

Content

Anaphora in L2 Teaching and Research Darcy Sperlich	1
Variations underlying Apparent English <i>Tough</i> Constructions Shigenori Wakabayashi, Yuki Takahashi	29
Japanese Children's Extensive English Learning Atsushi Miura, Natsumi Shibata, Shiro Ojima	77
Positive Cross-Linguistic Influence in L2 and Heritage Chinese Acquisition Shanshan Yan	117
L1 Speaker Modification in L1-L2 Conversation Tomoharu Yanagimachi	153
Time-Related Factors in Korean Learners' Vietnamese Acquisition Trang Phan, Lan Chu	179
Automatic Grammar Instruction Jiun-Shiung Wu, Hsin-Hui Chen, Hui-Ru Yang	205
English-Medium Instruction in Vegetable Disease Course Ying Ying Wong, Yuan-Min Shen, Cheng-Chun Huang	235

Taiwan Journal of Linguistics Book Series

ISBN 978-626-7239-67-4



Current Issues in L2 Teaching and Research

Current Issues in L2 Teaching and Research

Istano Kong, Yu-Hsin Huang, Miao-Chin Chiu
Kai-Lin Wu, and One-Soon Her

Current Issues in L2 Teaching and Research

Edited by
Stano Kong
Yu-Hsin Huang
Miao-Chin Chiu
Kai-Lin Wu
One-Soon Her

Contents

PREFACE	i
LIST OF EDITORS	vi
LIST OF CONTRIBUTORS	vii
 ANAPHORA IN L2 TEACHING AND RESEARCH: THEORY AND PRACTICALITIES Darcy Sperlich	 1
 VARIATIONS IN SYNTACTIC STRUCTURE UN- DERLYING APPARENT ENGLISH <i>TOUGH</i> CON- STRUCTIONS Shigenori Wakabayashi, Yuki Takahashi	 29
 IMPACTS OF EXTENSIVE ENGLISH EXPOSURE ON JAPANESE CHILDREN: LANGUAGE SKILLS AND BRAIN ACTIVITY Atsushi Miura, Natsumi Shibata, Shiro Ojima	 77
 A FOLLOW-UP STUDY ON THE POSITIVE CROSS-LINGUISTIC INFLUENCE OF SEN- TENCE-FINAL PARTICLES ON ACQUISITION: LEARNER, STRUCTURE, AND LANGUAGE PROFICIENCY Shanshan Yan	 117

Contents

FOR WHOM IS L1 SPEAKER MODIFICATION IN L1-L2 CONVERSATION?: RECONCEPTUALIZING FOREIGNER TALK FROM THE PERSPECTIVE OF INTERACTION ANALYSIS **153**

Tomoharu Yanagimachi

PROBING THE IMPACT OF TIME-RELATED FACTORS ON KOREAN LEARNERS' PROFICIENCY IN ACQUIRING THE VIETNAMESE LANGUAGE **179**

Trang Phan, Lan Chu

INTRODUCING AGIBIS (AUTOMATIC GRAMMAR INSTRUCTION BASED ON INPUT SENTENCES): A PILOT STUDY OF A CHATBOT FOR AI-ASSISTED L2 CHINESE LEARNING **205**

Jiun-Shiung Wu, Hsin-Hui Chen, Hui-Ru Yang

TEACHING AND LEARNING IN AN ENGLISH-MEDIUM INSTRUCTION COURSE IN UNIVERSITY: A TWO-YEAR EXPERIENCE OF THE VEGETABLE DISEASES COURSE FOR GENERAL EDUCATION **235**

Ying Ying Wong, Yuan-Min Shen, Cheng-Chun Huang

PREFACE

Welcome to the inaugural volume of *Current Issues in L2 Teaching and Research*, a collection arising from the first-ever international conference in Taiwan dedicated to exploring the dynamic world of second language acquisition and teaching. A second language (L2), in its broad sense, refers to any language a person uses or studies besides the first language (L1). The First Tunghai International Conference on Second Language Research, organized by the Department of Foreign Languages and Literature, the International Graduate Program of Teaching Chinese as a Second Language, and the English Center of Tunghai University, was held on October 21, 2023.

An international conference specifically on L2 research and teaching not only facilitates constructive interaction among teachers and researchers of various L2s but can also create virtuous synergy for all stakeholders. To enhance the quality and influence of the conference, the plan is for the organizers to select papers from the presentations and also issue an open call for submissions after the conference and publish a book after double-blind reviews. The program committee members serve as the editors. The conference is held every two years, with the publication of an edited book in the year following the conference.

The current volume reflects the multifaceted nature of L2 teaching and learning by bringing together diverse perspectives and going beyond the boundaries of individual languages. In the first chapter, Darcy Sperlich explores anaphora in language acquisition, focusing on Chinese and English, and argues for a multifaceted approach to examine the challenges Chinese learners face with anaphora and politeness in English. It concludes with teaching strategies and areas for future research. The second chapter, by Shigenori Wakabayashi and Yuki Takahashi, explores how Japanese learners interpret English *tough* constructions (TCs). It

finds more variation in learner grammar than in past research. Learners showed no clear link between TC interpretation and related grammar points. The results suggest that learners may use different strategies for TCs and that Universal Grammar functions differently in L2 acquisition, potentially influenced by the learner's existing grammar.

Atsushi Miura and Shiro Ojima's study is concerned with the impact of early L2 learning. Specifically, they investigate how massive exposure to an L2 affects children's L2 abilities and potentially even their L1 skills. Their experiments showed children can excel in spoken L2 with minimal formal instruction, but extensive L2 exposure might influence academic L1 abilities and brain activity. In the fourth chapter, Shanshan Yan analyzes positive cross-linguistic influence (CLI) in learners acquiring Mandarin sentence-final particles *le* and *ba*. The study combines data from three studies to examine factors like learner type (heritage vs. L2), proficiency, and sentence structure. Early exposure (heritage learners) did not always lead to an advantage. L2 learners showed positive CLI at higher levels, but positive CLI varied depending on the specific language feature. The research calls for a broader view of positive CLI and more precise categories.

Chapter 5, by Tomoharu Yanagimachi, examines how native English speakers adapt their speech, i.e., foreigner talk, during ski repair conversations in Japan. Unlike past research in controlled settings, the study analyzes real-world interactions, where the focus is on achieving mutual understanding and eliciting expert advice. Both parties adjusted communication to their roles, not language proficiency. This offers insights for language learning and highlights real-world applications of foreigner talk. In Chapter 6, Trang Phan and Lan Chu explore time-related factors affecting Vietnamese acquisition by 63 Korean learners at Korean and Vietnamese universities. Proficiency assessment included class placement, self-assessment, and a grammar test. Time-related factors, including duration of exposure to a native environment,

total study time, and starting age, were considered. The findings aim to improve Vietnamese teaching for Koreans and foreigners interested in the Vietnamese language and culture.

Chapter 7, written by Jiun-Shiung Wu, Hsin-Hui Chen, and Hui-Ru Yang, introduces a system that analyzes Chinese sentences and generates grammar explanations named AGIBIS, short for Automatic Grammar Instruction Based on Input Sentences. Unlike large language models, AGIBIS uses a small-data natural language engine to identify grammatical points within a sentence. The authors compare AGIBIS to existing grammar correction tools and discuss its potential for teaching Chinese as a second language. The final chapter by Ying Ying Wong, Yuan-Min Shen, and Cheng-Chun Huang illustrates how a graduate-level course on vegetable diseases was adapted for general university students with English as a second language. Despite some adjustments due to the pandemic, students gave positive feedback on the EMI format. Experiential activities like vegetable growing and farm visits helped meet learning goals. While students from different backgrounds had similar experiences, some Taiwanese students faced challenges compared to foreign students. The study's teaching practices offer a model for future EMI course design.

This inaugural volume of *Current Issues in L2 Teaching and Research* offers a glimpse into the exciting world of L2 research. By showcasing diverse research on various topics, this collection paves the way for future exploration and improved learning experiences for all L2 learners. We look forward to the second conference in 2025 and, hence, the second collection of research papers in 2026.

Stano Kong
Yu-Hsin Huang
Miao-Chin Chiu
Kai-Lin Wu
One-Soon Her
July 24, 2024

EDITORS

Stano Kong is a Lifetime Distinguished Professor in the Department of Foreign Languages and Literature (FLLD) and founding director of the International Graduate Program of Teaching Chinese as a Second Language (TCSL) at Tunghai University. He has previously served as Chair of the FLLD and Dean of the College of Arts. His research interests and publications are in second and multilingual acquisition within the framework of generative grammar. He is currently working on the acquisition of L3 Spanish wh-questions by L1 Chinese speakers of advanced L2 English.

Yu-Hsin Huang is an Associate Professor at Tunghai University's International Graduate Program of Teaching Chinese as a Second Language. She earned her PhD from the University of Cambridge, focusing on the L2 acquisition of Chinese syntactic structures. With a Master's in Linguistics and a Bachelor's in English and Linguistics from National Taiwan Normal University, Dr. Huang has extensive experience in language teaching and research. Her work includes numerous publications in prestigious international journals and conference presentations on second language acquisition, particularly in L2 Chinese. She has also been involved in various research projects and received multiple grants and scholarships throughout her academic career.

Miao-Chin Chiu, an Associate Professor of the English Language Center at Tunghai University, brings over 25 years of experience as an L2 instructor of Chinese and English. She holds a PhD in Linguistics, focusing her research on the application of cognitive linguistics in second language education. She has lectured in Mandarin and English language programs at various

institutes, including Harvard University, Princeton in Beijing, the East Asian Summer Language Institute (EASLI) at Indiana University, and the International Chinese Language Program (ICLP) at National Taiwan University. Bridging the gap between theory and practice is her passion and commitment.

Kai-Lin Wu is an Assistant Professor in the Department of Foreign Languages and Literature at Tunghai University, Taiwan. She takes pride in being not just a researcher but also a practitioner. Her research on second language writing, self and peer feedback, and contrastive rhetoric has practical applications in the classroom. She teaches English academic writing to both undergraduate and graduate students and has authored two English composition textbooks for EFL students in Taiwan, which are a testament to the practicality and relevance of her work.

One-Soon Her is Nan Lin & Alice Lin Chair Professor of the Department of Foreign Languages and Literature at Tunghai University and an adjunct Chair Professor at the Graduate Institute of Linguistics, National Chengchi University. He is the founder and editor of *Taiwan Journal of Linguistics*, a Diamond Open Access journal. His recent research interests include numeral classifiers, numeral systems, typology of questions, and language policies. Most of his publications can be downloaded from his homepage at <https://onesoonher.github.io/info/index.html>.

CONTRIBUTORS

Hsin-Hui Chen received her PhD in linguistics from National Chung Cheng University and is an assistant professor at the Chinese Language Center, National Chung Cheng University. She is a seasoned Chinese language teacher. Her research interests include psycholinguistic and cognitive aspects of teaching Chinese characters, lexical semantics, application of AI to Teaching Chinese as a Second/Foreign Language, and so on.

Lan Chu has a PhD in linguistics and is a tenured Lecturer at the University of Languages and International Studies, Vietnam National University Hanoi. Together, Trang Phan and Lan Chu co-authored the three trilingual Vietnamese-English-Italian textbooks *Vietnamese Vibes 1, 2, 3* (Cafoscarina Press, 2024).

Cheng-Chun Huang works with Ying Ying Wong and lends their support to a course on vegetable diseases at the National Taiwan University. They have documented two years of valuable teaching and learning experiences within an English-Medium Instruction (EMI) environment for this course.

Atsushi Miura is a PhD student at the Graduate School of Fundamental Science at Waseda University. His research interests lie in second language acquisition and processing, particularly neurolinguistic and psycholinguistic approaches.

Shiro Ojima is a professor at the Center for English Language Education Center School of Engineering (CELESE), Waseda University. His research interests lie in English language education and second language acquisition, with a particular focus on child language acquisition.

Trang Phan, Assistant Professor in the Department of Asian and North African Studies at Ca' Foscari University of Venice, completed her PhD at the University of Sheffield in 2013 and subsequently conducted postdoctoral research at Ghent University. In 2020-2021, she was a visiting scholar at Harvard Yenching Institute. Dr. Phan authored the monograph *The Syntax of Vietnamese Tense, Aspect, and Negation* (Routledge, 2023).

Yuan-Min Shen is an assistant professor in the Master Program for Plant Medicine at National Taiwan University (NTU). With over 40 scientific papers, Dr Shen specializes in plant pathogenic microorganisms and plant protection.

Natsumi Shibata is a research assistant at Waseda University and a part-time lecturer at Senshu University. She obtained PhD at Tokyo University. Her research interests are child development and child language learning.

Darcy Sperlich is an Associate Professor of Linguistics at the Department of Applied Linguistics, Xi'an Jiaotong-Liverpool University. He works both in theoretical and experimental linguistics. He works within the realms of syntax, semantics, and pragmatics, with a focus on anaphora and factivity. He focuses on the East Asian languages of Chinese, Japanese, and Korean. He has strong research interests in second language acquisition, from the viewpoint of using it as a testing ground for linguistic theory. On the experimental side, he focuses on the use of psycholinguistic methodology, such as eye-tracking, to understand language processing of linguistic phenomena.

Yuki Takahashi obtained his B.A. and M.A. from Chuo University under Professor Wakabayashi's supervision. He currently works in a trading company.

Shigenori Wakabayashi obtained his PhD from the University of Cambridge. His academic interest lies in applied linguistics, particularly in modeling second language learners' mental representations of morphology and syntax in the framework of generative grammar and psycholinguistics. Professor Wakabayashi was a foundation member of the Japan Second Language Association (J-SLA), Secretary, and President of this association. He serves on the editorial board of *Linguistic Approaches to Bilingualism*, among other journals.

Ying Ying Wong works with Cheng-Chun Huang and lends their support to a course on vegetable diseases at the National Taiwan University. They have documented two years of valuable teaching and learning experiences within an English-Medium Instruction (EMI) environment for this course. Miss Wong, who previously provided support as a teaching assistant, currently puts her expertise to work as a plant doctor at the National Taiwan University Plant Teaching Hospital.

Jiun-Shiung Wu received his PhD in linguistics from the University of Texas at Austin and is a Professor of Linguistics at National Chung Cheng University (CCU), Chiayi County, Taiwan. He serves as the Director of CCU's Chinese Language Center. His areas of expertise include formal semantics, formal pragmatics, discourse analysis, syntax-semantics interface, application of AI to Teaching Chinese as a Second/Foreign Language, and so on. He is the author of two books: *Temporal and Atemporal Relations in Chinese* *Taiwan Journal of Linguistics Monograph No.2* and *Intensification and Modal*

Necessity in Mandarin Chinese. He has published dozens of journal papers, book chapters, and conference presentations.

Shanshan Yan is an Assistant Professor at the School of Chinese as a Second Language, Peking University. Her research focuses on language acquisition, including Chinese as a second and heritage language, and bilingual and multilingual acquisition in adults and children. She also examines language contact, learning motivation, and the influence of sociolinguistic and individual factors on language acquisition. Her publications include a monograph *The Right Periphery in L2 Chinese* (by Routledge), and various research papers (by high-level academic journals).

Tomoharu Yanagimachi has a PhD from the University of Minnesota and is a Professor of Applied Linguistics in the English Department at Hokusei Gakuen University in Sapporo, Japan. His research interests include second language studies, discourse analysis, and conversation analysis, with a recent focus on English and Japanese as lingua franca in workplaces in Japan. He has authored and co-authored book chapters and journal articles on the microanalysis of interactions involving both L1 and L2 speakers. He has also served on the editorial board of the academic journal *Shakaigengokagaku (Japanese Journal of Language in Society)*.

Hui-Ru Yang received her MA in linguistics from National Chung Cheng University and is currently a PhD student at the Institute of Linguistics, CCU. Her research interests include the application of AI to Chinese teaching, Chinese teaching, syntax, and semantics.

