with TPC. In addition, the conversation transitioned between themes and rhema and the thematization process shifted back and forth with the original theme "regeneration," as shown in Figure 3.
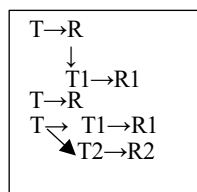
Figure 2. FSP type of "Replica" chatbot
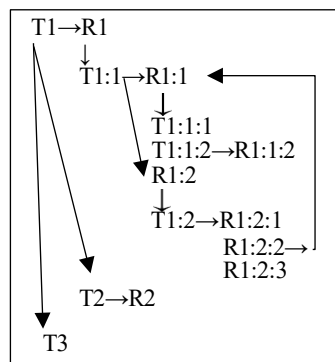
Figure 3. FSP type of human interlocutor

Human small talk tends to be more complex, by deviating from the main topic and freely shifting to topics of interest, rather than following a strictly organized progression.

Based on these findings, it can be concluded that conversations between humans demonstrate a depth of topic development characterized by the development of context space through new perspectives and framing. The disparity in topical shifts between chatbots and humans suggests potential areas for the development of unsupervised chatbot architectures.

Selected References

Amino, Kaoru., Chung, Mengyen., Zhou, Ruohong. 2023. "The analysis of conversational error on the unsupervised chatbot: Applying linguistic frameworks of IRF and its 6 acts," *Bulletin of the Society of East Asian Language and Culture Studies* 2,1-23.

Danes, Frantisek. 1974. "Functional sentence perspective and the organization of the text. Functional Sentence Perspective, The Hague": Monton, 106-128.

Fishman, Pamela. 1977. Interactional shitwork. *A Feminist Publication on Arts and Politics* 2, 99-101.

Murakami, Megumi, Kumatoridani, Tetsuo. 1995. "Cohesion in discourse topics and topic development structure," *Hyogen Kenkyu* 62, 101-111.

Reichman, Rachel. 1978. "Conversational coherency," *A Multidisciplinary Journal* 2, 283-327.

Sinclair, John, Coulthard, Malcolm. (1975) Towards an analysis of discourse. Oxford: Oxford University Press.

Weil, Henry. 1987. The order of words. Trans. by Charles W. Super. Boston: Ginn and Company.

Does Generative AI Dream of human-like Translation?
Implementing Obviation Control Rules into an LLM through Finetuning

Hisashi Morita

(Aichi Prefectural University, Japan)

[Introduction] Finding the right referents for anaphoric elements such as pronouns and definite expressions in texts, so called anaphora resolution, has been one of the most important and challenging topics in the field of computational linguistics. The research has gone through several phases since 1960s. Initially, researchers were "ambitious, theoretically oriented" (Mitkov 2002). However, they were domain-specific (e.g. targeting pronouns only or definite expressions only), required manual preprocessing, and lacked objective evaluation and reliable or robust applications. Then the research trend turned to the next phase (late 1990s) in which proposals were easy to implement and evaluate. It is noteworthy that they made use of large-scale corpora for evaluation while most of them did not incorporate theoretical, in particular, syntactic findings due to practical reasons: they are called "knowledge-poor" strategies (Mitkov 2002). Their methods and results continue to improve using machine learning. However, recent advancement in generative AI or LLMs (Large Language Models) such as ChatGPT and Claude seem to overcome anaphora resolution in translation miraculously, which may give one an impression that the research in computational (as well as theoretical linguistics) is no longer necessary.

This incredible feat is observed even when translating Japanese into English, the former of which allows zero pronouns in subject and object positions. Translating such an argument-ellipsis language into another language with no argument ellipsis option has been one of the hardest problems in computational linguistics: The translation machine somehow must recover what is not present in the original text. Nevertheless, generative AI seems to resolve tasks of recovering zero pronouns very easily.
[Problems] Despite the great success, generative AI has a difficulty in certain types of translation: obviation control phenomena. Currently, LLMs are likely to mistranslate the following types of Japanese into English:

(1) a. *pro* samu-**gar**-te.iru.   b. pro samu-i.
     pro feel.cold-gar-NonPast   pro feel.cold-NonPast
     'A person other than the speaker feels cold.'   'I feel cold.'
     AI translation: *"I feel cold."   AI translation: "I feel cold."

(2) John$_i$-wa [*pro*$_{*i/j}$ hasiri-**ta-gar**-te.iru   to] omow-ta. John-Top pro   run-want-gar-
     NonPast C think-Past 'John thought a person other than John wanted to run.' AI translation:
     *"John$_i$ thought he$_i$ wanted to run."

Japanese has subjective or personal experience predicates such as *samu* 'feel.cold' and *ta* 'want'. Generally, when subjects are covert, they are likely to denote 1st person in Japanese, and this strategy is exactly what LLMs adopt when transforming a Japanese example into an English translation as in (1)*b*. However, when such predicates are selected by *gar*, they exhibit the character of obviation control: in matrix clauses, the omitted subject is never the speaker as in (1)*a* while, in embedded, it is never the matrix subject as in (2). LLMs are initially unaware of such obviation rules, so they are likely to generate the wrong translations.

Such errors are attributed to the way LLMs are trained: the attention mechanism (Vaswani et al, 2017: 1st version). They are trained so that they predict what word is most likely to occur next. It is a type of collocation, but it is applied not only intra-sententially but also inter-sententially, the latter of which makes it possible for zero pronominals to refer to a Topic established in the preceding text. This

is why anaphora resolution is successful during translation. However, without preceding sentences, AI translates missing subjects or objects in Japanese examples into "I" or "me" in English in matrix clauses, hence, mistakes such as (1)*a*. There is another type of mistranslation by AI. Direct or indirect objects of addressive predicates such as *hanasikake* 'speak.to' and *shotaisur* 'invite' cannot refer to the speaker in the matrix clause (Hasegawa 2007) as in (3) or the matrix subject when embedded as in (4).

(3)   Bob-ga   {*pro*$_{*speaker}$/*watashi-ni} hanasikake-ta.
      Bob-Nom {pro/me-to}   speak.to-Past
      'Bob spoke to {somebody/*me}.'   AI translation: *"Bob spoke to <u>me</u>."

(4)   Mary$_i$-wa [Bob-ga *pro*$_{*i/j}$ hanasikake-ta to] omow-ta. John-Top Bob-Nom pro speak.to-Past
                                 C think-Past 'Mary thought Bob spoke to a person other than Mary.' AI translation: *"Mary thought Bob spoke to <u>her</u>."

This is another obviation control, and LLMs fail to translate them into English.

Such AI's mistakes are explainable because AI is exposed to only actual uses of language, i.e. Chomsky's E-language; accordingly, they cannot be equipped with I-language. In other words, they cannot judge whether a sentence is grammatical or not. Moreover, we do not understand AI's algorithms, so called the black box problem of AI; therefore, we are yet to know the way to implement negative rules such as obviation into their models directly.