ANAPHORA IN L2 TEACHING AND RESEARCH: THEORY AND PRACTICALITIES*

Darcy Sperlich

Xi'an Jiaotong – Liverpool University

ABSTRACT

This chapter is intended to give a selected overview of anaphora as related to linguistic and second language acquisition theory, while discussing the practical aspects of this in the second language classroom. The main language focus is on Chinese and English, but other languages are touched upon where pertinent. Reviewed are the syntactic, semantic, and pragmatic theoretical approaches found in accounting for anaphoric patterns. It is argued that a single route to anaphora is too restrictive, and it is better to take into consideration multiple aspects that affect anaphora, language dependent. The focus then shifts to the classroom, where it is discussed how a Chinese learner of English may fare when shifting between two anaphoric systems (from the first to second language), noting that it appears reflexive pronouns may present the biggest issue. This is followed by a discussion on the politeness aspects of the *tu/vous* system, and what obstacles a language learner may face when acquiring it. Then, teaching aspects are considered. Finally, future horizons are explored, with suggestions for future research.

Keywords: Anaphora, second language, syntax, pragmatics, politeness

* I would like to thank Stano Kong for the invitation to write this chapter, and also to both reviewers for their constructive comments.

1. INTRODUCTION¹

Anaphora is a basic part of language, and it appears that there is no language without it. There are different types of anaphora, where our focus is on NP anaphora, such as the well-known categories of pronominals (1), reflexives (2), and zero anaphora (3) (underlined, among other types), as exemplified below (in English and Chinese):²

(1) John told him a story.

(2) John shaved himself.

(3) 我今天去玩，Ø 明天休息。

Wǒ jīntiān qù wán, Ø míngtiān xiūxi.

1.S.PRN today go play tomorrow rest

‘Today I’ll play, tomorrow I’ll rest.’

How anaphora distributes over different languages will vary, each having subcategories of each type. For instance, pronouns can differ in case, reflexives in their morphological make-up, and zero anaphora in their phrasal positioning, among other contrasts. Contrastively comparing Chinese and English for example, English has pronominals with different case marking (e.g., accusative/nominative/genitive, *she/her/hers*), whereas Mandarin Chinese has a single 3rd person singular form *tā* without case marking, but with different character representations (e.g., 他 (general human)/ 她 (feminine)/ 它 (inanimate)). Considering reflexive pronouns, in English they are usually formed using the schema pronoun+self, such as *himself*. In Chinese, on the other hand, there are different forms of reflexives including pronoun+self (他自己 *tāzìjǐ*), self (自己 *zìjǐ*), and self+self (自己本身 *zìjǐběnshēn*). Chinese is well known for its zero anaphora, being able to occur both in subject and object positions, whereas in English zero anaphora is highly

¹ Abbreviations used in this chapter: 1 = 1st person; 2 = 2nd person; 3 = 3rd person; NEG = negative; POS = possessive; PRN = pronoun; RP = reflexive pronoun; S = singular.

² For the purposes of this chapter, Chinese refers to Mandarin Chinese.

restricted (e.g., occurring in coordinate structures such as *I ate an apple and then __ walked out*, here a form of ellipsis).³

The next question we turn to is, what are ways to account for these anaphoric patterns? In the theoretical realm, we find many different approaches in the syntactic, semantic, and pragmatic arenas that aim to capture the behavior of different types of anaphora cross-linguistically. For example, syntactic theories (such as the well-known Binding theory (Chomsky 1981) and its successors) capture local binding patterns of reflexives well in English, as a long-distance (LD) dependency between *John* and *himself* cannot be established in (4), due to the locality constraints licensing binding only within the subordinate phrase:

(4) John₁ thinks that David₂ hit himself_{*1/2}.

However, when given a sentence such as (5) below in Chinese, this is better captured by formal pragmatics (specifically, Huang's (2000) revised neo-Gricean pragmatic theory of anaphora), given the choice between the LD and local antecedent. Namely, it is possible for *ziji* to be anteceded by both *Zhangsan* and *Lisi*, which depends upon if a logophoric perspective or standard reflexive marking is intended—there are no such locality constraints as found in English.

(5) 張三說李四不喜歡自己。

Zhāngsān₁ shuo Lǐsì₂ bù xǐhuan zìjǐ_{1/2}.

Zhangsan say Lisi NEG like RP

'Zhangsan₁ said that Lisi₂ doesn't like him₁/himself₂.'

Theoretical issues aside, we now consider why this is an important topic for both research and teaching. Addressing the latter first, one may question the reason for learning about anaphora in the classroom explicitly, apart from learning the different forms and associated functions. Indeed, it would be commonly assumed that anaphoric systems are somewhat similar between languages, and anaphora is commonly found in both input and output that it can be quickly mastered

³ However, zero anaphora is much more widespread in conversation compared to writing (cf. Oh 2005, 2006).

without overt intervention. Consequently, following the argument, in terms of teaching it does not appear to be an important aspect of the language classroom.

However, simply by comparing English and Chinese, one cannot assume a seamless one-to-one transition without running into some meaning and grammatical issues. That is, apart from frequency of usage (e.g., overuse of an anaphoric form as compared to a native speaker (English learners of Italian), see Sorace and Filiaci 2006, and resulting Interface theory), an English learner of Chinese may use a pronominal instead of a zero anaphor, or a Chinese learner of English may use a reflexive pronoun to behave like a pronominal. However, overtime, it should be the case that the anaphoric system would develop naturally towards a native speaker's one, and the question becomes how might this be accelerated in the classroom.

Apart from this, another important development is understanding the correct use of the forms, the pragmalinguistic (knowledge of form use as related to the grammar) and sociopragmatic (knowledge of form use as related to society) competence of the learner. For example, in Korean, 2nd person pronominals are to be avoided when speaking to someone who is older (Kim 2009), but an English learner of Korean may overuse them due to lack of sociopragmatic knowledge. In another light, languages encode power relations in the pronominal system, such as French with *tu/vous*; it is a crucial part of the society to be able to use the right forms properly for the intended message (cf. Sperlich et. al 2021, for research on Korean learners of French).

In terms of research, anaphora is a very productive area in second language acquisition (SLA) for theoretical reasons. First, research in SLA can be used to understand how general linguistic theories' predictions hold in a second language (L2) learning environment, much like in first language (L1) research. Let us take the L2 acquisition of reflexive pronouns, for instance. Sperlich's (2020) ERA model of anaphora predicts that English learners of Chinese will transfer their L1 anaphoric binding rules into their L2. English uses syntax to bind, while Chinese uses pragmatics to assign antecedence. In other words, English learners of Chinese will locally bind in situations as (5) above, without realizing the LD antecedent possibility (until they become more

proficient with the pragmatic system)—due to negative transfer of their English binding rules. As reviewed in Sperlich (2020), this is certainly the case found in different experiments.

A general SLA theory (one that is not anaphora focused) to consider is the Shallow Structure Hypothesis (SSH, Clahsen and Felser 2006), learners in the beginning make more use of discourse cues than morphosyntactic cues in their L2 processing. Thus, when a Chinese learner of English begins learning English, in terms of anaphora, it could either be theorized that they transfer their pragmatic binding L1 strategies into their L2 (as predicted by Sperlich 2020), or instead it could be claimed that they make use of general discourse mechanisms, as predicted by the SSH. Both theories would be in line with the data found thus far (their binding patterns resemble what is found in their L1), but have very different base assumptions. However, how this might extend to English learners of Chinese is problematic for the SSH, as clearly, they bind locally at lower levels of proficiency given L1 syntactic transfer, whereby predicted discourse cues are not taken advantage of to allow LD binding (LD binding occurs at higher levels of proficiency as the pragmatic system is acquired). If a theory's predictions are correct, then this gives more credence to the assumptions the theory rests upon. Then, it is possible to utilize the underlying assumptions to power L2 teaching (indirectly at least), as teaching can be adjusted to appeal to the core aspects.

Therefore, it is crucial to see how theory holds in a language learning situation given the amount of multilingualism found in the world, rather than considering models developed with monolingualism (or specific languages) in mind. For instance, assuming Chinese anaphora is pragmatically regulated, then in the classroom the teacher can utilize context more to help guide antecedent judgement, rather than relying on a 'rule of thumb' approach, or worse, present examples which do not capture the range of possible antecedents (while considering the learner's L1).

The rest of the chapter is structured as follows. First, we review how one may approach the study of anaphora, and give a brief review of theories in the area. Following this, we then review how this might be incorporated into the SLA domain, in the sense of observing how

predictions from both linguistic and SLA theoretical perspectives may be incorporated. Then, we consider applied perspectives, including politeness, as the use of anaphora (particularly pronominals) encodes social relations, which is particularly important for a learner to understand. We next shift to teaching perspectives on what has been discussed. Finally, future horizons are considered.

2. THEORETICAL APPROACHES TO ANAPHORA

A pertinent question to ask is, how can one model anaphora formally? The answer to this, is: it depends on your beliefs, which stem from your background and training. In the theoretical linguistics surrounding anaphora, there are three major areas, being syntax, semantics, and pragmatics, each with their own goals and perspectives. However, even considering before which syntactic, semantic, or pragmatic approach to take, an even more basic question needs to be addressed: is language considered to be domain specific, general, or a mix? If the belief is language is domain specific, then one of the most influential schools of thought adhering to this assumption is Chomsky's Universal Grammar (UG), which has evolved over the decades into the Minimalist program (Chomsky 1995). Thus, it is likely that a syntactic approach is adopted for anaphora resolution. If language is considered not to be completely domain specific, such as found in the Emergentist approach (Christiansen and Chater 2016), then the starting point to antecedent resolution does not necessarily recourse to syntax—it could be that syntax plays a role with other mechanisms operating at the same time in antecedent resolution.

Assuming the adoption of a UG approach to language, a major anaphoric theory that has had very strong influence is Chomsky's Binding theory (Chomsky 1981). Namely, anaphors (such as reflexives) are locally bound (via Principle A), pronominals are not locally bound (Principle B), and Principle C focuses on r-expressions. An important aspect has been in resultant theories (such as Reuland 2011), is to deconstruct the Binding principles and instead attribute binding behavior to interactions in the grammar, resulting in the capture of local binding.

For example, in (4) above, *himself* is bound locally given its morphosyntax and grammatical environment (for details, see Reuland 2011). However, with (5) above, long-distance binding of the anaphor via syntactic mechanisms is not easily accomplished (although there are many theories that work on this), given the contextual influence one can induce. That is, if we consider (6) below, the LD antecedent now is ruled out given the contextual influence that usually a doctor will spend time to prepare him/herself for the surgery, rather than a patient:⁴

- (6) 病人說醫生明天會準備自己的心情開刀。
 Bìngrén₁ shuō yīshēng₂ míngtiān huì zhǔnbèi zìjǐ_{?1/2} de xīnqíng
 patient say doctor tomorrow will prepare RP POS mood
 kāidāo.
 operate
 ‘The patient₁ said that the doctor₂ will prepare his/her_{?1/2} mood for tomorrow’s operation.’

The main point is thus, anaphoric patterns seem to be straightforward when observing one language, however, when considering anaphoric behavior cross-linguistically, it becomes much more complex.

The study of pronominals, on the other hand, does not engender the same amount of theoretical problems, but it has received attention from first language acquisition researchers in the sense of pronominals may prove difficult for children dependent on language (cf. Di Sciullo and Agüero-Bautista 2008). Usually pronominals are not locally bound, and are usually seen to be in contrastive distribution with anaphors (however, see Huang 2000:21, for description of problems). Moreover, their antecedence can be found in the preceding discourse such as *John told him to go home*, where *him* must be traced back to another antecedent in a previous sentence given a normal conversation (or current context). However, this is not to say that learners do not have problems with pronominals, as it depends on what the possible antecedent being presented is (e.g., quantified antecedent, full NP; cf. Cummings et al. 2015; Slabakova et al. 2017, for experiments)

⁴ For notes on interaction of world knowledge and antecedent selection in Chinese, see Huang (2000:237).

Finally, the study of zero anaphora is a very productive area, and also theoretically interesting in understanding how antecedence is recovered (cf. Camacho 2013, for review). In languages that are morphologically rich, such as Spanish, the antecedent is easily identified using the morphological clues found on the verb. However, for a language such as Chinese, that does not have such morphological clues, it is impossible to use the same mechanism to understand antecedence. While there have been different syntactic attempts at this, there are also pragmatic routes to understanding the intended antecedent. Studies in this area pair different types of languages that may or may not have productive zero anaphora (see Contemori and Dussias 2020, for one).

Thus far, we have briefly considered anaphora from a syntactic standpoint. The implications of adopting a theory with syntax at the core, is the belief that antecedence would be syntactically regulated. Consequently, shifting to the SLA realm, the implication is that the learner is constructing a syntactic system and alike rules to guide binding. Therefore, with this assumption, it is then logical to adopt a theory of SLA that focuses on the syntactic system, such as Full Access/ Full Transfer (Schwartz and Sprouse 1996). The results of such studies then can help understand the predictions made from the SLA theory, and more deeply, the linguistic base assumed. However, studying SLA is not straightforward, as learners may bring to the fore strategies that are non-syntactic, or have cognitive limitations, complicating the interpretation of results, and the creation of studies going forward (e.g., see Hendriks 2016, for a cognitive perspective). This brings us to the pragmatic study of anaphora, however before this, we briefly look at semantic issues.

Many formal theories assume that syntax precludes semantics, in the sense that if one wishes to understand what a sentence means, one needs a syntactic representation first to order constituents, to be fed into a semantic representation. In other words, we understand well what *The dog was chased by the man* means, but if the order is jumbled as **The by the chased dog was man*, while all the words are there, it has no coherence due to lack of structure. Thus, when considering the study of anaphora, formal semantics focuses on antecedent resolution via predicate logic and Lambda calculus—the theory assuming that meaning in language is a logical system, and thus can be captured using principles

of logic (the idea stemming from Gottlob Frege, with some modern systems based on Montague Grammar). A famous sentence examined is *Every farmer who owns a donkey beats it*, while a normal and grammatical sentence, the pronoun *it* presents challenges for translation into formal logic (cf. Kamp et al. 2011, for exposition). Meaning representation is a difficult subject as semantic structure is not as observable as compared to syntax.

While it might be difficult to see how this might be applied to second language research, an important contribution of semantics is that of understanding anaphora through *de se* (Chierchia 1989). *De se* basically is the belief that one has about oneself, as seen through *John₁ thinks that he₁ looks good today*. As applied to Chinese, Pan (1995) argues that *ziji* is bound only in a *de se* context, by a self-ascriber. Therefore, armed with semantic machinery, anaphoric binding can be understood through meaning lenses. Moreover, it should be mentioned that perspective can be important when considering long-distance binding. An important point is that of logophoricity (to be discussed below), whereby the perspective of the sentence internal protagonist is being taken, rather than the sentence external utterer (cf. Huang 2000, for discussion). In (5) above, it is the case that the matrix subject is able to be a binder, sanctioned via logophoricity. Moreover, in Japanese, empathy is important when considering binding, which is roughly understood as the point-of-view taken (see Oshima 2007, and earlier references therein). Oshima (2007) takes the semantic perspective to capture these different effects (e.g., when considering the Japanese reflexive *zibun*), showing successful modelling, and this should be a factor taken into consideration when acquiring a language of this type.

In terms of SLA, it usually isn't the case that studies are designed to assess formal semantic logic, given that formal semantic theory isn't concerned with language acquisition generally, and predictions by the theory would be difficult to apply to a second language environment. However, this isn't to say that the ideas of *de se* can't be implemented, it is just difficult to juggle the syntax of the sentence that contributed to the semantic representation, and to hold *de se* effects separately in an experimental context. In saying this, ideas of logophoricity and the like can be approached in different ways (e.g., see Charnavel 2019, for a

syntactic approach to logophoricity), which leads us to a pragmatic approach.