[The aim and the organization] In this presentation I would like to show whether LLMs can learn to correctly translate zero pronominals in obviation control in Japanese into English by training a pretrained language model, so called finetuning. First, I will make a brief introduction of the history of anaphora resolution in computational linguistics. Then I will present a theoretical and novel approach to the obviation phenomenon in Japanese, an indexical shift approach in the framework of generative grammar. I will argue that *gar* (subject obviation) and *hanasikake* (object obviation) should be regarded as one type of shifty indexicals (Morita and Ramchand 2021; Morita 2024). Finally, I will present a series of experiments and their results: I have finetuned large language models, one of which is Elyza- japanese-Llama-2-7b (Sasaki et al 2023), with LoRA (Low Rank Adaptation of Large Language Models) through small sized data of Japanese and English pairs. I have finetuned the model with examples of *gar* in matrix and embedded context first, and then *hanasikake*. I will present several tips to enhance its performance, such as formatting of training data and prompts, use and positioning of syntactic tags, the parameters of learning rates and the number of epochs, and so on.

[Consequences] Theoretical findings (syntax, in particular) have been neglected since late 1990s in the computational linguistics and the trend continues even in the time of AI. However, the outcome of the present research demonstrates that theoretical findings can be employed to improve performance of (finetuned) LLMs. Moreover, the method proposed in this presentation can be applied to LLMs with more parameters and other shifty indexical languages easily, so it can contribute to creating more reliable LLMs.

Selected references

Hasegawa, N. (2007) "Modality-to Ninsho-Seigen (Modality and Person restriction)," in N. Hasegawa (ed.) *Nihongo-no Shubun Gensho (Root phenomena in Japanese)*, Hituzi Shobo.

Kameyama, M. (1985) *Zero Anaphora: The case of Japanese*, doctoral dissertation, Stanford University. Mitkov, R. (2002) *Anaphora Resolution*, Pearson Education, London.

Morita, H. (2024) "Indexical and Perspectival Shift in Japanese," a poster presentation at *SICOGG26/WAFL18*.

Morita, H. and Ramchand, G. (2021) "Compositionality and Context Shift in Japanese-An Account

without Monsters," manuscript, Aichi Prefectural University/University of Tromso.

Sasaki, A. et al (2023) ELYZA-japanese-Llama-2-7b, downloaded on 26th of March, 2024, from
https://huggingface.co/elyza/ELYZA-japanese-Llama-2-7b.

Sells, P. (1987) "Aspects of logophoricity," *LI* 19: 445-479

Vaswani, A. et al (2017: 1st version) "Attention is All you Need," in Guyon et al (eds.) *Advances in Neural Information Processing Systems*.

AI Assessment Tools in Language Learning: Bridging the Gap Between Theory and Practice

Jose Belda-Medina, Sandra Abad-Bataller, Kaoutar Smaili-Bensaad,
Andrea Gallardo-Ferrandez, Jorge Soto-Almela, Jose R. Calvo-Ferrer
(University of Alicante, Spain)

AI assessment tools are increasingly utilized in language learning to automate the assessment process, thus easing the workload for educators and enabling them to dedicate more attention to individual student needs (Alhrabi, 2023; Huang et al., 2023; Fagbohun et al., 2024). This research explores the effectiveness of using AI tools for essay correction compared to human grading practices among in-service teachers. The study aims to assess the integration of these tools by examining in-service teachers' prior knowledge of such tools, comparing corrections provided by teachers and AI based on established rubric criteria, and examining teachers' knowledge and perceptions of this new technology. Based on convenience sampling, 52 teachers from the Department of English Studies participated in this study. Employing a mixed- methods approach, the research utilized pre-post surveys to gather quantitative data focusing on teachers' technological affinities and perceptions of AI grading tools. Qualitative data were collected through written corrections of essays, evaluating adherence to a rubric covering argument strength, language use, coherence, and originality. The correction process consisted of three stages: pre-survey administration, intervention involving the correction of three short essays by both teachers and three AI grading tools, and a post-survey, along with semi- structured discussions, to assess changes in teachers' perceptions. Results indicated initial unawareness, reluctance, and distrust among in-service teachers toward utilizing such tools.

However, their acceptance grew moderately over time, primarily due to the considerable time saved and the increased systematic approach in the grading process facilitated by AI tools. Nonetheless, participants also highlighted certain drawbacks, including the potential for bias and subjectivity in grading. This study sheds light on how teachers can effectively integrate AI assessment tools into their teaching practice, advocating for better preparation in teacher training programs, and underscores the importance of understanding and addressing both the advantages and limitations associated with the adoption of AI tools in educational settings.

References

Alharbi, W. (2023). AI in the Foreign Language Classroom: A Pedagogical Overview of Automated Writing Assistance Tools. *Education Research International*, vol 2023, 1-15.

Fagbohun, O., Iduwe, N. P., Abdullahi, M., Ifaturoti, A., & Nwanna, O. M. (2024). Beyond Traditional Assessment: Exploring the Impact of Large Language Models on Grading Practices. *Journal of Artificial Intelligence, Machine Learning & Data Sci*ences, 2(1), 1-8.

Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023). Trends, research issues and applications of artificial intelligence in language education. *Educational Technology & Society*, 26(1), 112-131.

Themed Sessions
(Day 2)

Wh-doubling Constructions in Korean Sign Language

Haewon Jeon

(Korea Maritime and Ocean University, Korea)

Wh-doubling is a widely observed phenomenon in sign languages. Since the early days of the discipline, the syntax of wh-questions has been a highly researched topic in sign language linguistics. One notable aspect is that wh-phrases in numerous sign languages do not appear in a clause-initial position, as is commonly observed in most spoken languages that display overt wh-movement. In contrast, wh-phrases are frequently observed at the end of a clause in numerous sign languages worldwide, including German Sign Language (DGS) (Happ & Vorköper, 2014) and Korean Sign Language (KSL) (Jhang, 1999). Additionally, many sign languages exhibit the phenomenon of wh-doubling, where a wh-phrase appears twice in a question—both in clause-initial and clause-final positions. This phenomenon has been documented in DGS, American Sign Language (Petronio, 1993), Brazilian Sign Language (Nunes & Quadros, 2005), Italian Sign Language (Branchini et al., 2013), Flemish Sign Language (Pfau & Bos, 2016), and Russian Sign Language (Kimmelman, 2013), among others.

Despite the existing research, the theoretical analysis of KSL grammar remains limited. Given its status as a natural language, KSL is expected to exhibit syntactic characteristics that can be explained by linguistic theories seeking to uncover universal elements of human language. This research aims to analyze the syntactic properties of wh-doubling constructions in KSL using a generative approach to human language faculty.

(1) a. <u>   wh</u>
           WHAT PAUL BUY WHAT   'What did Paul buy?' (Bross, 2024)
   b. <u>   wh</u>
      <무엇> <철수> <샀다> <무엇>   'What did 철수 buy?' (Jhang, 1999)

In some cases, one of the doubles is base-generated, either the clause-initial one (e.g., Neidle, Kegl & Bahan, 1994) or the clause-final one (e.g., Petronio & Lillo-Martin, 1997). According to remnant-movement accounts, the clause-final double is a non-deleted copy of the wh-phrase that has been left behind by an intermediate movement (e.g., Bross, 2020b). One of the key challenges faced by these accounts is the distinct data pattern associated with wh-doubling. The primary issue that is extensively debated is that while the ultimate double can function as a simple wh-phrase (e.g., WHAT), it is incapable of serving as a complex wh-phrase (e.g., WHICH COMPUTER).

(2) a. *<u>   wh</u>
      WHICH COMPUTER PAUL BUY WHICH COMPUTER
      'Which computer did Paul buy?'
   b. <u>   wh</u>
      WHICH COMPUTER PAUL BUY WHICH
      'Which computer did Paul buy?'   (Bross, 2024:3)

   c. *_____ wh

       <어제> <철수> <누구> <차> <고치다> <누구> <차>

       'Whose car did 철수 fix yesterday?'

   d. _____ wh

       <어제> <철수> <누구> <차> <고치다> <누구>

       'Whose car did 철수 fix yesterday?'   (Jhang, 1999:18)

Selected References

Branchini, Chiara, Anna Cardinaletti, Carlo Cecchetto, Caterina Donati & Carlo Geraci. 2013. Wh-duplication in Italian Sign Language (LIS). *Sign Language & Linguistics* 16(2). 157–188.

Bross, Fabian. 2024. Wh-doubling in German Sign Language: Why not sluicing? In: Sign Language & Linguistics, 26(2).

Bross, Fabian. 2020a. Encoding different types of topics and foci in German Sign Language: A cartographic approach to sign language syntax. *Glossa: A Journal of General Linguistics* 5(1). 108.

Bross, Fabian. 2020b. *The clausal syntax of German Sign Language: A cartographic approach*. Berlin: Language Science Press.

Happ, Daniela & Marc-Oliver Vorköper. 2014. *Deutsche Gebärdensprache: Ein Lehr- und Arbeitsbuch [German Sign Language: A text- and workbook]*. Frankfurt am Main: Fachhochschulverlag.

Jhang, Se-Eun. 1999. 미국수화언어와 한국수화언어에서의 WH-의문문. 언어학, 7(3), 297-319.

Kimmelman, Vadim. 2013. Doubling in RSL and NGT: A pragmatic account. In *Information structure: Empirical perspectives on theory*, 99–118.

Nunes, Jairo & Ronice Müller de Quadros. 2005. Duplication of wh-elements in Brazilian Sign Language. In Leah Bateman & Cherlon Ussery (eds.), *Proceedings of NELS*, 35, 463. Amherst, MA: GLSA

Petronio, Karen. 1993. *Clause structure in American Sign Language*. Washington: University of Washington dissertation.

Petronio, Karen & Diane Lillo-Martin. 1997. WH-movement and the position of Spec-CP: Evidence from American Sign Language. *Language* 73. 18–57.

Pfau, Roland & Heleen Bos. 2016. Syntax: Simple sentences. In Anne Baker, Beppie van den Bogaerde, Roland Pfau & Trude Schermer (eds.), *The linguistics of sign languages: An introduction,* 117–147. Amsterdam: John Benjamins.

Pfau, Roland & Markus Steinbach. 2005. Relative clauses in German Sign Language: Extraposition and reconstruction. In Leah Bateman & Cherlon Ussery (eds.), *Proceedings of NELS* 35, 507–521.

Zeshan, Ulrike. 2004. Interrogative constructions in signed languages: Crosslinguistic perspectives. *Language* 80. 7–39.

Epistemic Modality Expression in Korean Sign Language

Ahmed Mai, Soon-Bok Kwon
(Pusan National University, Korea)

Modality is essentially a subjective personal attitude expression towards an utterance (Palmer 1986). Modality markers are classically classified into epistemic modality and deontic modality, where the first is concerned with expressions of possibility and uncertainty the latter is a classification of expressions of obligation and duty (Cornillie & Pietrandrea 2012). Wilcox and Shaffer (2006) pointed out that signed languages can be analyzed phonologically, morphologically and even syntax wise, they also noted how a part of grammar develops over time from a lexical sign. To give example for that, they explained how in the American Sign Language the modal auxiliary verb 'can' has developed from the sign for 'strong' which reflects ability. But while modality has been investigated in Korean spoken language and other sign languages, it has not been researched in KSL. In this paper we are investigating the expression epistemic modality, specifically expression of assumption or uncertainty in KSL.

Method

We searched for the expressions of epistemic modality in KSL using the Korean Sign Language Dictionary by the National Institute for Korean Language and the Korean Sign Language Dictionary by Kim et al. (2005). And to look for the sign we have derived a list of modality expressions used in spoken Korean, as both dictionaries are not annotated or concerned with classification of grammatical expressions. The list was based on previous studies of modality expression in spoken Korean (such as, Ahn 2012 etc.). An analysis of marker composition was then carried out to find out the characteristics of uncertainty expression in KSL.
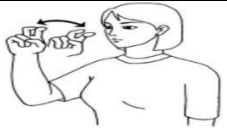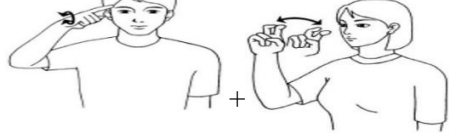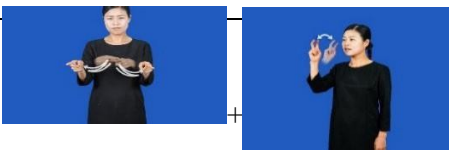
Result

Table 1 demonstrates the signage for uncertainty and assumption in KSL. As indicated in the table the sign for "ama" MAYBE (Table 1 upper left corner) is part of composition of many signs that express epistemic modality. Signs such as ASSUME is a compound of the sign for "pyeongka" EVALUATE + "ama" MAYBE, the sign for "teutsipta" SEEM is also a compound of "saengkak" TO THINK + "ama" MAYBE. The placement of "ama" MAYBE after the action lexical is consistent with modality marker placement in the sentence as it follows the verb in KSL (Jo et al. 2022). This is also consistent with what Wilcox and Shaffer (2006) have discussed about the grammaticization of a sign as it first starts as a lexical sign which then changes into a grammatical meaning, as in "ama" is modifying the meaning of the lexical to make up the modal marker.

Conclusion

The sign for "ama" plays a big role in expressing epistemic modality in discourse. It is also an important factor in the modality marker grammaticization process. An additional note about epistemic modality markers in both KSL and spoken Korean is that KSL shows limitation in modality expression in comparison to spoken Korean. To further investigate this, a study of a KSL corpus is needed to investigate actual use of modality markers in KSL discourse. This will be the research aim in further study.

Table 1. signs that convey uncertainty and assumption in KSK

Marker	Sign	Marker	Sign
-l (-(eu)l keosida) ** Apeuro + ama (MAYBE)	+	ama* **, -(eu)l keosida*, -ket- (MAYBE)*	
teuthada*, teutsipta* (SEEM)	+	-ket-** (WILL)	
Chucheuk** Jimjak** saengkak+ama (GUESS)	+	-(eu)l su itta* ** kaneunghada* su** (CAN)	
Kajeong** pyeongka+ama (ASSUME)	+		

Note: *: the National Institute for Korean Language dictionary, **: Kim et al. (2005) dictionary (signs' illustrations are collected from the National Institute for Korean Language dictionary)

References

Ahn, Joo-Ho. 2012. *Korean Education and Modality Expression*. Seoul: Hankookmunhwasa.

Cornillie, Bert and Pietrandrea, Paola. 2012. "Modality at work. Cognitive, Interactional and Textual
  Functions of Modal Markers." *Pragmatics 44*(15), 2109-2115.