The pragmatic approach to anaphora discussed here is neo-Gricean, using Huang's (2000) theory, the revised neo-Gricean pragmatic theory of anaphora. Namely utilizing neo-Gricean pragmatics (Levinson 2000), Huang aims to capture anaphoric patterns using formal pragmatic principles. Without diving into the theory (refer to Huang's works for details), the key idea is that the focus is on the use of form, and how it can be used to send an intended message. Thus, the assumption is an uttered sentence that is both grammatical and meaningful, its intended meaning can depend on the use of one form over another. For example, considering (7) below, where the anaphoric position can grammatically contrast between different forms, different meanings can arise:

- (7) a. 醫生說病人知道明天給他開刀。
Yīshēng₁ shuō bìngrén₂ zhīdào Ø_{1/3} míngtiān gěi tā₂
doctor say patient know tomorrow for 3.S.PRN
kāidāo.
operate
'The surgeon₁ says that the patient₂ knows that
(he_{1/3}/I/you/we/they...) will operate on him₂ tomorrow.'
- b. 醫生說病人知道他明天給他開刀。
Yīshēng₁ shuō bìngrén₂ zhīdào tā_{1/3} míngtiān gěi tā₂
doctor say patient know 3.S.PRN tomorrow for 3.S.PRN
kāidāo.
operate
'The surgeon₁ says that the patient₂ knows that he_{1/3} will operate
on him₂ tomorrow.'
- (Huang 2000:240)

This being the case, it is clearly not the result of a different syntax given the sentence structure is the same, and it only depends on the forms used. Thus, it stands to reason, if a native speaker is aware of these different meanings, and effortlessly understands different meanings arising from different form usage, then this can be modelled formally,

which is what Huang's theory aims to achieve. Considering then, that the formal pragmatic system behind anaphoric resolution is considered to be a universal, it stands to reason that this will be a factor in learning how to use anaphoric elements. However, there is one major caveat with this approach—it might not be applicable to all languages, not due to non-universal availability, but the choice of the language to use it. That is, the theory predicts that languages lie on a syntactic-pragmatic continuum in relation to what anaphoric strategies are present in the language, resulting in languages that are more syntactic, and languages that are more pragmatic (anaphorically speaking). What this means, is that dependent on the criteria observed (cf. Huang 2000, for list), such as reflexive pronouns, we find English utilizing syntactic strategies to resolve antecedence (given the clear preference to bind locally); Chinese, on the other hand, again with reflexives, shows it resolves antecedence via pragmatics (considering the influence of context).

Considering the SLA acquisition of anaphoric elements through the pragmatic perspective discussed here, my own work utilizes this approach. For instance, Sperlich (2013) researched the SLA of the Chinese reflexive *ziji* by English and Korean learners of Chinese. It was hypothesized that English learners of Chinese initially transfer their syntactic binding strategies into their L2 Chinese, resulting in a type of negative transfer (with a strong preference to locally bind). Korean learners of Chinese, on the other hand, transfer their pragmatic strategies into their L2 Chinese, resulting in positive transfer (allowing LD binding and being affected by context). The results of the study showed just this (see also Lee 2008). There are also other ways that pragmatics is incorporated into studying the acquisition of anaphora, for instance see García-Tejada (2022) for English learners of Spanish and clitics.

The linguistic road forks into different directions, and each fork forks further dependent on the theory of the area. Combined with SLA theories, a researcher must carefully weigh which linguistic theoretic approach to adopt, and pair it against an SLA theory that has similar base assumptions. However, this does not necessarily need to be the case, as it is also possible to work with theories which incorporate different linguistic modes. For example, Reuland's (2011) theory posits binding at different levels (syntax, semantics, discourse), however, his theory and

ongoing works posit heavily syntactic-based solutions. In my own work, Sperlich (2020), a middle ground is taken positing both syntactic and pragmatic processing to anaphoric resolution. Ultimately, what the researcher believes to be the best fit to the research problem will guide which theory to adopt—this should be done consciously and not according to underlying assumptions prevalent in the literature.

It should be noted however, that the SLA researcher most likely has better knowledge about SLA theories rather than linguistics theories on anaphora. For instance, it might be the case that a researcher wishes to test the Shallow Structure Hypothesis (Clahsen and Felser 2006), which posits early use of discourse mechanisms while morphosyntactic cues are being established to gain more prominence, or a more UG orientated theory of Full Access/Full Transfer (Schwartz and Sprouse 1996), among many others. Then, this researcher, having an interest in anaphora, works towards understanding how, for example, reflexives and pronominals are interpreted in the L2, by comparing different L1 speakers learning the same L2. At this point, given the pragmatic and syntactic factors that influence binding, the researcher needs to be aware of the L1 strategies. If one assumes all binding is syntactic, then apart from missing contextual issues, the interpretation of the data will be one-sided, only working with SLA theory which is syntactically based (for better or for worse). If one assumes that different languages may have more than one strategy to bind, then it would be better to adopt an SLA theory that includes multiple pathways. Thus, mixing theories from the linguistic and SLA domains requires special attention.

Finally, it should be pointed out that one may hold a typical underlying assumption that there is a distinction between the grammar and parser, or in UG terms, competence versus performance. Much of what has been discussed has been focused on the grammar perspective, i.e., the acquisition of the underlying grammar system, and the assessment of its acquisition. However, another perspective on the problem is to observe the parsing of that system, which takes a more psycholinguistic perspective. That is, understanding the real-time processes can cast light on how different systems interact, or how different pressures affect the use of the system, or of the parser itself. For example, cue-based processing investigates how morphosyntactic cues

are taken advantage of with special reference to memory (Lewis and Vasishth 2005). Example of anaphoric (L1) studies in this experimental realm are Parker and Philips (2017) and Parker (2019), among many others, and continue to be a productive area of study. SLA anaphoric studies with a more psycholinguistic perspective are fewer, Cunnings et al. (2017) being a good example. In understanding and interpreting experimental results, it also depends on one's theoretical perspective—given that experimental data are open to interpretation. Finally, it should be noted that it is not a universal assumption that there is a distinct parser or grammar—Cognitive Linguistics makes no such assumptions and it has been applied to anaphora (cf. Hirose 2014, for application to Japanese). O'Grady's (2005, 2022) theory assumes there is no syntactic representation (as Sperlich 2020, follows, with the inclusion of a neo-Gricean pragmatic processor), but rather a direct form-to-meaning parse using specific language algorithms via a language specific sentence processor.

In summary, we have reviewed particular linguistic approaches to anaphora covering the syntactic, semantic, and pragmatic realms. How these might apply in different languages require careful consideration; thus crosslinguistically, one needs to be vigilant when considering the data and arguments. How these theories might meld (or not) with SLA theories has been touched upon, driving the point that careful consideration needs to be made in respect to the standpoint one takes. In the next section, we now consider more applied approaches to anaphora, in terms of teaching and learning.

3. APPLIED PERSPECTIVES TO ANAPHORA

The type of theoretical research discussed above is not straightforwardly integrated into teaching. As discussed in the Introduction, anaphora is generally assumed to be learned without much effort, and the focus of language learning is on developing the many other complex aspects of language. Of course, anaphora is anything but simple; however, our ability to assign antecedence quickly and accurately makes it look deceptively simple. Here, discussed are two pertinent aspects of

anaphora that are complex—how language differences can hinder/support acquisition of anaphora in the target language (leading to negative/positive transfer), and the issue of politeness.

An important aspect of language learning that language teachers are well aware of, is language transfer from the L1. Let us deal with an example to guide our discussion: what difficulties or advantages would a Chinese learner of English have in acquiring English anaphora? First, there are general issues of a possessive pronoun in Chinese which differs to English, having the form pronoun + POS (e.g., *his/her* vs. 他的 *tā de*), which doesn't appear to be much of a challenge, or plural marking which again is spilt (*them* vs. 他們 *tāmen*).⁵ One important aspect is the gender marking on English pronominals, which isn't found in Chinese (see Dong et al. 2015, for an interesting psycholinguistic study that takes advantage of this). That is, Chinese pronominals are not specified for gender, thus one will occasionally hear in a learner's English use of *he* when referring to a female, and quickly rectifying the mishap. What is the possible cause of this, since the input is rich enough to make distinctions between pronominals referring between male/female? This can be attributed to the genderless Chinese pronominal 他 *tā*, which is contrasted against another written form with a female radical, feminine 她 (introduced into Chinese in the early 20th century (Huang 2009); however, 他 is used generally regardless of gender). Therefore, it can be understood that the same strategy would be applied into English, to treat *he* as default until attention paid to gender becomes engrained. Accordingly, if the language teacher believes that languages naturally make such a distinction in the grammar, then confusion might result when dealing with learner errors such as these. Shifting briefly to a Korean learner of English, Korean is a language that has pronominal forms, but their usage varies. For example, ㅁ *ku* (3rd person singular pronoun) was introduced into the language again in the 20th century and has different usage patterns and frequency as compared to English (cf. Kim and Han 2016). Consequently, the Korean learner of English may have different pronominal usage in his/her L2 due to L1 language norms.

⁵ It is important to note that the anaphoric systems of Chinese languages have differences. For example, in Minnan found in Taiwan, the 2nd person plural pronoun is monomorphemic 恁 *lín*, whilst in Mandarin, it is separately marked.

In this sense, while a form is recorded formally, one needs to further investigate its usage in the specific language.

The next line of inquiry is the reflexive pronoun. Reflexives come in different shapes and sizes; generally, there are three major types, self, pronoun+self, and self+self.⁶ Staying with our Chinese learner of English, we find that there is partial overlap with English, the pronoun+self form (e.g., 他自己 *tāzìjǐ* vs. *him/herself*). Considering Chinese, on the face of it, this does not seem to be too problematic grammar-wise given the similarity of word order between the two languages as seen in (5). However, in Chinese, reflexives can occupy a subject position, whereas in English they cannot. Chinese has a self reflexive, whereas English does not (apart from being found in headlines). The former does not seem to be an important issue as quickly it will be discovered that there is no nominative form of the reflexive, but the latter does have more implications.

The implications of English not having a self pronoun, in terms of a L2 learner's L1 that does, is the issue of logophoricity. That is, as discussed above, we see that logophoricity allows licensing of LD binding in Chinese. This is an important reason why we do see LD binding in Chinese learners of English—they transfer a logophoric perspective into their L2 English, but without the simplex self form present in English, they substitute the complex form. Due to L1 transfer, they might expect the same results due to word order similarities. Hence, in terms of L1 transfer, the L2 learner is transferring the notion of logophoricity into his/her L2 English along with the required machinery, with incorrect results. It should be noted that under this scenario, it is assumed that the learner has transferred his/her L1 pragmatic binding strategies into his/her L2, thus interfering with the construction of English syntactic binding.

In this sense, we might be tempted to simply focus on form transfer—English and Chinese both have reflexives, and one which looks very similar is pronoun+self. In this respect, however, it isn't the

⁶ This is an oversimplification of the variety of forms reflexive pronouns can take, e.g., some languages use body parts (Huang 2000). It should also be noted that some languages lack reflexive pronouns, using different strategies to indicate reflexive relationships.

form that is being transferred, rather it is what the form is being used for. For instance, in (8) below, it gives rise to a contrastive/ emphatic reading (he and not someone else), as compared to the use of *zìjǐ*:

- (8) 他說他自己要去。
Tā shuō tāzìjǐ yào qù.
3.S.PRN say 3.S.PRN.RP want go
'He/she said that he/she him/herself wants to go.'

In English, only a contrastive/ emphatic meaning can be found in the same position, rather than a reflexive relation, giving similar usage. However, within a standard SVO pattern, in both languages a standard reflexive relation arises. Thus, instead of considering form transfer as cause, it is important to see what the function of that form is, and how it can be understood in different sentence positions.

Finally, regarding zero anaphora, as the constructions found in Chinese licensing its use aren't found in English, it appears that this should not be a clear problem past the beginning stages in English as its use will lead to ungrammatical sentences given subjects are not dropped productively (apart from imperatives).

Regarding these three types of anaphora, we see that issues arise with pronominals and reflexives for the Chinese learner of English, considering this specific L1-L2 combination. Whether or not other language combinations will cause learners the same issues depend on both what anaphoric forms are available in the L1 and L2, and how they are licensed. For example, German also has a self-self form *sich selbst*, however, it is not used as a typical reflexive but instead used as an intensifier (Reuland 2017). English does not have a self-self form, and Rankin (2017) notes that in writing, German learners of English tend to overuse *himself* as an intensifier, which adheres to transfer of using L1 style intensification in the L2.

All this assumes a language transfer hypothesis. Another perspective to take is that of the Shallow Structure Hypothesis, which theorizes that less proficient learners use discourse mechanisms to process the L2 (in particular cases), due to a lack of established morphosyntactic cues in the system. Consequently, instead of appealing to language transfer, we

could instead argue that learners process antecedence via discourse mechanisms, until the morphosyntax becomes dominant in the L2 processing. This is attractive as instead of appealing to language transfer, there is more of a universal early mechanism at play. This indeed could be the case and help explain the issues faced with learners, but as we know it, usually learners from different languages show different kinds of errors (not just in anaphora usage, which presumably can be traced to their L1). For instance, English learners of Chinese do not LD bind and prefer local binding, instead of allowing LD binding through discourse (only occurring at latter stages of proficiency); and morphosyntactic cues are not so helpful in deciding antecedence in a pragmatically-based anaphoric system. Therefore, language transfer has an important role to play (which is assumed by many theories, such as Rothman's (2011) Typological Primacy Model).

Working with Huang's (2000) syntactic-pragmatic continuum thesis, Sperlich (2017) theorized that learners need to shift from one anaphoric system to another. In the case of Chinese learners of English, they have to shift from a pragmatic-based anaphoric system to a syntactic-based one, and English learners of Chinese *vice versa*. Thus, for the language teacher, the focus here should be on providing input and output opportunities to encourage this transition. Again, it is noted that acquiring the anaphoric items is not the issue, rather it is the system that governs their usage. One possibility that can be considered is to show where the use of context is not relevant to binding choices; for (9a) in a stereotypical situation, students generally worry about the professor's opinion about him/herself (the student), a reading that is not available in English but is available in Chinese (9b).

(9) (a) The student₁ is worried that the professor₂ hates him/herself_{*1/2}.

(b) 學生擔心教授恨他自己。

Xuéshēng₁ dānxīn jiàoshòu₂ hèn tāzìjǐ_{1/2}.

student worry professor hate 3.S.PRN.RP

'The student₁ worries that the professor hates him/her₁/him/herself.'

However, this does conflict with examples such as *David₁ told John₂ a story about himself_{1/2}*, so-called ‘picture sentences’ where *himself* is not found in an argument position; it is ambiguous who the story is about—context is the key in these circumstances. In this sense, it also becomes important to understand argument relations. In any case, this is part of the transition, and it should also be noted that the transition is not something which becomes finalized or complete. What this means, is that, due to L1 influence, it is possible to revert back L1 practices in the L2. For instance, McCormack (2003) researched Japanese learners of English, and traced their binding preferences when they returned to Japan. It was found that over time, the learners slowly started showing patterns of binding from their L1 in their L2 English. In this sense, L2 systems are susceptible to attrition without sufficient maintenance opportunities (see Schmid and Mehotcheva 2012, for overview).

3.1 Politeness considerations

Much of our discussion has revolved around the basic anaphoric system, and how this is affected in the L2 through language transfer. Another very pertinent issue is the question of how anaphoric forms integrate with power relations, typically understood as politeness (cf. Brown and Levinson 1987). Politeness strategies vary from language to language, and this can also be built into the language grammatically (consider honorifics in Korean and Japanese), including specialized terms of address. The phenomena discussed here is the well-known *tu/vous* system of address (here represented by French), the ‘close’ *you* versus ‘distant’ *you* in terms of human relationships (or solidarity vs. power). Originally formalized by Brown and Gilman (1960), the basic pattern of using *tu* in relations with people known such as friends, and *vous* with people such as strangers, is found across many languages, such as German and Spanish. Considering French, this type of ‘rule of thumb’ usage does not come close to capturing the variation found in French, as *tu/vous* is closely embodied within French culture. There have been many different studies done on *tu/vous* style systems, and also within the classroom (cf. Sperlich et al. 2021, and references therein).