Jo, Jinwoo, Lee, Bo-Kyung and Choi, Youngju. 2022. "Word order and clause structure in Korean Sign
  Language." *Studies in Generative Grammar* 32(1), 97-152.

Kim, Samchan, Lee, Seonho, et al. 2005. *Korean Sign Language Dictionary*. Gyeonggi: Hyungseul.

Palmer, Frank R. 1986. *Mood and Modality*. Cambridge: Cambridge University Press.

Wilcox, Sherman and Shaffer, Barbara. 2006. *Modality in American Sign Language. In The Expression
  of Modality*. Frawley, W. (ed.), Berlin: Mouton De Gruyter.

National Institute for Korean Language. Korean Sign Language
        Dic.: https://sldict.korean.go.kr/front/main/main.do#

Double Mapping Constraint and KSL Metaphors

Youngju Choi
(Chosun University, Korea)

Introduction: Meir (2010) and Meir and Cohen (2018) argue that due to the Double Mapping Constraint (DMC), metaphors, which are prevalently observed in spoken languages, are not possible in sign languages when a lexical sign is iconic using body parts that are closely related to the meaning of the sign. Meir (2010) provides evidence such as EAT vs. NIBBLE, TAKE- OFF (as an airplane) vs. TAKE-OFF (in general), and others in Israeli Sign Language (ISL). For example, metaphorical expressions are not interpretable with EAT, but they are with NIBBLE (see (1) & (2)). Meir claims that the difference comes from the DMC, which dictates that metaphorical mappings have to be structurally preserved in iconic mappings. The 'consumption' meaning in the metaphorical expressions is not reflected in the iconic mapping of EAT, where putting food into one's mouth is represented (<Figures 1 & 2>). However, NIBBLE can be used metaphorically since the 'consumption' meaning occurs in the iconic mapping where the teeth consuming food is represented (<Figures 3 & 4>).

Revision of DMC: In this paper, based on the metaphorical examples of Korean Sign Language (KSL), it will be claimed that the DMC has to be revised to include some exceptional examples. The metaphor UNDERSTANDING IS DIGESTING in KSL demonstrates that pragmatic inferences of iconic signs can be metaphorically conceptualized even when the highlighted element in metaphorical mappings is not involved in iconic mappings. DIGEST in KSL shows how our stomach moves to break down food (<Figure 5>). The sign does not show how the stomach absorbs nutrients. However, the cognitive process UNDERSTANDING highlights how we infer the main idea after analyzing information. The iconic and metaphorical mappings do not match (<Figure 6>), but DIGEST is still used in metaphorical expressions. Therefore, the DMC should include pragmatic inferences.

The metaphor MORALITY IS PHYSICAL CLEANLINESS in KSL shows that iconic mappings have to be consistent all the way through metaphorical conceptualization, even when pragmatic inferences are involved. CLEAN and DIRTY in KSL use one's body parts: face for CLEAN and nose for DIRTY (<Figures 7 & 8>). As in (2) and (3), metaphorical expressions with CLEAN and DIRTY are possible. However, moral restoration with cleansing, as observed in *cleaning up corruption* or *washing away sin*, is not interpretable with WASH-HANDS or WASH-BODY in KSL (KSL does not have a stand-alone WASH sign). In spoken languages, when cleanliness is metaphorically conceptualized as moral purity, cleansing can be conceptualized as moral restoration through pragmatic inferencing, but in KSL, it is not possible. The reason is that WASH-HANDS and WASH-BODY use different body parts, hands and upper body, from CLEAN and DIRTY, violating consistency in iconicity.

Conclusion: It is concluded that the DMC should include pragmatic inferences and, in pragmatic inferencing, iconic consistency must be preserved. Meir and Cohen (2018) identify the source of the DMC as inhibition: It is difficult to inhibit what is iconically expressed. Thus, we can say that perspective shift, as well as inhibition, is not easy with iconic signs.

(1)   a. His wife's illness **eats at** him.
         b. The car **ate up** all my savings.
(2)   The acid **nibbled at** the iron key.

<Figure 1: EAT (ISL)>   <Figure 2: Double Mappings of EAT>   (Meir 2010, 879)

<Figure 3: NIBBLE (ISL)>   <Figure 4: Double Mappings of NIBBLE>   (Meir 2010, 880)

<Figure 5: DIGEST (KSL)>   <Figure 6: Double Mappings of DIGEST>

<Figure 7: CLEAN (KSL)>   <Figure 8: DIRTY (KSL)>

(3) MY FRIEND STRAIGHT CLEAN 'My friend is honest and morally clean'
(4) BEHAVIOR DIRTY HABIT CUT 'He stopped immoral behaviors'

References
Meir, I. (2010). Iconicity and metaphor: Constraints on metaphorical extension of iconic forms. *Language* 86.4: 865-896.
Meir, I & A. Cohen. (2018). Metaphor in sign languages. *Frontiers in Psychology* 9: 1025.

Conceptual metaphor and double mapping in Chinese Sign Language

Charmhun Jo

(Southwest University, China)

Conceptual metaphor is a linguistic mechanism that utilizes concrete experiences to understand and express abstract concepts. This is considered as a universal linguistic phenomenon that is widely observed not only in spoken languages but also in sign languages.

The situation, however, differs for conceptual metaphors derived from sign languages, as a result of the strong association between metaphor and iconicity (Wilcox 2000; Taub 2001). In other words, sign languages are known for their rich iconicity, which means that when signs are extended metaphorically, they undergo double mappings that involve both metaphorical and iconically-related transfers.

This study thus aims to investigate the applicability of the double mapping model for Chinese Sign Language (CSL), which has received less attention thus far. To determine if the double mapping model, which connects metaphorical and iconic dimensions, is applicable within the linguistic context of CSL, this study examines the metaphorical mapping mechanism of the abstract verb FORGET in CSL.

The results indicate that the sign for FORGET in CSL exhibits a layered structure of metaphorical and iconic elements, demonstrating that the double mapping model is valid for CSL.

Figure 1. CSL abstract verb FORGET (*Standard CSL Dictionary*, p. 886)

Table 1. Double mapping for CSL abstract verb FORGET

	Iconic mapping	Metaphorical mapping
ARTICULATORS	**Source**	**TARGET**
All fingers gathered	An object	A thought
Forehead	Head	Locus of thought
All fingers gathered touch forehead	Object located in head	Thought generated by originator
All fingers gathered move beside head	Relocate an object out of head	Difficulty in recalling thought
All fingers gathered change to all open-extended fingers beside head	Release an object out of head	Failure of recalling thought

The dual-layered mapping not only enhances the expressiveness of the sign but also highlights the cognitive interplay between concrete experience and abstract concept. The findings of this study emphasize the significance of considering both metaphorical and iconic dimensions when analyzing conceptual metaphor in sign languages and confirm that the double mapping model provides a robust framework for examining the intricate connection between metaphor and iconicity.

Selected References
Standard CSL Dictionary. 2019. Beijing: Huaxia Publishing House.
Taub, Sarah F. 2001. *Language from the Body: Iconicity and Conceptual Metaphor in American Sign Language*. Cambridge: Cambridge University Press.
Wilcox, Phyllis Perrin.2000. *Metaphors in American Sign Language*. Washington, DC: Gallaudet University Press.

Grammaticalization of Adverbial Clauses in Spoken and Sign Languages

Seonhye Lee

(Pukyong National University, Korea)

The relationship between social change and language evolution is a fundamental aspect of human society. Language reflects societal transformations and acts as a catalyst for cultural change. In the field of linguistics, grammaticalization stands out as a prominent phenomenon, particularly in lexical change. It encompasses a variety of definitions and processes that drive language change and transformation.

According to Antoine Meillet (1912), who coined the term, grammaticalization involves the transition of lexical items into grammatical elements. This process entails the evolution of less grammatical components into more grammatical ones, which adds grammatical characteristics to lexical items. Another notable definition of grammaticalization is provided by Pfau and Steinbach (2006), who describe it as a process in which lexical elements, such as nouns or verbs, evolve into free grammatical forms, such as pronouns or adverbs, and subsequently develop into dependent grammatical forms, such as affixes for agreement or tense. They refer to this progression as the "grammaticalization path."

Grammaticalization presents distinct manifestations across various spoken and sign languages, including German Sign Language (DGS), American Sign Language (ASL), and Italian Sign Language (LIS), among others. This is due to the distinct modalities of these communication forms.

This study primarily aims to explore the concept of grammaticalization by analyzing its definitions and expressions in both spoken and sign languages. By examining the universal characteristics of grammaticalization processes, we can gain deeper insights into the mechanisms underlying language evolution and change.

Subordination is generally defined as a relationship in which a clause serves a function similar to that of a word in a simple clause. When a clause functions as the subject or object of another clause, it is referred to as a complement clause. If a clause modifies a noun, akin to an adjective, it is identified as a relative clause. When a clause occupies an adjunct position, modifying another clause, it is classified as an adverbial clause. Among these noun clauses, adjective clauses, and adverbial clauses, this study focuses on adverbial clauses and aims to investigate the grammaticalization processes occurring within them.

More specifically, this research investigates the grammaticalization processes of subordinate conjunctions in subordinate clauses of both sign and spoken languages, focusing on the development and use of adverbial clauses, particularly temporal, causal, factual conditional, and counterfactual conditional clauses.

The study aims to elucidate the mechanisms supporting language change and evolution by examining the fundamental structures of complex sentences. Through a detailed analysis of various sign and spoken languages, it demonstrates how grammaticalization in sign languages differs from that in spoken languages, considering the unique modalities and cognitive processes involved. It explores the extent to which sign languages employ similar or different pathways to achieve grammatical complexity, particularly in the formation of adverbial clauses.

By examining specific examples from various sign languages and comparing them to spoken languages, the study provides valuable insights into how different languages encode temporal, causal, and conditional relationships. This research addresses key questions regarding the universality of grammaticalization processes and their implications for our understanding of language evolution.

Our findings contribute to the broader field of linguistics by providing insights into the interaction between language modality and grammatical development. By investigating the evolution

of complex sentence structures, we gain a deeper understanding of the dynamic nature of language and its capacity for adaptation and change. This study not only enhances our comprehension of grammaticalization in diverse linguistic contexts but also underscores the importance of cross-modal research in capturing the full spectrum of human linguistic capability.

Keywords: grammaticalization, sign languages, adverbial clauses, language evolution, complex sentences, language variation

Lexical Development in Young Deaf Children:
A Comparative Longitudinal Study of Korean Sign Language Acquisition

Se-Eun Jhang
(Korea Maritime and Ocean University, Korea)
Sook Ki Lee
(Korea Nazarene University, Korea)
Byung-Cheon Youn
(Chosun University, Korea)
Haewon Jeon
(Korea Maritime and Ocean University, Korea)
Seonhye Lee
(Pukyong National University, Korea)

Abstract:

This study examines the lexical development of two young deaf children, with a particular focus on their acquisition of Korean Sign Language (KSL) between the ages of 6 and 36 months. Employing a detailed longitudinal methodology, it not only observes their developmental milestones but also situates their progress within a comparative framework involving American Sign Language (ASL), Japanese Sign Language (JSL), and others. This approach employs a comparative methodology, linking the children's developmental milestones in KSL to broader phenomena observed across various sign languages. This approach provides a more comprehensive understanding of the universal and specific aspects of sign language acquisition.

Key Words: Korean Sign Language, language acquisition, lexical development, young deaf children, early sign development, comparative sign linguistics

Methodology:

The study examines the acquisition of Korean Sign Language (KSL) in two deaf girls born to deaf parents and native KSL users. We observed the young deaf twins on a daily basis in their natural environment, engaging in activities such as reading picture books, viewing KSL children's video books, and participating in various play-based learning activities. These sessions, which lasted between 10 and 15 minutes each, were essential for collecting data on the expressive language used by the twins. The length of the sessions was adjusted based on the children's engagement levels. The observations focus on key sign parameters, including handshape, movement, location, and palm orientation, from the initial simple gestures to more complex linguistic structures. The children's initial signs, misarticulations, and progress towards more standardized sign forms are documented and compared with equivalent stages in American Sign Language (ASL), Japanese Sign Language (JSL), and other languages based on existing literature (e.g., Anderson and Reilly 2002; Berger et al. 2024; Caselli et al. 2020; Marentette and Mayberry 2000; Mayberry and Squires 2005; Woolfe 2010).

Comparative Analysis:

At six months, the initial findings indicate the emergence of basic signs, such as "milk" and "eat," which are characterized by simplified handshapes in proximity to major body parts, including the chest and mouth. As the children progress, their signing becomes increasingly complex, as evidenced by the emergence of signs such as "bath" and "mom" by the twelfth month. By 24 months, the children's signing has become significantly more complex, with signs such as "book" and "cat" indicating lexical

enrichment. It is notable that by 36 months, the children exhibit two-handed signing patterns that are similar to those observed in adults, which represents a significant milestone in the acquisition of KSL. The observation of these early signs, which were exposed to their deaf native KSL parents by 36 months, was conducted to conduct a comparative analysis with other sign languages, such as ASL and JSL. This analysis reveals both shared and distinct developmental patterns. For example, the simplification observed in early KSL handshapes mirrors trends seen in ASL and JSL, where infants initially adopt a limited set of phonological parameters. This study further examines how the phonological constraints of early sign production are adapted by children in different linguistic environments, informed by existing research on manual babbling and first sign errors in various sign languages.

Results:

The results of this study indicate that daily observations in natural settings can provide insights into the role of environmental factors, such as interaction with native signers and exposure to sign language media. These influences are crucial for understanding the varied trajectories of sign language acquisition among deaf children. The children's evolving ability to mimic complex signs indicates both cognitive and motor skill development, contextualized within a broader comparative analysis.

Conclusion:

The findings contribute to a more profound comprehension of the processes underlying sign language acquisition and highlight the importance of a supportive linguistic environment. They underscore the necessity for early and consistent exposure to sign language for optimal lexical development in deaf children. Moreover, this study enhances our comprehension of how developmental patterns in sign language acquisition might vary yet also align across different cultural and linguistic contexts.

Selected References

Anderson, Diane and Judy Reilly. 2002. "The MacArthur Communicative Development Inventory: Normative Data for American Sign Language," *Journal of Deaf Studies and Deaf Education* 7(2), 83-106.