An anecdotal example given of the complexity is that of a service encounter. In a class I was teaching on sociolinguistics, the discussion was on *tu/vous*, and how there might be variations to its usage. A French student mentioned that when she was a restaurant server, the requirement was to use *vous* with customers to be polite. However, it is well understood that peers (those that appear close in age), are free to use *tu* with one another. Thus, she was somewhat aggrieved that some peer customers would abuse this privilege as they would use *tu*, and she would be forced to use *vous* in return, when they should have used *vous* in solidarity. Considering this, among many other deviations from the ‘rule of thumb’, what hope does an L2 learner have by learning simple *vous* for strangers, and *tu* for friends? The answer is clear, one must reside in the country and be immersed in the environment, as is often the case.

In terms of language transfer, learning *tu/vous* can be influenced by the L1 both in terms of sociopragmatic concerns (connections between society and usage), and pragmalinguistic forms (connections between usage and form). Beginning with the latter, a language teacher could assume that the L1 of an L2 learner of French has 2nd person pronouns such as ‘you’; with an English speaker this is the case, however, with a Korean speaker, this is true but the same usage assumptions do not hold. Namely, in English, it is common to use ‘you’ when addressing someone, being a very productive form. In Korean however, the use of ‘you’ is restricted—it is not considered polite to address someone directly with the pronoun (especially elders). Thus, while the L1 forms are found, certainly they are mapped differently in terms of sociopragmatic functions.

In terms of an English learner of French, it will be the case to understand the two different forms and their functions in relation to the sociopragmatics, whereas the Korean learner of French, he/she must first come to terms of comfortably using the pronominal form to address others, while understanding how *tu/vous* differ—the starting points of the learners are different due to the L1 cultural background. We could consider a German learner of French—also with a *du/sie* system—the case would be then to consider the sociopragmatics as German and French cultures operate differently, and one cannot simply transfer the

practices of the form usage into another language without problems arising. Regarding a Chinese learner of French on the other hand, Chinese appears to have a similar system with 你/您 *nǐ/nín*. However, again, second person pronominals are generally avoided in addressing someone, and 您 *nín* usage is highly restricted to contexts such as formal interactions between students and professors, or service encounters where the server wants to display a clear attitude of the customer being superior (cf. Wang 2022, for a study on Chinese learners of Spanish). Language transfer here, clearly can be problematic, culturally speaking.

To summarize this section, we have briefly observed how the anaphoric system may cause problems for learners in terms of language transfer—the astute teacher being aware of the learner’s L1 can pave the way for a quicker acquisition of the L2 anaphoric system. We also discussed politeness considerations, which are important for societal reasons. Again, if a teacher is aware of the politeness distinctions made by the learner’s L1, then again, the teacher will be better able to guide the learner into understanding the L2 cultural norms in using these forms, as compared to the learner’s L1.

4. TEACHING CONSIDERATIONS

In terms of the above discussion on anaphora and politeness, what practical implications are there for the classroom? Considering anaphora in general, as discussed, it doesn’t appear that there are languages that lack anaphora, and a teacher generally can assume that this is an understood component in the classroom. However, given the variety of forms and their usage over different languages, there cannot be a one-fits-all solution to its teaching. If a teacher knows the L1 of the student, then this can be taken advantage of in the classroom to design specific teaching activities. Considering a Chinese learner of English, it can be implicitly taught that LD binding does not occur in English, and show how logophoricity is represented through targeted dialogue. For example, showing how a person’s perspective is represented through contrasting reflexives and pronouns in differing sentence positions. For English learners of Chinese however, they must be made to understand that LD

binding is permissible in Chinese, which requires understanding the use of the self form 自己 *zìjǐ*. In terms of politeness, the same principles apply. If we consider a French learner of Chinese, there might be a mismatch between the L1 *tu/vous* and L2 你/您 *nǐ/nín*, which have different sociopragmatics. Assuming the teacher does indeed know enough about the sociopragmatics of these forms, then again communicative activities can be designed (such as service encounters) to show when the forms would, or would not be used (a term of address would be preferred though). Of course, classroom time is limited, thus including these points throughout a student's course is recommended as receiving constant input will help structure the L2 system towards L1 norms. We now move to the final section, considering future directions.

5. FUTURE HORIZONS

What does the future hold for research in anaphora? For the majority of languages, anaphora is well understood in terms of the forms used and how reference works. In terms of theory however, the debate between which theory best predicts the data remains a hot topic for many years to come. Is it the case that all binding can be captured by a syntactic or pragmatic theory, or is it the case that multiple routes need to be taken? Considering recent trends, it appears that multiple routes to anaphoric establishment will lead us into the future, in terms of dividing the labour over different areas of linguistic processing. As is per usual, new theories will be created to reflect up-to-date data patterns found (hopefully, cross-linguistically), to be measured against previously established theories in terms of what advantages may be had.

In my own work (Sperlich 2020), the ERA model's direction taken leads to a more psychological orientated theory of linguistics, one that takes into account processing pressures as advocated by O'Grady (2005), with Emergentism as the key guiding tenet. In this sense, my work eschews syntax as a medium assumed by syntactic theories such as Minimalism, advocating instead for a direct form-meaning algorithmic parse via a sentence processor, with a pragmatic processor working in tandem on the meaning representation. What this means is that the ERA

model recognizes both syntactic and pragmatic traditions of understanding anaphoric reference resolution, however, it does so without recourse to tree structure, and has integration with parsing. Ultimately, the theoretical perspective one takes depends on one's overall aims.

Studies on teaching on the other hand, are rather sparse when it comes to understanding how a learner might grapple with anaphoric reference. As discussed previously, the assumption that the learner will easily acquire the forms of anaphoric items is not misplaced, and usually the classroom is focused on communicative competency. In this light however, this area shows a clear gap for application to the classroom—what methods could be implemented by the teacher to speed up acquisition of the system? As noted previously, in regard to reflexives, it appears only at the advanced stages of proficiency do learners bind in a native-like fashion—this is rather alarming, and it shows that the system is far more complex than understanding a few items. However, note again that theoretically there are many possible perspectives to take, thus any application in this area should take base linguistic assumptions into account.

In terms of politeness on the other hand, this is a very productive area of research, especially considering the sociopragmatic implications it has for students in terms of communication. The literature is rich, providing many possible routes for a researcher. For example, what type of teaching method can hasten the acquisition? How should learner attention be drawn? These are all active areas of research. In terms of more targeted research, there seems to be less understanding on what has been discussed here in terms of Korean for example, how the use of a second person pronominal in the L1 affects *tu/vous* acquisition, or in Chinese, how a pseudo *tu/vous* system might affect the acquisition. Thus, mixing learner L1-L2 languages with *tu/vous* systems should be a target of research if one wishes to further understand the interplay between pragmalinguistics and sociopragmatics between alike systems.

Finally, in drawing together theory with data, it is important to note that apart from exploring established patterns of binding, it is also important to keep an eye on different forms that escape attention. Taking reflexives as an example, Stern (2019) investigates reflexive forms that

are generally not found in Standard English, yet they are used productively. Japanese has a wide range of reflexives, yet they have not been explored fully (cf. Kogusuri 2019). Huang (2000) details a wide range of anaphoric variation one can find in different languages, which is ripe for in depth exploration. With better descriptive data, paired with experimental data, can we come closer to an all-encompassing theory of anaphora.

6. CONCLUSION

In this chapter, we have reviewed theoretical trends in researching anaphora over the syntactic, semantic, and pragmatic realms. It was argued that instead of taking a single route to understanding anaphora, it is better to work with a theory that allows different routes to antecedent resolution. We then shifted to a more practical perspective in the sense of understanding how different anaphoric phenomena may present issues for learners through language transfer lenses. We also discussed how politeness and pronominals are an important linguistic issue for learners considering societal aspects. We briefly observed how teaching may incorporate these points. Finally, we considered future horizons for research. Anaphora remains a pertinent topic in linguistics due to the complexity of systems involved, and the future looks bright for its continued study.

REFERENCES

- Brown, Rodger, and Albert Gilman. 1960. The pronouns of power and solidarity. In *Style in Language*, ed. T. A. Sebeok, pp.252-281. Cambridge: MIT Press.
- Brown, Penelope, and Stephen C. Levinson. 1987. *Politeness*. Cambridge: Cambridge University Press.
- Camacho, José A. 2013. *Null Subjects*. Cambridge: Cambridge University Press.
- Charnavel, Isabelle. 2019. *Locality and Logophoricity*. New York: Oxford University Press.
- Chierchia, Gennaro. 1989. Anaphora and attitudes de se. In *Semantics and Contextual Expression*, eds. R. Bartsch, J. van Benthem, and P. van Emde Boas, pp.1-32. Berlin: De Gruyter Mouton. DOI: 10.1515/9783110877335-002
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge: MIT Press.
- Christiansen, Morten H., and Nick Chater. 2016. *Creating Language: Integrating Evolution, Acquisition, and Processing*. Cambridge: MIT Press.
- Clahsen, Harold, and Claudia Felser. 2006. Grammatical processing in language learners. *Applied Psycholinguistics* 27:3-42.
- Contemori, Carla, and Paola E. Dussias. 2020. The processing of subject pronouns in highly proficient L2 speakers of English. *Glossa: A Journal of General Linguistics* 5(1):38. DOI: 10.5334/gjgl.972
- Cummings, Ian, Patterson, Clare, and Claudia Felser. 2015. Structural constraints on pronoun binding and coreference: Evidence from eye movements during reading. *Frontiers in Psychology* 6:840. DOI: 10.3389/fpsyg.2015.00840
- Cummings, Ian, Fotiadou, Georgia, and Ianthi Tsimpli. 2017. Anaphora resolution and reanalysis during L2 sentence processing: Evidence from the visual world paradigm. *Studies in Second Language Acquisition* 39(4):621-652.
- Di Sciullo, Anna Maria, and Calixto Agüero-Bautista. 2008. The delay of Principle B Effect (DPBE) and its absence in some languages. *Language and Speech* 51(1-2):77-100. DOI: 10.1177/00238309080510010601
- Dong, Yanping, Wen, Yun, Zeng, Xiaoming, and Yifei Ji. 2015. Exploring the cause of English pronoun gender errors by Chinese learners of English: evidence from the self-paced reading paradigm. *Journal of Psycholinguistic Research* 44:733-747. DOI: 10.1007/s10936-014-9314-6
- García-Tejada, Aída. 2022. Direct object anaphora resolution in L1 English-L2 Spanish referring clitics and DPs. *Revista Española de Lingüística Aplicada* 37(2):591-627. DOI: 10.1075/resla.22014.gar
- Hendriks, Petra. 2016. Cognitive modeling of individual variation in reference production and comprehension. *Frontiers in Psychology* 7:506. DOI: 10.3389/fpsyg.2016.00506
- Hirose, Yukio. 2014. The conceptual basis for reflexive constructions in Japanese. *Journal of Pragmatics* 68:99-116.

- Huang, Xingtao. 2009. "Tā" zì de wénhuà shǐ "她"字的文化史 [The Cultural History of the Chinese Character "ta"]. Fuzhou: Fujian Jiaoyu Chubanshe.
- Huang, Yan. 2000. *Anaphora: A Cross-linguistic Study*. Oxford and New York: Oxford University Press.
- Kamp, Hans, Genabith, Josef van, and Uwe Reyle. 2011. Discourse Representation Theory. In *Handbook of Philosophical Logic* (2nd ed.), eds. D. M. Gabbay, and F. Guenther, pp.125-394. Cham: Springer.
- Kim, Jaegu. 2009. *Descriptive Study of Korean E-mail Discourse*. Louisiana: Louisiana State University dissertation.
- Kim, Kyeong-min, and Chung-hye Han. 2016. Inter-speaker variation in Korean pronouns. In *The Impact of Pronominal Form on Interpretation*, eds. Patrick Grosz, and Pritty Patel-Grosz, pp.349-373. Berlin: De Gruyter Mouton.
- Kogusuri, Tetsuya. 2019. On the emphatic reflexive in Japanese: With special reference to zibun-de. *Gengo Bunka Kyodo Kenkyu Purojekuto: Nintikino Gengogaku* [Language and Culture Joint Research Project: Cognitive-functional Linguistic Research] 4:31-40.
- Lee, Kum-Young. 2008. The role of pragmatics in reflexive interpretation by Korean learners of English. In *Selected Proceedings of the 2007 Second Language Research Forum*, ed. by M. Bowles, pp.97-112. Somerville: Cascadia Proceedings Project.
- Levinson, Stephen. C. 2000. *Presumptive Meaning*. Cambridge: MIT Press.
- Lewis, Richard. L., and Shravan Vasishth. 2005. An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, 29(3):375-419.
- McCormack, B. G. 2003. *The Maintenance and Loss of Reflexive Anaphors in L2 English*. Durham: Durham University dissertation, United Kingdom.
- O'Grady, William. 2005. *Syntactic Carpentry: An Emergentist Approach to Syntax*. Mahwah: Lawrence Erlbaum.
- O'Grady, William. 2022. *Natural Syntax – An Emergentist Primer* (3rd ed.). Unpublished manuscript. <https://manoa.hawaii.edu/linguisticswp-content/uploads/Natural-Syntax-%e2%80%93-An-Emergentist-Primer-3rd-ed.-2022.pdf>
- Oh, Sun-Young. 2005. English zero anaphora as an interactional resource. *Research on Language and Social Interaction* 38(3):267-302. DOI: 10.1207/s15327973rlsi3803_3
- Oh, Sun-Young. 2006. English zero anaphora as an interactional resource II. *Discourse Studies*, 8(6):817-846. DOI: 10.1177/1461445606067332
- Oshima, David Y. 2007. On empathic and logophoric binding. *Research on Language & Computation* 5:19-35.
- Pan, Haihua. 1995. *Locality, Self-ascription, Discourse Prominence, and Mandarin Reflexives*. Austin: The University of Texas dissertation.
- Parker, Dan. 2019. Cue combinatorics in memory retrieval for anaphora. *Cognitive Science* 43(3):e12715. DOI: 10.1111/cogs.12715
- Parker, Dan, and Collin Phillips. 2017. Reflexive attraction in comprehension is selective. *Journal of Memory and Language* 94:272-290.
- Rankin, Tom. 2017. The distribution of reflexive intensifiers in learner English. *International Journal of Learner Corpus Research* 3(1):36-60.