Berger, Lauren, Jennie E. Pyers, Amy M. Lieberman and Naomi K. Caselli. 2024. "Parent American Sign Language Skills Correlate with Child-but Not Toddler-ASL Vocabulary Size." *Language Acquisition* 31(2), 85-99. doi:10.1080/10489223.2023.2178312

Caselli, Naomi K., Amy M. Lieberman and Jennie E. Pyers. 2020. "The ASL-CDI 2.0: An Updated, Normed Adaptation of the MacArthur Bates Communicative Development Inventory for American Sign Language." *Behavior Research Methods* 52(5), 2071-2084. doi:10.3758/s13428-020-01376-6

Marentette, Paula, F. and Rachel, I. Mayberry. 2000. "Principles for an Emerging Phonological System: A Case Study of Early ASL Acquisition," in Charlene Chamberlain, Jill P. Morford, & Rachel I. Mayberry, eds, *Language Acquisition by Eye,* 71-90. Mahway, NJ: Lawrence Erlbaum Associates.

Mayberry, Rachel. I. and B. Squires. 2005. "Sign Language Acquisition," in Keith Brown, ed., *Encyclopedia of Language and Linguistics*, 291-296, Elsevier.

Woolfe, Tyron, Rosalind Herman, Penny Roy and Bencie Woll. 2010. "Early Vocabulary Development in Deaf Native Signers: a British Sign Language Adaptation of the Communicative Development Inventories." *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 51(3), 322-31. doi:10.1111/j.1469-7610.2009.02151.x

Impersonally Married: A Comparative Genre Analysis of Reddit and ChatGPT-4 Spouse Appreciation Posts

Erika Kristine Arcenal, Licca Pauleen Capistrano, Marielle Jessie De Guzman,
Micaela Isabel Forrosuelo, Janeson Miranda
(De La Salle University Integrated School, The Philippines)

The prevalence of generative Artificial Intelligence (AI) software has been a helping tool to its users in various fields and contexts, and with this rapid advancement, a recurring topic of inquiry is its ability to mimic human-written ouputs. One AI model that has been receiving extensive attention is the language-processing chatbot, ChatGPT, recording 100 million weekly users a year within its release (Malik, 2023).

While ChatGPT's projected ethical concerns, validity, and potential in education, programming, and business have been explored, there is an evident scarcity in studies concerned with the AI's ability to imitate human language. In academic contexts, several researchers have found ChatGPT's outputs coherent and concise yet lacking factual accuracy (Herbold et al., 2023; Ariyaratne et al. 2023; Johansson, 2023), while an even greater gap is identified in texts highlighting storytelling and personal experience. Hence, in this study, we explore the current capabilities of ChatGPT-4 at mimicking human appreciatory language by comparing its prompt-based generated texts to the appreciative posts of married individuals on Reddit, specifically subreddit r/Marriage.

One hundred fifty posts, evenly split between curated Reddit Spouse Appreciation Posts and ChatGPT-4 Generated Posts, were analyzed using Swales' (1990) moves analysis as genres and later compared together. Furthermore, four categories of obligatory moves were established to better distinguish between the texts as observed in (1)

(1) Obligatory Classification of Moves

Rate of Move Appearance (RMA)	Classification
0.00 - 0.24	Not Obligatory
0.25 - 0.49	Obligatory but Infrequent
0.50 - 0.74	Obligatory but Supplementary
0.75 - 0.99	Obligatory
>=1.00	Clearly Obligatory

Findings revealed that ChatGPT-4 imitated 11 out of the 13 moves present in the Reddit posts as seen in (2).

(2) Moves Identified in Reddit Posts and ChatGPT-4 Generated Text

Moves	Reddit	RMA	ChatGPT	RMA
Abridge and Edits (AAE)	✓	0.31	✓	0.00
Acknowledgment of Readers (AOR)	✓	0.17	✓	2.21
Appreciatory Sentiment (APS)	✓	1.16	✓	1.96
Cheers and Salutations (CAS)			✓	0.61
Contextual Narrative (CON)	✓	2.48	✓	2.37
Description of Original Poster (DOP)	✓	0.36	✓	0.04
Description of Subject (DOS)	✓	0.28	✓	0.01

Moves	Reddit	RMA	ChatGPT	RMA
Editable Bracket (EDB)			✔	0.33
Inclination to Post (ITP)	✔	0.25	✔	1.28
Mention of Future (MOF)	✔	0.11	✔	0.25
Point of Reflection (POR)	✔	0.93	✔	1.69
Relationship Background (REB)	✔	0.52	✔	0.08
Section Indicator (SEI)	✔	0.07	✔	1.20
Support for Subreddit (SFS)	✔	0.03		
Title (TIE)	✔	1.00	✔	1.00

In (2), ChatGPT-4 was also observed to have more clearly obligatory moves and was more consistent in following a certain sequence than texts written by a human being, as seen in (3).

(3) Sequence of Moves

	Move Sequence								
Reddit	TIE	CON	APS	*Varied*					
ChatGPT-4	SEI	TIE	AOR	ITP	CON	POR	APS	CAS	EDB

Additionally, ChatGPT-4 generated texts had a higher average word count of 359 over Reddit's 216 and gave an impersonal impression due to its lack of the following moves: Description of Subject, Description of Original Poster, and Relationship Background.

With these findings, we recommend further linguistic studies on Reddit posts and other non-academic social media forms as a genre and the continuous use of genre analysis to circumscribe, more precisely, the AI imitative capabilities of human natural language and its implications to future applications.

Selected References

Ariyaratne, S., Iyengar, K. P., Nischal, N., Babu, N. C., & Botchu, R. (2023). A comparison of ChatGPT-generated articles with human-written articles. Skeletal Radiology, 52(9), 1755– 1758. https://doi.org/10.1007/s00256-023-04340-5

Herbold, S., Hautli-Janisz, A., Heuer, U., Kikteva, Z., & Trautsch, A. (2023). A large-scale comparison of human-written versus ChatGPT-generated essays. Scientific Reports, 13(1). https://doi.org/10.1038/s41598-023-45644-9

Johansson, I. (2023). A Tale of Two Texts, a Robot, and Authorship : A Comparison Between a Human-Written and a ChatGPT-Generated Text. DIVA.

https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1767661&dswid=3136 Malik, A. (2023, November 6). OpenAI's ChatGPT now has 100 million weekly active users. TechCrunch. https://techcrunch.com/2023/11/06/openais-chatgpt-now-has-100-million-weekly-active-users/?guccounter=1

Swales, J. M. (1990). Genre analysis: English in academic and research settings (1st ed.). Cambridge University Press.

Applying AI in Teaching Academic Writing: A Pilot Study

Evgeniya Aleshinskaya

(National Research Nuclear University MEPhI, Russia)

Scientific discourse has been shown to largely reflect the flow of intelligent processes in a scientific text (Osipov Devyatkin Kuznetsova & Shvets 2019). A scientific text comprises a series of cognitive communicative actions, which correspond to mental actions in the author's cognitive activity. This feature of scientific discourse is essential for language research and teaching, as it allows applying artificial intelligence as a means of optimizing and facilitating the process of teaching academic writing. The current study suggests using AI in assessing students' academic writing assignments. It proposes to employ an intelligent system of excerpts from authentic texts in different academic genres, such as abstracts, introductions or conclusions, that would enable the extraction and systematization of necessary data and in this way provide automatic assistance in the assessment of students' academic writing.