Darcy Sperlich

- Reuland, Eric. 2011. *Anaphora and Language Design*. Cambridge: MIT Press.
- Reuland, Eric. 2017. Why is reflexivity so special? Understanding the world of reflexives. *Studia Linguistica* 71(1-2):12-59.
- Rothman, Jason. 2011. L3 syntactic transfer selectivity and typological determinacy: the typological primacy model. *Second Language Research* 27(1):107-127.
- Schmid, Monika S., and Teodora Mehotcheva. 2012. Foreign language attrition. *Dutch Journal of Applied Linguistics* 1(1):102-124. DOI: <https://doi.org/10.1075/dujal.1.1.08sch>
- Schwartz, Bonnie D., and Rex A. Sprouse. 1996. L2 cognitive states and the full transfer/full access model. *Second Language Research* 12:40-72.
- Slabakova, Roumyana, White, Lydia, and Natália Brambatti Guzzo. 2017. Pronoun interpretation in the second language: effects of computational complexity. *Frontiers in Psychology* 8:1236. DOI: 10.3389/fpsyg.2017.01236
- Sorace, Antonella, and Francesca Filiaci. 2006. Anaphora resolution in near-native speakers of Italian. *Second Language Research* 22(3):339-368.
- Sperlich, Darcy. 2013. *The Acquisition of Long-distance Reflexives in Chinese as an Interlanguage: An Experimental Study*. Auckland: The University of Auckland dissertation, New Zealand.
- Sperlich, Darcy. 2017. The instantiation of binding through pragmatic and syntactic processes. In *Key Issues in Chinese as a Second Language*, eds. Istvan Kecskes, and Chaofen Sun, pp.188-213. New York: Routledge.
- Sperlich, Darcy. 2020. *Reflexive Pronouns: A Theoretical and Experimental Synthesis*. Cham: Springer.
- Sperlich, Darcy, Leem, Jiaho, and Eui-Jeen Ahn. 2021. Explicit and implicit dimensions of sociopragmatic and pragmalinguistic competence. *International Review of Applied Linguistics in Language Teaching* 59(3):369-394. DOI: 10.1515/iral-2018-0319
- Stern, Nancy. 2019. Ourselves and themselves: grammar as expressive choice. *Lingua* 226:35-52.
- Wang, Yixin. 2022. Politeness and pragmatic transfer in L2 pronominal address usage: variation of T/V pronouns in Spanish by Chinese learners. *Spanish in Context* 19(1):25-47. DOI: 10.1075/sic.19018.wan

Darcy Sperlich
Department of Applied Linguistics
Xi'an Jiaotong – Liverpool University
Suzhou, China
Darcy.sperlich@xjtlu.edu.cn

二語習得教學與研究中的回指：理論與實踐

Darcy Sperlich

西交利物浦大學

本章旨在概述語言學和二語習得理論中關於回指現象的研究，並探討這些理論在二語課堂中的實際應用。主要關注的語言是中文和英文，但在相關情況下也會涉及其他語言。文章首先回顧了在解釋回指模式時發現的句法、語義和語用理論方法，並指出單一的回指方法過於局限，最好根據不同語言的特點，綜合考慮影響回指的多個因素。接著，文章將重點轉向二語課堂，討論了中國英語學習者在兩種回指系統（從一語到二語）之間轉換時可能遇到的問題，特別指出反身代詞可能是最棘手的問題。此外還討論了*tu/vous*系統中涉及禮貌的方面，以及語言學習者在使用該系統時可能面臨的障礙。隨後文章進一步探討了教學方面的問題。最後，本文探討了該領域未來的前景，並對未來的研究提出了建議。

關鍵字：擬聲詞、二語、句法、語用學、禮貌

VARIATIONS IN SYNTACTIC STRUCTURES UNDERLYING APPARENT *TOUGH* CONSTRUCTIONS IN ENGLISH*

Shigenori Wakabayashi and Yuki Takahashi
Chuo University

ABSTRACT

Examining the grammar of second language (L2) learners, this paper focuses on the interpretation of *tough* constructions (TCs) in English and the acceptability of object drops and pseudo-*tough* constructions by Japanese-speaking learners. Departing from prior studies, our investigation revealed that the grammar of L2 learners exhibits more variations than previously suggested. We experimented with a truth-value judgment task and an acceptability judgment task. The results suggest that L2 learners' interpretation of TCs in English relates neither to the acceptance and interpretation of object drops nor the acceptability of the judgment of pseudo-TCs. They also suggest that some learners adopt one set of operations for structuring apparent TCs while others employ an alternative approach. Beyond that, learners in our study responded to each type of sentence consistently, and their patterns showed that Universal Grammar constrained their L2 grammar. Therefore, we propose that Universal Grammar operates differently between first language and L2 acquisition. In L2 grammar, the acquisition of lexical items, including features and sets of operations, is affected by the learner's grammar at the point of development, and principles of economy specific to L2 grammar are likely to operate.

Keywords: *tough* construction, English, Japanese, Universal Grammar, variation

* We thank two anonymous reviewers for their valuable comments. Special thanks go to Stano Kong for supporting the whole editing process, including finalizing the abstract in Chinese, and John Archibald for collecting data. We also thank everyone who participated in the experiment. Our research was supported by a Chuo University Personal Research Grant and Chuo University Grant for Special Research awarded to the first author.

1. INTRODUCTION

In this paper, we discuss characteristics of the grammar of second language (L2) learners, with particular focus on Japanese-speaking learners' understanding of *tough* constructions (TCs) in English. Ultimately, we reveal that the grammar of L2 learners may exhibit more variations than suggested by previous studies (e.g., Kim and Schwartz 2022). We also suggest that even when learners' behavior is target-like, their grammar may differ from that of native speakers, just as the operation of Universal Grammar (UG) in L2 acquisition differs from that in first language (L1) acquisition.

TCs are interesting to researchers who study L1 and L2 acquisition because they contain a so-called Plato's problem, also called logical problem in language acquisition or a "poverty-of-stimulus situation." Noam Chomsky's (1964:66–67) famous examples clearly show that the syntactic structures underlying the two similar sentences in (1) are different:

- (1) a. John is easy to please.
b. John is eager to please.

Although the sentences in (1) are identical in terms of the linear order of categorized constituents (i.e., a subject DP *John*, the "be" verb *is*, an adjective, and the infinitive verb *to please*), *John* is the patient or theme but also the agent or causer of *please* in (1a) and (1b), respectively. In English, sentences such as (1a) are called "TCs," which are known to be difficult to learn in L1 acquisition. Carol Chomsky (1969) and others have additionally found that children less than 5 years old experience problems with interpreting TCs in adult-like ways.

In L2 acquisition, the same Plato's problem applies. Some learners interpret sentences such as (1a) in a target-like way, but others do not. Based on experimental data, Kim and Schwartz (2022) have suggested that a learner's grammar was identical to a native speaker's grammar.

Although Kim and Schwartz's (2022) data are relevant, their argument contains some problems. The authors seem to consider that L2 learners have one L2 grammar at a certain stage of development as long as they share an L1. We argue, however, that such may not be the case.

We experimented with Japanese-speaking learners of English (JLEs) and found many patterns in their behavior concerning TCs and related phenomena, including object drops and pseudo-TCs (see Section 2.1). Some learners may adopt one set of operations to structure apparent TCs, whereas others employ another. Although UG constrains both L1 and L2 grammar, its operation may differ between those two types of language acquisition and thereby result in more variations in L2 grammar than L1 grammar.

In what follows, we present native speakers' grammar for TCs in English and its (near) equivalent in Japanese in Section 2, after which we review several studies concerning the acquisition of TCs in Section 3. In Section 4, we provide a theoretical description of object drops in Japanese in order to describe L2 grammar in greater detail. Next, we describe our research question and experiment in Sections 5 and 6, respectively, and we discuss the results and their implications for L2 acquisition in Section 7. We conclude the article in Section 8.

2. LINGUISTIC FACTS

This section describes the derivation of English TCs and Japanese equivalents.

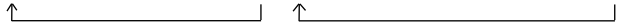
2.1 TCs in English

In English, TCs contain adjectives that denote difficulty or ease and precede the infinitival clause.¹ In the English TC in (2), the matrix subject *John* has to be interpreted as the patient or theme of *catch*, not the agent or actor. Therefore, its truth condition is the same as the one in (2b):

¹ Sentences with adjectives denoting emotions (e.g., *agreeable*, *annoying*, *boring*, *comfortable*, *exciting*, *irritating*, *joyful*, and *(un)pleasant*); value judgments such as *(in)appropriate*, *bad*, *beneficial*, *(in)convenient*, *good*, *(un)healthy*, *nice*, and *useful*; and some nouns (e.g., *a bitch*, *a breeze*, *a piece of cake*, *a snap*, and *a cinch*) in the place of *easy* in (1a) are also regarded as TCs in general (cf. Chiba 2019). This paper does not address those predicates because Japanese has no similar or equivalent structures. However, regardless of whether and, if so, how JLEs acquire sentences with those adjectives, our proposals in this paper require no change.

- (2) a. John is easy to catch.
b. It is easy to catch John.

The questions asking what syntactic structures underly TCs are ones that have attracted numerous linguists (e.g., N. Chomsky 1977, 1982; Lasnik and Fiengo 1974; Postal 1971). On that count, we follow Hicks (2009), who has proposed that A'-movement and A-movement occur in sequence, as shown in (13): the noun phrase [*Op [John]*] merges with the verb *catch* as its complement and undergoes A'-movement to Spec-CP, after which *John* takes the subject position via A-movement. Such movement is called “smuggling” (Collins 2005a, 2005b).

- (3) John_j is easy [CP [*Op [John_j]]*]_i *PRO* to catch [*Op [John_j]]*]_i.


(based on Hicks 2009:547)

Because *catch* is a mono-transitive verb, the sentence becomes ungrammatical if another noun phrase follows it, as in (4a). Beyond that, because A'-movement is involved, unaccusative or unergative verbs cannot be used, as in (4b) and (4c):

- (4) a. *John is easy to catch Mike.
b. *John is hard to jump. (unergative)
c. *Taro is easy to arrive. (unaccusative)

Yip (1995) has dubbed those sentence “pseudo-TCs,” a term that we also adopt in this paper.

2.2 Tough Constructions in Japanese

Although some theoretical linguists have argued that Japanese has no construction equivalent to English TCs, others disagree. From the perspective of L2 acquisition, it is important to note that Japanese has constructions that L2 learners are likely to interpret as translations equivalent to English TCs. The adjectival items used in such sentences are

twofold—namely, *-yasu(i)* “easy” and *-niku(i)* “hard/difficult/tough” positioned after the infinitival form of a verb. We call those types of sentences “TCs in Japanese” (cf. Inoue 1978, 2004).

Inoue (1978, 2004) has classified TCs in Japanese into four types, among which Type 1 is arguably equivalent to TCs in English.² In (5a), the subject noun phrase *kono zisyo*, “this dictionary,” with a nominative case *-ga* is interpreted as the theme of the verb *tukai*, meaning “to use.” Sentence (5b) is the translation equivalent to the English sentence (1a).

- (5) a. kono zisyo-ga tukai-yasu-i
this dictionary-NOM use-easy-PRES
“This dictionary is easy to use.”

b. John-ga yorokobase-yasu-i
John-NOM please-easy-PRES
“John is easy to please.”
- (based on Inoue 2004:76)

In Type 2, the subject is interpreted as an agent of the infinitival verb. In (6a), the second conjunction has an empty category *ec* subject, interpreted as the agent or actor of the unergative verb *aruki*, meaning “to walk.” The antecedent of *ec* is *watasi*, “I”—that is, the subject of the first conjunction. That sentence can be considered to be syntactically equivalent to the English pseudo-TC in (4b):³

- (6)⁴ Watasi_i-wa kega-o site-ite, *ec_i* aruki-niku-i
 I-TOP injury-ACC do-and *ec* walk-hard-PRES
Lit. “*I have an injury, and I am hard to walk.”
 “I have an injury, and it is hard for me to walk.”
 (based on Inoue 2004:77)

² See also Kaneko (1994), Kuroda (1987), Ohkado (1993), and Takezawa (1987).

³ The reason for using *watasi* but not *John* here is that Japanese experiential/psychological adjectives, such as *-nikui* “hard” and *-hosii* “wanting,” can be used only with the first person subject. If we replace *watasi* with *John*, the sentence (*John wa/ga aruki nikui*) is degraded although some Japanese accept such an expression. In order to avoid the degradation, a raising verb (e.g., *yooda* “appear”) needs to be used (*John wa aruki nikui yooda*. Lit. “John appears to be hard to walk”).

⁴ In this article, an empty category is presented as *ec* in the places where whether it is *pro*, an elided item, or a copy left after movement does not have to be discussed.

Type 3 is a passive TC and an unaccusative TC. In (6a), the Japanese passive morpheme (*r*)*are* is used, and the subject is the theme or patient of *damas(u)*, “deceive.” In (7b), an unaccusative verb *toke(ru)*, “dissolve,” is used, and the subject plays the role of theme:

- (7) a. Roojin-ga damas-are-yasu-i
elderly people-NOM deceive-PASS-easy-PRES
Lit. “*Elderly people are easy to be deceived.”
“Elderly people are easily deceived.”
- (based on Inoue 2004:104)
- b. Senzai-ga yu-ni toke-yasu-i
detergent-NOM warm-water-in dissolve-easy-PRES
Lit. “*Detergent is easy to dissolve in warm water.”
“Detergent dissolves easily in warm water.”
- (based on Inoue 2004:82)

Last, in Type 3, the adjective *-yasu(i)* means “tend to” and *-nikui* means “tend not to.” The subject NP marked with the nominative case marker is the agent or actor of the infinitive verb, and the object NP marked with an accusative case marker is the theme.

- (8) a. Awatemono-ga ziko-o okosi-yasu-i
hasty-people-NOM accident-ACC cause-easy-PRES
Lit “*Hasty people are easy to cause accidents.”
“Hasty people tend to cause accidents.”
(based on Inoue 2004:85)

In Type 1, 2, and 3 TCs, any change-of-state or activity verbs may be used in Japanese TCs, and their thematic relations with sentential subjects and objects differ from those in English TCs.⁵

⁵ An anonymous reviewer raised a question whether the range of the types of verbs that may appear in Japanese TCs may influence JLE's acquisition of English TCs. Although the underlying structures of Japanese TCs clearly influence JLEs' use and acceptance of English TCs, as shown below, it is unclear whether the range of the types of verbs may affect SLA. We leave that question for further studies.

Several analyses of syntactic structures underlying Type 1 Japanese TCs have proposed different base-generate positions of the subject noun phrase and movement in the derivation. In this paper, we follow the analysis in which the nominative noun phrase is base-generated and stays in the complement of the infinitival verb (Inoue 2004; Taguchi and Niimura 2009; Kishimoto 2017), because it describes the four types of Japanese TCs in a principled way. According to that analysis, the Type 1 sentence in (9a = 5a), has the structure in (9b), in which the theme—the noun phrase *kono zisyo* (-ga)—merges with the infinitival verb *tukai*. Meanwhile, the agent *PRO* merges with the same constituent, thereby resulting in a Verb Phrase (VP), which merges with the predicative adjective *yasu(i)*, meaning “easy,” as its complement and constitutes a VP. A *pro*, or lexical noun phrase, thus merges with the rest of the constituent in the subject position.⁶

- (9) a. kono zisyo-ga tukai-yasu-i
 this book-NOM use-easy-PRES
 b. *pro* [_{VP} *PRO* kono zisyo-ga tukai] yasu-i
 pro PRO this book-NOM use -easy-PRES
 ‘This dictionary is easy to use.’
(based on Taguchi and Niimura 2009:1876)

The sentence (9a) has non-canonical case markings—that is, the complement of the infinitive verb is marked with a nominative case *ga*, for nominative phrases appear to be the complements of stative predicates in Japanese (Koizumi 2008; Kuno 1973; Shibatani 1999),⁷ and the *tough* adjectives *-yasu(i)/-niku(i)* are stative predicates. Therefore, Type 1 TCs in Japanese, in which a noun phrase with a nominative case is interpreted

⁶ The rationale of the non-overt subject *pro* being the external argument and the DP with a nominative case being the internal argument stems from several syntactic tests, such as the interpretation of a reflexive anaphor *zibun*, the availability of honorific agreement, and resultative predicates. See Taguchi and Niimura (2009) and Nagamori (2017).

7 Similar constructions to the nominative object construction are found in Hindi, Icelandic, Kannada, Korean, Russian, and Tibetan, among others (Koizumi 2008), which indicates that it is not exceptionally rare or marked that an object is assigned a nominative case.

as the theme of the infinitival verb, differ in syntactic structure from English TCs.

In Type 2 TCs, the nominative noun phrase is the agent of the infinitival verb.⁸ The syntactic structure of that type of sentence corresponds to the English control sentence in (10b), in which *PRO* is the agent of the infinitival verb and is controlled by the subject in the matrix clause (cf. Ohkado 1993):

- (10) a. Watasi_i-wa kega-o site-ite, *pro*_i [*PRO*_i aruki]-niku-i
 I -TOP injury-ACC get-and *pro* *PRO* walk -hard-RES
 (Lit.) “*I get an injury, and I am hard to walk.”
 “I am injured, and it is hard for me to walk.”
 b. Taro_i is eager [*PRO*_i to walk.]

We do not discuss the structures of Types 3 and 4 in this paper because the discussion in this subsection fulfills our paper’s purpose of showing that the types of Japanese TCs have different structures. Because different types of Japanese TCs have different structures, one sentence may exhibit ambiguity. The sentence (11a) may have a Type 1 structure, as in (11b), or a Type 2 structure, as in (11c):

- (11) a. Taro-ga mitsuke yasu-i
 Taro-NOM find easy-PRES
 b. *pro* [*PRO* [VP Taro-ga mitsuke]] yasu-i
 “Taro is easy to find.”
 c. Taro-ga [*PRO* [VP *pro* mitsuke]] yasu-i
 (Lit.) “*Taro is easy to find *pro*.”
 “It is easy for Taro to find someone.”

In sentence (11a), *Taro* may mean the theme or patient or else the agent of the infinitival verb *mitsuke* “find.” In the latter case, *pro* can be replaced by an overt noun phrase as in (12). That possibility shows that a gap in the complement position of the infinitival verb in a TC is not required in Japanese TCs.

⁸ Interested readers are referred to Inoue (2004).

- (12) Taro-ga kakureteiru ko-o mitsuke yasu-i
Taro-NOM hiding child(ren)-ACC find easy-PRES
“It is easy for Taro to find the hiders / Taro can easily find the hiders.”

In short, similar to English TCs, a Japanese TC may have a noun phrase, an adjectival predicate, and an infinitival verb phrase and can be interpreted in ways similar to English TCs. However, the syntactic structures in those two TCs are distinct. Japanese TCs are likely to involve neither an A'- nor an A-movement, whereas English TCs involve both.⁹

In this subsection, we have described the syntactic structures of different Japanese TCs. Because Japanese allows null arguments and has case assignments unlike those in English, even Type 1 TCs, although superficially equivalent to English TCs, have a structure different from English TCs. In the next section, we review previous studies on the L2 acquisition of English TCs by native Japanese and Korean speakers.

3. PREVIOUS STUDIES ON L2 ACQUISITION

In this section, we review three important studies on the acquisition of English TCs by Japanese- and Korean-speaking learners of English.

Yoshimura et al. (2017) have investigated how JLEs interpret sentences with English TCs containing a prepositional phrase (PP)—for instance, *for Betty* in (13). In their study, participating JLEs in the lower group ($n = 28$; mean TOEIC score: 443.36) and higher group ($n = 25$; mean TOEIC score: 732.92) were asked to read target sentences written in English and answer a multiple-choice comprehension question in Japanese after reading each sentence. The following provides an example:

⁹ Movement analyses for Japanese TCs have been proposed and discussed. For example, Nakagawa (1997) argues that the empty category at the complement of a verb is a null resumptive pronoun in some syntactic conditions but the trace of an empty operator in others. We do not discuss this further because it does not affect our main proposal in this paper.

(13) Elizabeth is always difficult for Betty to please on her birthday.

Question

Dare-ga tanzyoobi-ni yorokobaseru -no-ga muzukasii
Who-NOM birthday-on please-nominalizer-NOM difficult
desu ka
copula Q
“Who is difficult to please on her birthday?”

Answer options

A: Elizabeth B: Betty C: both D: I don’t know
(Yoshimura et al. 2017:249)

The mean percentages of correct responses were 73% for the lower group and 92% for the higher group. Yoshimura et al. (2017) attributed the native like performance of both groups to positive L1 transfer: A particular type of Japanese TCs are similar to English TCs in their word order and interpretation, even though both types of TCs are distinct in their underlying syntactic structures, as shown in the previous section. The apparent similarity on the surface may have caused the participating JLEs to interpret the English TCs in target-like ways, which implies that JLEs’ grammar needed for the correct interpretation differs from the grammar of native speakers. We should also note that the English sentence with the TC in the experiment included an agent PP—that is, *for Betty* in (13)—which likely led participants to abandon the interpretation of the subject noun phrase *Elizabeth* in (13) as the agent of the infinitival verb *yorokob-aseru*, meaning “please,” in (13). Thus, the results of the experiment may have obscured the possibility that JLEs understood the subject noun phrase as the agent of the infinitival verb if there was no such PP, in a reflection of their L1 grammar.¹⁰ Most important, even if participating JLEs understood that the subject was the agent in the target sentence, they may have chosen the appropriate answer *Elizabeth* because the Japanese question is ambiguous: The

¹⁰ Why Yoshimura et al. (2017) introduced a PP in their experiment is not clearly mentioned.

subject *dare*, meaning “who,” can be interpreted as either the agent or patient, as described in (11) in the previous section.

In another experiment using sentences containing TCs without agent PPs, Okuma (2018) administered a truth-value judgment task to two groups of JLEs: an intermediate group ($n = 14$; mean TOEIC score: 560) and an elementary group ($n = 19$; mean TOEIC score: 404).¹¹ The participants read some context and a sentence with a TC and subsequently judged whether the context matched the sentence. Four conditions were set with two parameters: (i) the transitivity of infinitival verbs (i.e., obligatory transitive such as *beat* vs. intransitive-biased optionally transitive such as *fight*) and (ii) sentence–context matching (i.e., match vs. mismatch). An example of some context and the corresponding sentences to be judged appear in (14):

(14) Context: Lisa and Sam love to box. Lisa has special gloves that detect the other boxer’s movements and make her give punches, so she always wins. However, today Lisa lent the special gloves to Sam and is wearing normal gloves, so she cannot punch him as she used to. In contrast, Sam can give punches much more quickly and accurately than Lisa.

- a. Sam is easy to beat today. True False I don’t know
b. Lisa is easy to beat today. True False I don’t know
(Okuma 2018:20)

Sentence (14a) mismatches the context, while (14b) matches it. Although the results showed that native speakers of English (NSEs) performed as expected by theory, both JLE groups tended to interpret subjects as agents or actors, not as patients or themes. Both groups chose response option “True” in the mismatch conditions, as in (14a), but “False” in the match conditions, as in (14b), regardless of the transitivity of the verbs. The analysis of individual participants revealed that only four learners surpassed the level of pure chance (i.e., 50%). Given those

¹¹ Okuma (2018) is an advancement of Kim (2014), the relevant part of which is published as Kim and Schwartz (2022), which we review below.

results, Okuma (2018) concluded that JLEs have difficulties with interpreting English TCs in targeted ways.

More recently, Kim and Schwartz (2022) examined Korean-speaking learners of English (KLEs) and suggested that some KLEs acquired target-like grammar.¹² To be sure, Korean is similar to Japanese in many morphosyntactic aspects. For example, (15a) is ambiguous because it contains two possible underlying syntactic structures as in (15b) and (15c), similar to (11) in Japanese:

- (15)¹³ a. Kyoswunim-i manna-ki-ka elyep-ta
 professor-NOM meet-NML-NOM difficult-DECL
 (based on Kim and Schwartz 2022:263)
 b. *pro* [*PRO* [_{VP} Kyoswunim-i manna-ki-ka] elyep-ta
 “The professor is difficult to meet.”
 c. Kyoswunim-i [*PRO* [_{VP} *pro* manna-ki-ka]] elyep-ta
 (Lit.) “*The professor is difficult to meet *pro*.”
 “It is difficult for the professor to meet someone.”

Attributing the difference between English TCs and Korean TCs to the availability of a null object *pro*, Kim and Schwartz (2022) assumed that KLEs likely acquired native-like grammar if they had unlearned *pro* as a means to disallow structures used in Korean.

In the same research, Kim and Schwartz (2022) conducted an experiment with two types of truth-value judgment tasks (TVJT) and an acceptability judgment task (AJT) with three groups of KLEs: low ($n = 13$), medium ($n = 18$), and high ($n = 18$). Added to those KLEs, 10 native Korean speakers participated in the Korean TVJT (KTVJT). Ten native English speakers also participated in the English TVJT (ETVJT) and the English AJT.

¹² Adult and child KLEs were involved in the experiment. Only the data from adults are discussed here.

¹³ The subjecthood and objecthood of noun phrases are illustrated with honorifics in Kim and Schwartz (2022) as discussed in theoretical linguistics (see references cited there). We do not describe such properties here to keep the discussion concise. To mention that Korean and Japanese are similar to each other, including TCs, is sufficient for our argument in this paper.

In TVJTs, participants were asked to read sentences that provided context that were accompanied with pictures and to judge whether a target sentence matched the context. Two conditions were manipulated in the TVJT: one concerning the verb’s transitivity and the other concerning the context’s bias. For the verb’s transitivity, obligatorily transitive verbs (i.e., *beat*, *catch*, *find*, and *touch*) and intransitively biased, optionally transitive verbs (i.e., *fight*, *hunt*, *move*, and *race*) were used.¹⁴ The authors predicted that the latter type of verbs might result in non-target-like reading, such that the sentential subject might be interpreted as an agent because no empty category was required after them. The context’s bias was manipulated by designing two contexts: neutral and subject-biased. Figure 1 presents the neutral context, in which Bert is playing tag and wearing roller skates. The roller skates help Bert regardless of whether he is the catcher or the target. If he is the catcher, then he can move faster than usual and catch Elmo easily. Meanwhile, if he is the target, then he can move faster than usual and flee from Elmo, which makes it difficult for Elmo to catch Bert, and the expected response is “No.”

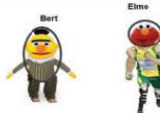
	Spoken stimuli	Visual stimuli
Context	Elmo is the fastest in the class. He is better than Bert at playing tag because he can run fast. He can chase Bert and run away from him.	
	But today, Bert is wearing <u>roller skates</u> . He is very good at roller skating. When Bert and Elmo play tag, Bert is much faster than Elmo.	
Test sentence	The bear says, “Bert is easy to catch today.” (Target: NO)	
	The bear asks, “Why?” (Target: Bert has skates.)	

Figure 1. An example of a neutral context from a truth-value judgment task (Kim and Schwartz 2022:269)

¹⁴ The intransitively-biased, optionally transitive verbs are those with a transitivity bias of less than 50% (Connine et al. 1984; Gahl et al. 2004).

In Figure 2, for the subject-biased context, Bert is wearing Spider-Man gloves, which unlike the roller skates benefit Bert only if he is a catcher. Even if Bert is wearing the gloves, they do not affect his ability to flee from Elmo. In that sense, the type of context created a bias toward reading Bert as the agent (i.e., subject) and was therefore called a “subject-biased context.” If KLEs behaved in a target-like way regardless of those manipulations, then they may have unlearned Korean grammar, as long as full transfer can be assumed as suggested in Schwartz and Sprouse (1994, 1996). Meanwhile, TVJTs were presented in English (ETVJT) and Korean (KTVJT) by using Korean-translated equivalents to the materials in ETVJTs to compare KLEs’ patterns of interpretation between L1 Korean speakers and L2 English speakers.

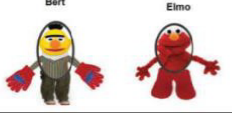
	Spoken stimuli	Visual stimuli
Context	Bert and Elmo like to play tag. Elmo has <u>spider-man gloves</u> . Elmo can hold anything easily when he is wearing the magic gloves. He is better than Bert at playing tag because his gloves make him good at grabbing hold of Bert.	
	Today, Bert has borrowed the magic gloves from Elmo. Bert and Elmo are playing tag. Today, Bert is better than Elmo at playing tag.	
Test sentence	The bear says, “Bert is easy to catch today.” (Target: NO)	
	The bear asks, “Why?” (Target: Before, Elmo used the gloves. Today, he doesn’t.)	

Figure 2. An example of subject-biased contexts
(Kim and Schwartz 2022:269)

Another task was an acceptability judgment task (AJT) asking participants to judge whether provided sentences sounded natural or not on a scale from 1 (*sounds completely unnatural*) to 6 (*sounds completely natural*). Two conditions were manipulated: one for verb transitivity, the

other for the presence of an object, as shown in (16). The eight verbs used in the ETVJTs were also used in the AJTs.

- (16) a. Kim is fighting Lee.
(intransitively-biased, optionally transitive + an object)
b. Kim is beating Lee. (obligatorily transitive + an object)
c. Kim is fighting ____.
(intransitively-biased, optionally transitive + no object)
d. *Kim is beating _____. (obligatorily transitive + no object)
(Kim 2014:216)¹⁵

In (16a) and (16b), the intransitively biased, optionally transitive verb *fight* and the obligatorily transitive verb *beat* were used with an object, respectively. In (16c) and (16d), meanwhile, the same verbs were used without an object, which resulted in ungrammaticality in (16d) because the obligatory transitive verb *beat* required an object. Therefore, the target response was to accept (16a), (16b), and (16c) and to reject (16d).

Regarding the results of the KTVJTs, native Korean speakers showed a strong bias toward the interpretation in which the subject noun phrase was the agent across all four conditions. KLEs more often upheld interpretations of a patient or theme than native Korean speakers, and that result may have exhibited a test effect, for KLEs completed the ETVJTs immediately prior to completing the KTVJTs.

The result of the ETVJTs showed that NSEs displayed 100% of expected interpretations regardless of the verb's transitivity or the types of context. By contrast, most KLEs more often upheld interpretations that identified an agent or actor. Even KLEs in the high group upheld interpretations of a patient or theme with obligatory transitive verbs only at a rate greater than chance. The main effects were observed among all factors (i.e., verb transitivity, type of context, and learner group). Moreover, the individual learners' data showed variability. Even in the high group, KLEs' performance varied, with one participant showing 100% target-like interpretations but another showing none.

¹⁵ Kim and Schwartz (2022) is based on Kim (2014), from which the experimental data are cited. Note that progressive form is used here. The progressive aspect with some punctual verbs such as *find*, *catch*, and *touch* may have degraded the sentences or made them sound unnatural without context.

Kim and Schwartz (2022) analyzed data regarding AJTs by comparing the acceptability of the sentences with obligatory transitive verbs in the intransitive framework (i.e., with no overt noun phrase following the verb), as in (17d). They found that native speakers rejected such sentences with a very low judgment based on a mean (M) rating of 1.3 in a range from 1 to 6. They reported that KLEs in the high group rejected them, with the data in English ($M = 2.3$), which was significantly lower than their judgment of Korean equivalents ($M = 3.8$).¹⁶ Based on those data, Kim and Schwartz suggested that KLEs had unlearned *pro*, which was assumed to have been transferred from their L1 (i.e., Korean).¹⁷

Because seven learners in the high group exhibited native-like performance on ETVJTs and AJTs,¹⁸ Kim and Schwartz (2022) suggested that they had acquired native-like grammar for English TCs, including *tough* movement. They suggested that sentences such as (17) might have triggered the acquisition of target-like grammar because preposition stranding shows that empty categories and movement are involved in TCs:

- (17) a. John is easy to talk to ____.
 b. Mary is hard to fight with ____.
 c. Bill is fun to laugh at ____.
- (Kim and Schwartz 2022:281)

Although Kim and Schwartz's (2022) contribution to our understanding of KLEs' behavior is significant, we also suggest that their argument about KLEs' grammar has a few major drawbacks. One is that whether KLEs behave similar to native speakers or not should be discussed based on the cross-examination of ETVJT-related data and AJT-related data. Although the authors present individual learners' data for the ETVJTs, they do not present such data for AJTs, which prevents understanding whether, for example, any participants who answered in native-like ways at a rate of 75% (i.e., 12/16) or higher on the ETVJTs

¹⁶ This is true to all proficiency groups although child Korean groups, which we do not discuss here, failed to reject the use of null objects.

¹⁷ The mean score in the medium score group and low score group are also below 3.

¹⁸ Kim and Schwartz (2022) set a criterion for a target-like performance: 12 times out of 16 trials (75%) are interpreted as target-like.

rejected null objects in target-like ways, whatever the criterion, on AJs. The more serious problem is that unlearning *pro* and interpreting English TCs in target-like ways do not guarantee that KLEs have native-like grammar. The acquisition of those two properties may be a necessary condition but is not a sufficient condition to conclude that they have native-like grammar. English TCs involve a complicated syntactic derivation involving both A'-movement and A-movement, as shown in Section 2. It is possible (and likely) that KLEs may have learned that, in English, an empty category is not allowed after a transitive verb and that a sentential subject in a TC is interpreted as a patient independently of each other, all without acquiring the native-like grammar.

In this paper, we verify those problems by analyzing experimental data. However, before describing our experiment, we need to further discuss null objects and TCs in Japanese.