This investigation presents the results of a pilot study into the cognitive model of a research article involving mental operations fulfilled by native speaking authors in the introduction section. Theoretically and methodologically, the study is based on the create-a-research-space (CARS) model suggested by Swales (1981). In accordance with this model, research article introductions follow a particular organizational pattern that reflects the basic mental actions materialized through a series of moves serving to transmit the author's intention in a written scientific text and steps having specific communicative functions within a move realized by particular linguistic features. The model comprises three moves: (1) establishing a research territory; (2) establishing a research niche; (3) occupying the niche. The basic moves include several steps, three of which are obligatory (literature review in Move 1, indication of a gap in the literature in Move 2, and description of the purpose or nature of the research in Move 3), and the rest of which are optional depending on the scientific area. The units of analysis to identify steps within the moves of the model are individual sentences. The identification of steps is based on determining linguistic features, which indicate mental operations involved in the generation of a research article introduction. The steps belonging to each of the three basic moves are illustrated with examples of linguistic features representing mental operations in (1).

(1) Intelligent structure of a research article introduction

Move	Step	Verbal representation
Move 1: Establishing a research territory	Opening generalization	Many, often, increasing
	Claiming centrality	Essential, important, central
	Claiming relevance	Recent(ly), current(ly)
	Claiming novelty	New, for the first time
	Providing background information	Represent, consist of, involve
	Providing evidence from previous research	Report, explain, example, previous work
	Defining key terms	Is defined, is characterized as
Move 2: Establishing a research niche	Extending previous knowledge	Proved efficient for, we demonstrate its application to
	Presenting positive justification	Motivated, helpful, prospect
	Counterclaiming	Although, however, challenge

	Question raising	Raise several questions
	Raising a need	Needed, necessary, require
	Indicating a gap	No study has investigated, few studies have dealt with
	Presenting a hypothesis	Hypothesized, if, then
	Outlining the research aim(s)	The purpose is, is aimed at
Move 3: Occupying the niche	Stating the nature of the research	This paper examines, we focus on, we explore
	Announcing the main outcome(s)	We demonstrate, we show, we find, these results indicate that
	Stressing the value	The first main contribution, a clear advantage, benefit
	Stressing the applicability	Can perform well in, can be applied in, wide applicability
	Outlining the structure of the paper	The rest of the article, the following structure, then, next

Regarding the order of the intelligent activities within the structure of article introductions, the qualitative analysis of the selected articles allowed identifying the main sequences of mental operations. Of interest is an observation that the sequence of mental operations is not always linear, when move 1 is followed by move 2 and then move 3. The nature of the research can be stated several times in the article introduction, taking place in a series of cycles. One of such sequences is presented in (2).

(2) opening generalization (sentences 1-2) > question raising (sentences 3-8) > stating the nature of the research (sentence 9) > defining a key term (sentence 10) > providing background information (sentences 11-13) > claiming centrality (sentences 14-16) > providing evidence from previous research (sentences 17-19) > counterclaiming (sentence 20) > stating the nature of the research (sentence 21) > announcing the main outcomes (sentences 22- 37) > stressing the value (sentences 38-42).

The results of the pilot study demonstrate the possibility of constructing a cognitive model of a scientific text, which largely reflects its intelligent structure. They allow for the identification of intelligent operations as elements of academic writing aimed at further assessment of students' texts based on such sets of evaluation criteria as task fulfilment, style compliance, organization and coherence. Further automation of the intelligent structure identification will significantly relieve the burden on the teacher and facilitate the effort- and time-consuming task of checking numerous scientific texts while teaching academic writing.

References

Osipov, Gennady S., Devyatkin, Dmitry A., Kuznetsova, Yulia M., Shvets, Alexander V. 2019. "The Possibilities of Intelligent Analysis of Scientific Texts by Construction of Their Cognitive Models," *Scientific and Technical Information Processing 46*(5), 337-344.

Swales, John M. 1981. *Aspects of Article Introductions*. Birmingham: University of Aston.

Grammar Plays a Role in Human-Computer Interaction, Huh?

Matthew Galbraith, Preeti Kumari
(Southwest University, China)
Martina Wiltschko
(ICREA, Universitat Pompeu Fabra, Spain)

Introduction. One key goal in conversational design for human-computer interaction (HCI) is to mimic the naturalness of human-human interaction (HHI), often by formalizing aspects of language with syntactic and semantic rules. However, natural language also involves knowing how to **use** language, and all languages have means to regulate communicative interaction. Following Wiltschko (2021), we distinguish between propositional language (p-language) and interactional language (i-language). For example, in "*The alarm is set for 7 AM, huh?*", the propositional content is "*The alarm is set for 7 AM…*", while the interactional "*huh?*" has two functions in HHI: i) it indicates that the speaker assumes some shared knowledge with the addressee, thus regulating the management of common ground (CG); ii) it signals that the speaker requests a response, thus managing turn-taking (TT). In addition to elements that can serve both functions (like the sentence-final "*huh?*"), there are also elements that are dedicated to just one of the two functions (CG or TT). The question we are exploring is whether human users perceive the use of i-language in HCI as natural. And in particular, whether there is a difference between i-language used for CG management and i-language used for TT management. Answering these questions may shed light on the question of how human users perceive computers as interactants, a question that is currently debated (Clark & Fischer 2022). Our findings suggest that there is no blanket ban on the naturalness of i-language in HCI, rather, only i-language required for CG management is considered unnatural in HCI, whereas i-language used to manage TT is considered natural.

Methodology. An acceptability judgment experiment was conducted through a survey wherein 200 native English speakers assessed the naturalness of HCI employing "*huh?*" in two contexts: 1) Other Initiated Repair (OIR), a case of TT-management and 2) Request for Confirmation (RFC), a case of CG management. Participants were given five dialogues containing OIR and RFC each. For OIR, participants evaluated two hypothetical HCI scenarios - one, an interrogative formulated with p- language only (OIRp) and the other using i-language "*huh?*" (OIRi), see Table 1 (where H = human user and C = computer).

	OIRp	OIRi
H:	Set an alarm for 8 o'clock at night.	Set an alarm for 8 o'clock at night.
C:	OK, I can do that for you.	OK, I can do that for you.
H:	**What was that?**	**Huh?**

Table 1. OIR Dialogue

In the RFC scenario, subjects rated the naturalness of four target utterances - a standard interrogative (*Interrogative* in Table 2), a confirmational question comprising of a declarative sentence followed by i-language in the form of the tag question particle "*huh?*" (*RFCi* in Table 2), a declarative (*Declarative* in Table 2), and a confirmational question comprising of a declarative followed by the p-language tag question "*Is that true?*" (*RFCp* in Table 2). We used these contexts to see how human users perceive the difference between a canonical interrogative and an RFC question (columns 2 & 3) as well as determining whether there is a difference in perception between RFCs with and without i-language (columns 2 & 4). Finally, we also tested the naturalness of an RFC using a declarative, which is possible in HHI.

	Interrogative	RFCi (i-language)	Declarative	RFCp (p-language)
H to C	Do I have an alarm set for today?	I have an alarm set for today, huh?	I believe that I have an alarm set for today.	I believe that I have an alarm set for today, is that true?