4. NULL AND OVERT OBJECTS IN JAPANESE

So far, we have assumed, following Kim and Schwartz (2022), that the empty category at the complement position of transitive verbs is *pro*. However, it is widely accepted that an empty category at that complement may be a form of argument ellipsis (AE).¹⁹ Consider the sentences in (18). In (18a), *zibun-no-oziiisan* “self’s grandfather” in the first clause refers to William’s grandfather, while the one in the second clause refers to Tom’s grandfather. If an empty category appears in the second clause, as in (18b), then the sentence is grammatical, and the same reading (i.e., sloppy reading) is possible. Therefore, the empty category can be the result of an ellipsis.

¹⁹ Among other theoretical accounts, the most well-known one is the VP stranding verb phrase ellipsis (e.g. Otani and Whitman 1991), which we do not discuss here.

- (18) a. William-ga zibun-no oziisan-o tukamae-te,
 William-NOM self's grandfather-ACC catch-and
 Tom-mo zibun-no oziisan-o tukamae-ta.
 Tom-ADD self's grandfather-ACC] catch-PAST
 Lit. "William caught self's grandfather and Tom caught self's
 grandfather, too."
 "William_i caught his_i grandfather, and Tom_j caught his_j
 grandfather, too."
- b. William-ga zibun-no oziisan-o tukamae-te,
 William-NOM self's grandfather-ACC catch-and,
 Tom-mo *ec* tukamae-ta.
 Tom-ADD catch-PAST
 "William_i caught his_i grandfather, and Tom_j caught his_j
 grandfather, too."
 "William_i caught [his_i grandfather]_j, and Tom caught [his_i
 grandfather]_j, too."

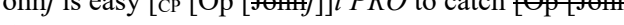
Sloppy reading is possible only in (18a), because *zibun* is bound by the subject noun phrase, *Tom*, in the preceding conjunction. By contrast, the empty category in (18b) may refer to William's grandfather, which constitutes a strict reading, as well as Tom's grandfather, which is a sloppy reading. When the sentence has strict reading, the empty category is a null pronoun *pro*.

Because Japanese allows both argument ellipsis and *pro*, JLEs may transfer one or both of those properties, thereby resulting in the acceptance of null objects in L2 English. However, if some JLEs' grammar is similar to that of native English speakers, as expected from Kim and Schwartz's (2022) discussion, then they should allow neither type of null object. Whether that expectation is correct can be tested, as done in Section 6.

Another issue concerns the relationship between the availability of an empty category and the target-like obligatory operator movement in English TCs. Kim and Schwartz (2022) investigated whether KLEs accepted null objects by examining their judgment of null objects with two types of transitive verbs—obligatorily transitive verbs (e.g., *beat*) and optionally transitive verbs (e.g., *fight*)—without any context. They

found that KLEs rejected sentences such as **Kim is catching* but not ones such as *Kim is fighting* and concluded that they had unlearned *pro*. The authors thus considered that when KLEs had acquired the target-like interpretation of English TCs and unlearned *pro*, they had attained native-like grammar in terms of using strings of words that make apparent English TCs. However, that logic may not be entirely correct.

As described in Section 2, a crucial difference between English, on the one hand, and Korean and Japanese, on the other, is that the former does not allow sentences such as (19a) because the noun phrase following the infinitival verb—for instance, *Mike* in (19a)—receives neither its theta role nor case, as shown in (19b). However, Korean and Japanese allow sentences such as (19c) because the predicate equivalent to the adjective *easy* in (19a), *-yasu(i)*, can be used in Japanese (and Korean) TCs, which are in fact control sentences, as shown in (19d):

- (19) a. *John is easy to catch Mike. (= 4a)
 b. *John_j is easy [_{CP} [_{Op} [John_j]]_i *PRO* to catch [_{Op} [John_j]]_i Mike].

 c. John-ga Mike-o tukamae yasui.
 John-NOM Mike-ACC catch easy-PRES
 “It is easy for John to catch Mike.”
 d. John-ga [*PRO* [_{VP} Mike-o tukamae]] yasui-i
 (Lit.) “*John is easy to catch Mike.”
 “It is easy for John to catch Mike. / John easily catches Mike.”

If JLEs acquire native-like grammar underlying an English TC, then they should not accept pseudo-TCs such as (19a). In our experiment, we examined whether that prediction is correct.

5. RESEARCH QUESTIONS

To investigate L2 learners' grammar regarding English TCs, we examined how they interpret sentences such as (20a) and whether they accept sentences with object drops such as (20b), which are ungrammatical in the grammar of native English speakers. Although Kim

and Schwartz (2022) conducted a similar experiment, the materials that they used for the latter question were different.

- (20) a. Sam is easy to beat today. (=14a)
b. *Kim is beating. (=16d)

Kim and Schwartz (2022) assumed that Korean learners of English who interpreted sentences with English TCs and rejected sentences with object drops had native-like grammar, which could be learned only with UG. However, such behavior among learners may be based on grammar unlike that of native speakers. As a case in point, L2 learners may construct non-target-like grammar that leads them to behave in native-like ways. To test that possibility, we scrutinized the materials to investigate L2 learners' grammar in its own light.

To that end, we formulated four questions. The first addressed how JLEs interpret English TCs—more precisely, whether they interpreted the subject of the sentence—for instance, *Sam* in (20a)—as the patient or theme or as the agent or actor of the infinitival verb—for instance, *beat* in (20a). The second question was whether JLEs allow object drops in sentences with an obligatory transitive verb—for instance, (20b)—and if so, then how they interpret the empty category following the verb. If they transfer their grammar from Japanese, then they may allow sloppy reading and strict reading in relevant contexts (see Section 4). If they allow only one type of reading, however, then their L2 grammar may differ from their L1 or native English grammar.

The third question was whether learners accept sentences with the pseudo-TCs such as (21):

- (21) *Edward is easy to buy expensive rings.
Intended meaning: Edward finds it easy to buy expensive rings.

If they do, then JLEs may not have a target-like grammar for English TCs. L2 learners with target-like grammar should reject that type of sentences because English TCs require the complement of the infinitival verb to be an empty category.

The last and most important question is whether JLEs behave like native speakers with all three items. If their L2 grammar is target-like,

then they should respond to all of the items in target-like ways—that is, they should interpret English TCs as native speakers do (i.e., interpret the subject of a sentence as the patient or theme of the infinitival verb), reject null object sentences, and reject pseudo-TCs. That combination is necessary to assume that JLEs have target-like grammar for English TCs.

6. EXPERIMENT

Our experiment contained three primary tasks: an English TVJT (ETVJT) similar to the ones in Kim and Schwartz (2022) and Okuma (2018); a Japanese TVJT (JTVJT), which is the translated equivalent of the ETVJT; and an AJT in English. The Oxford Quick Placement Test (OQPT; University of Cambridge Local Examinations Syndicate 2001) was also administered to L2 learners.

6.1 Participants

The participants were 26 NSEs and 27 JLEs. To explore JLEs' preference in interpreting Japanese TCs free from any effect from the ETVJT, we had 10 native speakers of Japanese complete the JTVJT. Table 1 summarizes the participants' characteristics.

Table 1. Participants' characteristics

	<i>n</i>	<i>Age</i>	<i>Sex (F/M)</i>	<i>Mean Age of First Exposure</i>	<i>Mean Duration of Learning</i>
NSEs (ETVJT)	13	32.38 <i>SD</i> =16.72	8/5	N/A	N/A
NSEs (AJT)	13	31.54 <i>SD</i> =14.65	10/3	N/A	N/A
JLEs	27	20.07 <i>SD</i> =1.41	6/21	9.78 <i>SD</i> =3.62	10.48 <i>SD</i> =3.65
NSJs (JTVJT)	10	21 <i>SD</i> =1.41	3/7	10.4 <i>SD</i> =2.84	10.5 <i>SD</i> =3.17

Based on the OQPT scale, JLEs were grouped into elementary (EL), lower intermediate (LI), or upper intermediate (UI).

Table 2. Proficiency groups

		<i>OQPT Score Range</i>	<i>Mean</i>	<i>SD</i>
JLEs	Elementary (EL) (n=5)	24-29	27.8	1.94
	Lower-Intermediate (LI) (n=11)	30-38	34.45	2.84
	Upper-Intermediate (UI) (n=11)	40-52	42.27	3.44

JLEs signed a consent form, answered a background questionnaire, and completed the ETVJT, AJT, JTVJT, and OQPT in that order. The order was set to minimize the JTVJT's effect on the English tasks (i.e., ETVJT and AJT) and to separate the two judgment tasks. Completing the primary tasks took 60–90 minutes, and the JLEs rested between tasks if desired. The experiment was conducted on Google Forms in a room with the second author serving as the experimenter, although the OQPT was conducted online. JLEs who completed all tasks received a gift card to Amazon.com worth JPY 2,000 as remuneration. Meanwhile, NSEs completed the task as volunteers without remuneration; we divided them into two groups so that the task burden was as light as possible.²⁰ One group ($n = 13$) took the ETVJT online, whereas the other ($n = 13$) took the AJT online. Figure 3 summarizes the order of tasks. Last, although we conducted our experiment as described, we do not discuss our JTVJT-related data in this paper because Japanese TCs are ambiguous, as described in Section 2.2.

²⁰ We thank John Archibald for his kind help with recruiting volunteers. We also thank everyone who participated in the experiment.

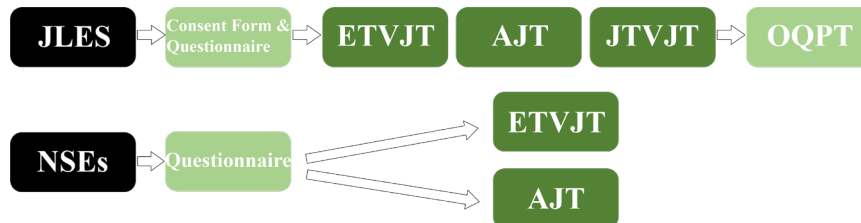


Figure 3. Order of the tasks

6.2 Materials

English Truth-Value Judgment Task (ETVJT). The ETVJT materials consisted of a story providing context, a picture providing context, the target sentence, and the answer options (i.e., “True” or “False”). The target sentence matched the context in the true condition but not in the false condition. The stories for context were presented in English. Examples appear in Figure 4 (i.e., true condition) and Figure 5 (i.e., false condition).

In Figure 4, the context states that *Today, Mary can move faster than Kathy thanks to roller skates*. The context implies that it is difficult for Kathy to catch Mary. Because the target sentence matches the context, the answer option “True” is correct. The context in Figure 5, by contrast, states that *Today, Elena can run faster than Hanako because Hanako is sick*. That context implies that it is not difficult for Elena to catch Hanako. Because the target sentence does not match the context, “False” is the correct answer option.

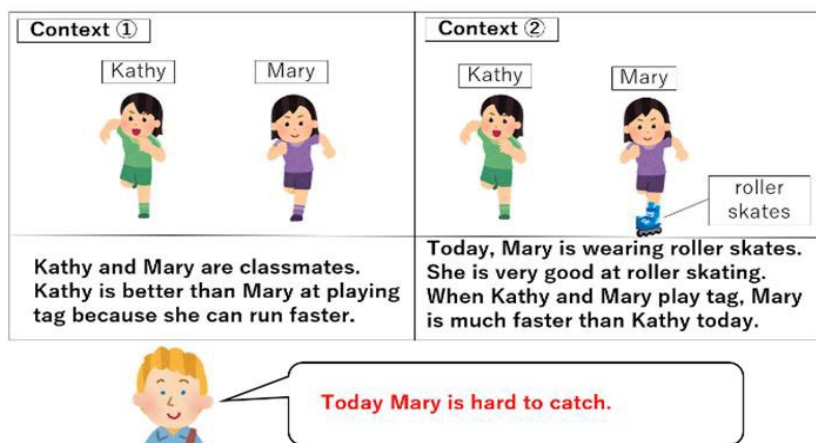


Figure 4. A true condition in the English truth-value judgment task

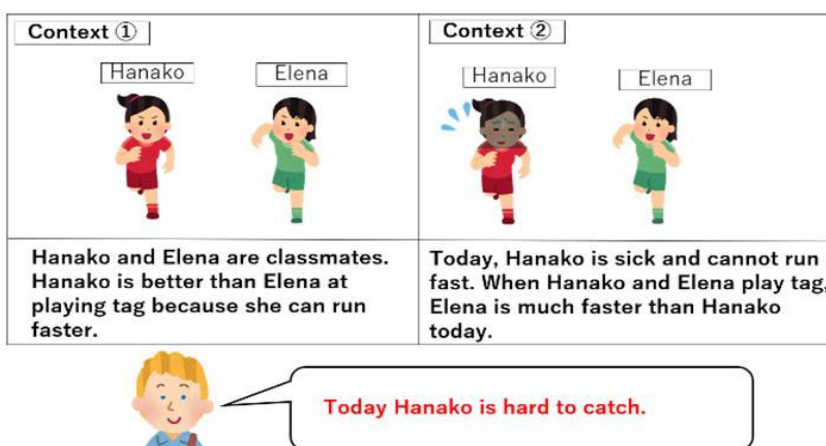


Figure 5. A false condition in the English truth-value judgment task

The verbs used in the ETVJT were the four obligatorily transitive verbs (i.e., *beat*, *catch*, *find*, and *touch*) used by Kim and Schwartz (2022), along with two more (i.e., *protect* and *lift*) for both the true and false conditions. As for adjectives, *easy* and *hard* were used, following Okuma (2018) and Kim and Schwartz (see footnote 1). The sentential subjects

were all animate because inanimate subjects may facilitate the interpretation of a theme (Ito 2018; Nguyen 2018). The numbers of words in the target sentences of paired conditions (i.e., true and false) were equal.

In the ETVJT, each type had six tokens, and 30 fillers were added, half of which were true and the rest false. Participants judged 42 tokens in total. The order that items were presented was pseudo-randomized so that sentences with the same type would not appear consecutively.

AJT. In the AJT, participants read a story for context, along with a picture for context, and judged whether the target sentence was acceptable for describing the situation on a 7-point scale from -3 to 3. Participants received instructions in their native languages so that they could judge the acceptability of the target sentence based on the target sentence's grammaticality, not the sentences or pictures for context, with grammatically correct and incorrect materials for examples. All materials had a story in black and white, a color picture, and a target sentence in red. The material included seven types. The first two types involved checking whether JLEs knew empty categories were unacceptable in English. Figure 6 is an example.

Context : When Ken, William, and Tom were playing tag, Ken and William touched Tom at the same time.

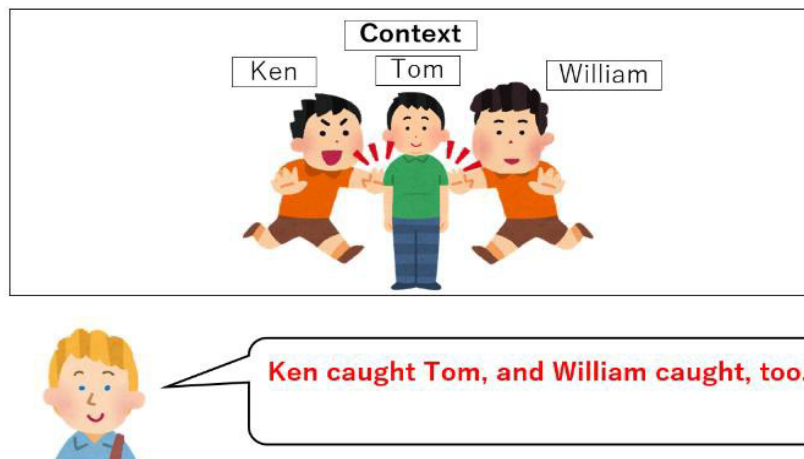


Figure 6. Type 2 (*strict *ec*) in the acceptability judgment task

- (22) a. Tom caught Ken, and Simon caught him, too. (Type 1: strict DP)²¹
 b. *Ken caught Tom, and William caught, too. (Type 2: strict *ec*)

In the picture, Tom and Simon caught Ken in the strict identity reading,²² so *pro* is assumed to occupy the object position of the second verb, *caught* (cf. Hoji 1985; Kuroda 1965; Ohso 1976).