Table 2. RFC Dialogue

Results. To analyze the collected data, ordinal regression analyses were conducted on the Likert scale responses for each context. This method enabled us to quantify user preferences and assess the statistical significance of variations between response types. As for OIRs, we observed a tendency towards p-language, but we did not find a statistically significant result for *"huh?"* being classified as natural or unnatural. For RFCs, interrogatives were rated significantly natural, whereas RFCi were rated significantly unnatural, with the overall order of naturalness being as follows: *Interrogative > RFCp > Declarative > RFCi*

Discussion. In OIR dialogues, differences emerged in user preferences between responses with and without "*huh?*". Responses lacking "*huh?*" showed higher naturalness, but the inconsistent statistical significance suggests further investigation is needed to reach a definitive conclusion. Specifically, we need to control for the possibility that *"huh?"* is used by the human user to express confusion without necessarily considering the computer as an interlocutor. If so, i-language might still generally be considered unnatural, even for TT-management.

As for RFCs, the analysis showed a significant preference for direct questions over other types. Confirmations, statements, and questions with "*huh?*" were less preferred, with the latter being the least favored. This suggests that i-language which regulates CG is considered unnatural in HCI. It remains to be determined whether elements of p-language that are used to access the CG are considered equally unnatural or whether the observed unnaturalness is restricted to i-language.

Conclusion. The results of our pilot study indicate that one of the core attributes of natural language used in HHI is affected in HCI. This is not surprising as i-language is highly context-specific, and its use depends in part on who knows what in the ongoing conversation. We submit that i-language can be used as a window into the way human users view computers as interactants without having to rely on post-hoc questionnaires (Bartneck et al 2009).

References.

Bartneck, C., Kulić, D., Croft, E., & Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of Robots. *International Journal of Social Robotics*, *1*(1), 71–81. https://doi.org/10.1007/s12369-008-0001-3

Clark HH, & Fischer K. (2022). Social robots as depictions of social agents. Behav Brain Sci. Mar 28;46:e21. doi: 10.1017/S0140525X22000668. PMID: 35343422.

Dingemanse, M., Torreira, F., & Enfield, N. J. (2013). Is "Huh?" a Universal Word? Conversational Infrastructure and the Convergent Evolution of Linguistic Items. *PLoS ONE*, *8*(11), e78273. https://doi.org/10.1371/journal.pone.0078273

Wiltschko, M. (2021). The Grammar of Interactional Language (1st ed.). Cambridge University Press. https://doi.org/10.1017/9781108693707

Wiltschko, M., & Heim, J. (2016). The Syntax of Confirmationals. A neo-performative analysis. In G. Kaltenböck, E. Keizer, & A. Lohmann (Eds.), *Outside the clause: Form and function of extra-clausal constituents* (pp. 303–340). John Benjamins.

Biography of Invited Speakers

John Beavers (The University of Texas at Austin, USA)

John Beavers is the Chair of the Department of Linguistics and the Robert D. King Centennial Professor of Liberal Arts at The University of Texas at Austin. He is also the Editor of *Language,* the flagship journal of the Linguistic Society of America. He received his PhD from Stanford University in 2006 and subsequently taught at Georgetown University from 2006-2007. Prof. Beavers' research explores the syntax and semantics of the world's languages, with a primary focus on the nature of word meanings, how word meanings are decomposed into more basic semantic pieces, and how a word's meaning determines how it is used grammatically. This work has involved in-depth studies on a variety of languages, including English, Indonesian, Japanese, Spanish, Colloquial Sinhala, and Kinyarwanda, as well as cross-linguistic studies. He is co-author of *The Roots of Verbal Meaning*, published in 2020 by Oxford University Press, and he has published articles in many top journals in linguistics, including *Language*, *Natural Language and Linguistic Theory*, *Journal of Linguistics*, *Journal of Semantics*, and *Glossa*.

Thomas Hoffmann (KU Eichsättt-Ingolstadt, Germany)

Thomas Hoffmann is "Furong Scholar" Distinguished Chair Professor of Hunan Normal University and Full Professor and Chair of English Language and Linguistics at the Catholic University of Eichstätt-Ingolstadt. His main research interests are usage-based Construction Grammar, language variation and change, and multimodal communication. He has published widely in international journals such as *Cognitive Linguistics, Journal of English Linguistics, English World- Wide* and *Corpus Linguistics and Linguistic Theory*. His 2011 monograph *Preposition Placement in English* as well as his 2019 book *Comparing English Comparatives* were both published by Cambridge University Press. He has recently published a textbook on *Construction Grammar: The Structure of English* for the Cambridge Textbooks in Linguistics series in 2022 and co-authored *Copilots for Linguistics* for the Cambridge Elements in Construction Grammar series in 2023. He is also Area Editor for *Syntax of the Linguistics Vanguard* and Editor-in-Chief of the Open Access journal *Constructions*. Currently, he is working on psychological as well as cognitive aspects of linguistic creativity.

Benjamin Bruening (The University of Delaware, USA)

Benjamin Bruening is Professor of Linguistics at the University of Delaware. He received his PhD from MIT in 2001. His research seeks to understand the nature of syntax and its interfaces with semantics and morphology. Prof. Bruening's research has involved in-depth studies on various phenomena, with a primary focus on the syntax and morphology of Algonquian languages, ditransitive constructions in various languages, subject-auxiliary inversion in English, and the connection between morphology and phonology in Semitic languages. His other interests include neurolinguistics and sentence processing. His current research projects include argument structure; double object constructions; experimental evidence bearing on the issue of reconstruction in *wh*-movement; and reflexive clitics in Slavic languages. He has published articles in many influential journals and edited volumes, including *Linguistic Inquiry, Natural Language and Linguistic Theory* and *The Wiley Blackwell Companion to Syntax*.

Bernd Kortmann (The University of Freiburg, Germany)

Bernd Kortmann is Full Professor of English Language and Linguistics at the University of Freiburg, Germany. His main research interest over the last decade has been the grammar of non-standard varieties of English around the world, especially from a typological perspective. He has published on semantics, semantic change, grammaticalization, typology, language complexity, history of linguistics, and English grammar. Prof. Kortmann serves as co-editor of two book series, both published by De Gruyter Mouton: *Topics in English Linguistics* and *Dialects of English* (formerly Edinburgh University Press). He also serves as one of three editors of the journal *English Language and Linguistics* (Cambridge University Press). Apart from his duties as a series and journal editor, he is a member of the editorial boards of the journals *English Today* and *Transactions of the Philological Society.* Since 2012 he has been a member of the selection panel of the Alexandervon Humboldt Foundation Fellowships. From April 2008 until September 2009 he was a Senior
Research Fellow in the School of Language and Literature of the Freiburg Institute of Advanced Studies (FRIAS). In December 2013 he was appointed Director of the FRIAS for the Humanities and Social Sciences, and as of October 2015 Speaker of the Board of Directors (Executive Director) of FRIAS. Prof. Kortmann holds an Honorary Doctorate from the University of Eastern Finland and is an elected member of the Academia Europaea.

The 2024 Seoul International Conference on Linguistics (SICOL-2024)

Reenvisioning linguistic creativity:
Language and AI in interaction