We also used four types of sentences to investigate JLEs' grammar about the argument ellipsis. Two types are illustrated with the examples in (23).

- (23) a. Simon_i caught his grandfather, and Ken caught his grandfather,
 too.
 (Type 3: sloppy DP)
 b. *William caught his grandfather, and Tom caught, too.
 (Type 4: sloppy *ec*)

Sentence (23a) is grammatical, whereas (23b) is ungrammatical without an object noun phrase after the second verb, *caught*. To examine whether JLEs accept sentences such as (22b), we provided the material with a story for context, a picture for context, and the target sentence, as exemplified in Figure 7. In that context, William caught his grandfather, and Tom caught his. Argument ellipsis induces a sloppy reading of identity.

²¹ Here, DP is used instead of D to make the presentation parallel to Type 2 below. Theoretically, the pronoun *him* is a maximal projection of D.

²² Kim and Schwartz (2022) used progressive sentences in the AJT. The progressive aspect is not compatible with some punctual verbs such as *find*, *catch*, and *touch*, and this incompatibility may have led participants to reject the target sentences. Therefore, the sentences such as (22b) were used in this experiment.

Context : When William played tag with his own grandfather, William touched his grandfather. ^{*}
Tom also played it with his own grandfather and touched him.

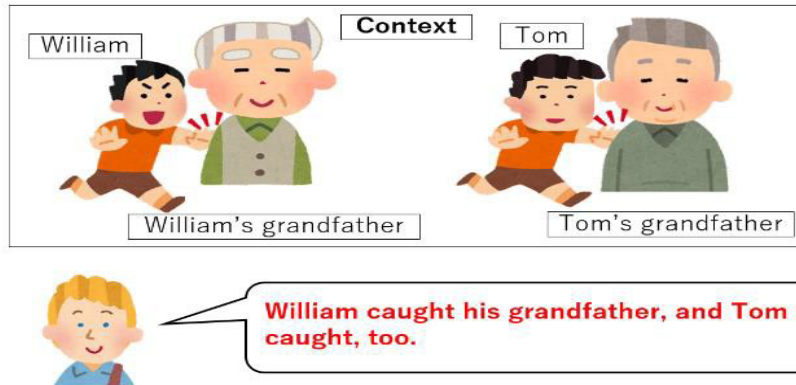


Figure 7. Type 4 (*sloppy *ec*) in the acceptability judgment task

Another type of sentence was used to investigate JLEs' knowledge of pseudo-TCs. An example appears in (24):

(24) *Edward is easy to buy expensive rings. (Type 5: *pseudo-TC)

In that type of sentence, six verbs (i.e., *read*, *cook*, *buy*, *touch*, *sing*, and *use*) and two adjectives (i.e., *easy* and *hard*) were used.

In the AJT, each type had six tokens, and 24 fillers were added; half were grammatical, whereas the rest were ungrammatical. Participants judged 54 tokens in total. The order in which the items were presented was pseudo-randomized so that sentences with the same type would not appear in consecutively.

6.3 Data Analysis and Results

No screening of the data was conducted. The data representing groups and the data representing individual learners are described and analyzed in Sections 6.4 and 6.5. Although the results of analyzing the data of the groups were similar to Kim and Schwartz's (2022), the data also seemed to obscure the grammar of individual learners. As shown in

Section 6.5, individual learners' data indicated that individuals should be divided into different groups even if they shared an L1 and had a similar proficiency in the L2.

6.4 Group Data and Individual Data in Each Group

In this subsection, we present the results of each task before examining relationships between them so that we can discuss the grammar of learners underlying their behaviors in the next section.

ETVJT. Figure 8 shows the ratios of the target-like interpretation of ETVJT. The left bars show the ratios of the target choice (i.e., “False”) for the target–context-mismatched items, whereas the right bars show the ones for the target–context-matched items.

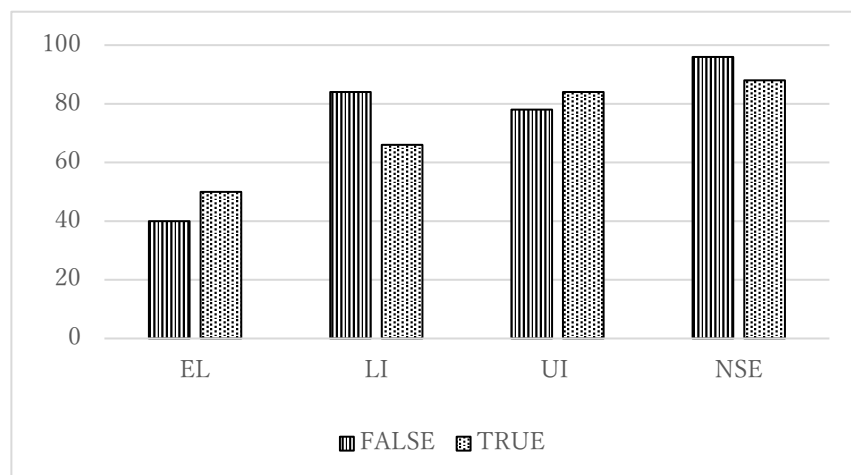


Figure 8. The ratios of target-like answers in ETVJT (%)

NSEs behaved as expected, with a very high rate of 92%. As proficiency increased, JLEs came to employ target-like interpretations of English TCs. When a Kruskal–Wallis test was conducted in R, the results showed a significant difference between the four groups, $\chi^2(3) = 13.864, p = .003$. Post hoc Wilcoxon rank-sum tests with Bonferroni correction indicated a significant difference between EL learners and NSEs ($p = .022$) and between LI learners and NSEs ($p = .022$). No significant differences were detected between UI learners and NSEs or between the groups of JLEs, probably due to the small number of learners in each group. Figure 9 shows the individual learners' data and that the number of target-like responses varied within each group of learners.

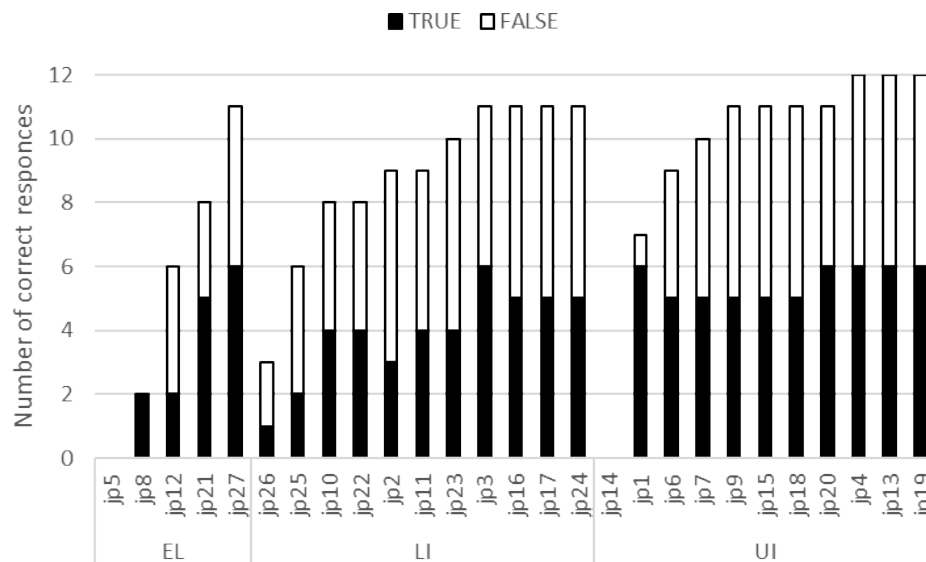


Figure 9. The number of target-like responses by JLEs

AJT: Object Drops & Pseudo-TCs. We collected data for the AJT using a 7-point scale ranging from -3 to 3, including 0. The results from NSEs and the three groups of JLEs appear in Figure 10, in which columns show the mean scores of each type. Type 1 (i.e., strict) sentences had an overt noun phrase and corresponded to the sentences in Type 2 (i.e., *strict *ec*; see Figure 6), which had an empty category to be co-indexed with the referent of the antecedent noun phrase. Types 3 and 4 corresponded with Types 1 and 2, respectively, whereas the former should be interpreted with sloppy reading (i.e., *sloppy AE; see Figure 7). Example sentences appear in (25):

- (25) a. Tom caught Ken, and Simon caught him, too. (Type 1: strict DP)
 (=22a)
 b. *Ken caught Tom, and William caught, too. (Type 2: *strict *ec*)
 (=22b)
 c. Simon_i caught his grandfather, and Ken caught his grandfather,
 too.
 (Type 3: sloppy DP) (=23a)
 d. *William caught his grandfather, and Tom caught, too.
 (Type 4: *sloppy *ec*) (=23b)
 e. *Edward is easy to buy expensive rings. (Type 5: *pseudo-TC)
 (=24)

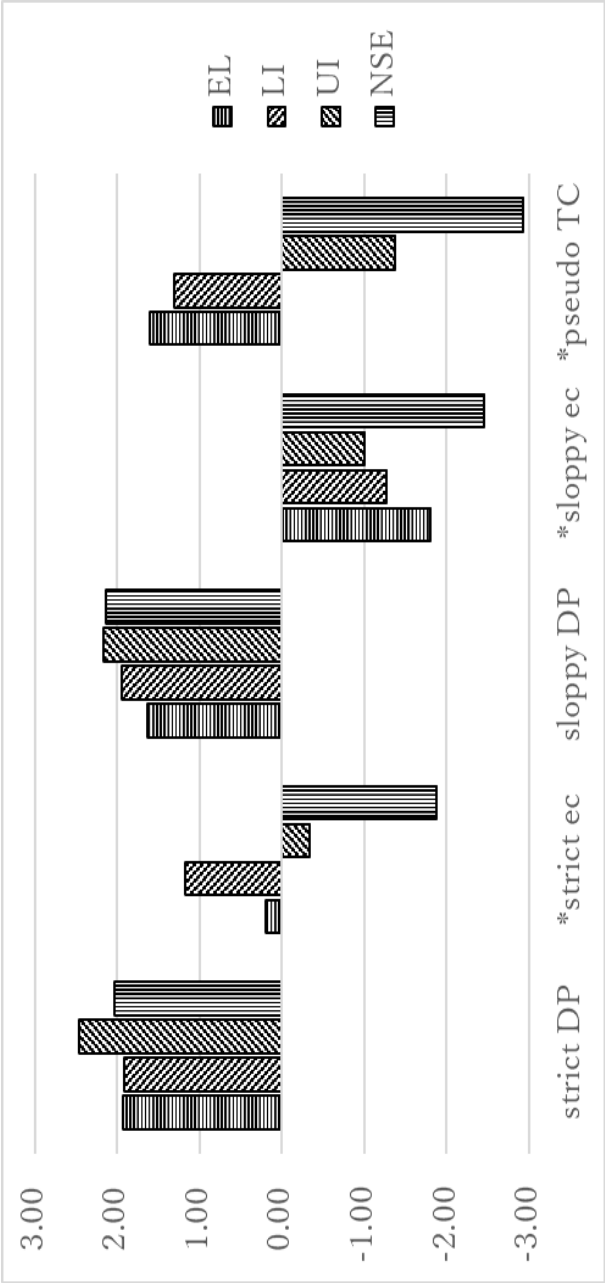


Figure 10. AJT data

NSEs accepted the grammatical types and rejected the ungrammatical types as expected. JLEs, however, accepted the grammatical sentences (i.e., *strict* and *sloppy*) but did not reject the sentences with null objects as much as expected. Interestingly, their responses to *sloppy AE showed that the more proficient learners accepted sloppy reading with empty categories more often than the less proficient learners.

A two-way repeated-measures analysis of variance (i.e., 4 participant groups $\times$ 5 sentence types) was conducted in js-Star, assuming $N_h = 8.72$ as the harmonic mean. The results revealed a significant interaction between the participant groups and sentence types, $F(12, 195) = 5.87, p < .01$, partial $\eta^2 = .3285$). The responses to Type 2 and Type 5 sentences differed significantly between the groups of learners (i.e., EL, LI, and UI), $F(9, 39) = 4.93, p < .01$ and $F(9, 39) = 16.10, p < .01$, respectively. Multiple comparisons using Bonferroni correction showed significant differences between Types 1 and 2 by the UI ($p < .05$) and NS groups ($p < .05$). Therefore, it is safe to conclude that proficient JLEs differentiated those two types. The differences between responses to Type 3 versus Type 4 differed significantly between all participant groups ($p < .05$). Regarding Type 5 sentences, significant differences ($p < .05$) were observed between EL and UI, EL and NS, LI and UI, and LI and NS but not between EL and LI or UI and NS.

If learners who rejected null objects also rejected pseudo-TCs, then there should not have been any differences between Types 2, 4, and 5, but such was not the case.²³ The LI group showed significant differences ($p < .05$) between Type 2 and Type 4. By contrast, no significant differences emerged among any learner groups in responses to Type 4 sentences. Significant differences were observed, however, between Type 2 and Type 5 sentences in LI and between Type 4 and Type 5 sentences in the EL and LI learner groups' responses. Those data showed that L2 learners' acceptance of object drops varied depending on their interpretations and was not likely to be related to the rejection of pseudo-TCs.

²³ An anonymous reviewer suggested, "But even so, learners may encounter difficulties especially with different proficiency levels." The reviewer suggests that less proficient learners may have difficulty (e.g., lack of confidence), which may be reflected in the data. The results of the analysis in Section 6.5 respond to that problem.

6.5 Individual Learners' Data

In this subsection, to confirm that individual JLEs varied in their acceptance of null objects with the strict reading from ones with the sloppy reading and from ones with the pseudo-TCs, we examined individual participants' behavior by comparing AJT-related data for each participant. Afterward, we compared those data with ETVJT-related data. Because Figure 8, depicting JLEs' responses to the ETVJT, shows that their performance was not straightforwardly related to their OQPT scores, we analyzed JLEs' data without referring to the proficiency-based groups. Responses to the AJT were converted to target-like responses when they had positive values (i.e., 1, 2, or 3) to Type 1 (i.e., strict DP) and Type 3 (i.e., sloppy DP) and when they had negative values (i.e., -1, -2, or -3) to Types 2 (i.e., *strict *ec*), 4 (i.e., *sloppy *ec*), and 5 (i.e., *pseudo-TC).

Because Types 1–4 related to object drops and the interpretation of the empty categories, we examined the responses to those types first. JLEs' behavior could be considered to be consistent if their responses were either positive or negative in five or six tokens in one type. Ultimately, their responses were consistent in 95 out of 108 instances (i.e., 27 participants $\times$ 4 types). Therefore, it is safe to say that their behavior was rule-governed and not at random.

We also found eight patterns in individual JLEs' responses, as illustrated at the bottom of Figure 11. One participant, jp16, was exceptional in not accepting Type 1 representing Pattern 1. All other JLEs correctly and consistently accepted that type. Four JLEs (i.e., jp24, jp25, jp18, and jp9) accepted all types of sentences, so their responses in Types 1 and 3 were target-like, but Types 2 and 4 were not (Pattern 2). One participant, jp17, differed from the others due to consistently accepting Type 1 but rejecting Type 4 representing Pattern 3. Another participant, jp6, was also exceptional in responding to Types 2 and 4 inconsistently, thereby representing Pattern 4. Among other results, six JLEs (i.e., jp22, jp12, jp27, jp02, jp10, and jp20) accepted Types 1, 2, and 3. Those learners accepted object drops only when the empty category referred to the person in the contexts (i.e., *strict *ec*) representing Pattern 5. Four JLEs (i.e., jp13, jp21, jp14, and jp5) consistently accepted Type 1 and

Shigenori Wakabayashi, Yuki Takahashi

rejected Types 2 and 4 but failed to accept Type 3 and thus represented Pattern 6. Three JLEs (i.e., jp23, jp03, and jp07) rejected Type 4 but failed to reject Type 2 and thus represented Pattern 7. Last, six JLEs (i.e., jp8, jp11, jp01, jp04, jp15, and jp19) responded consistently in target-like ways to all sentence types and thus represented Pattern 8.

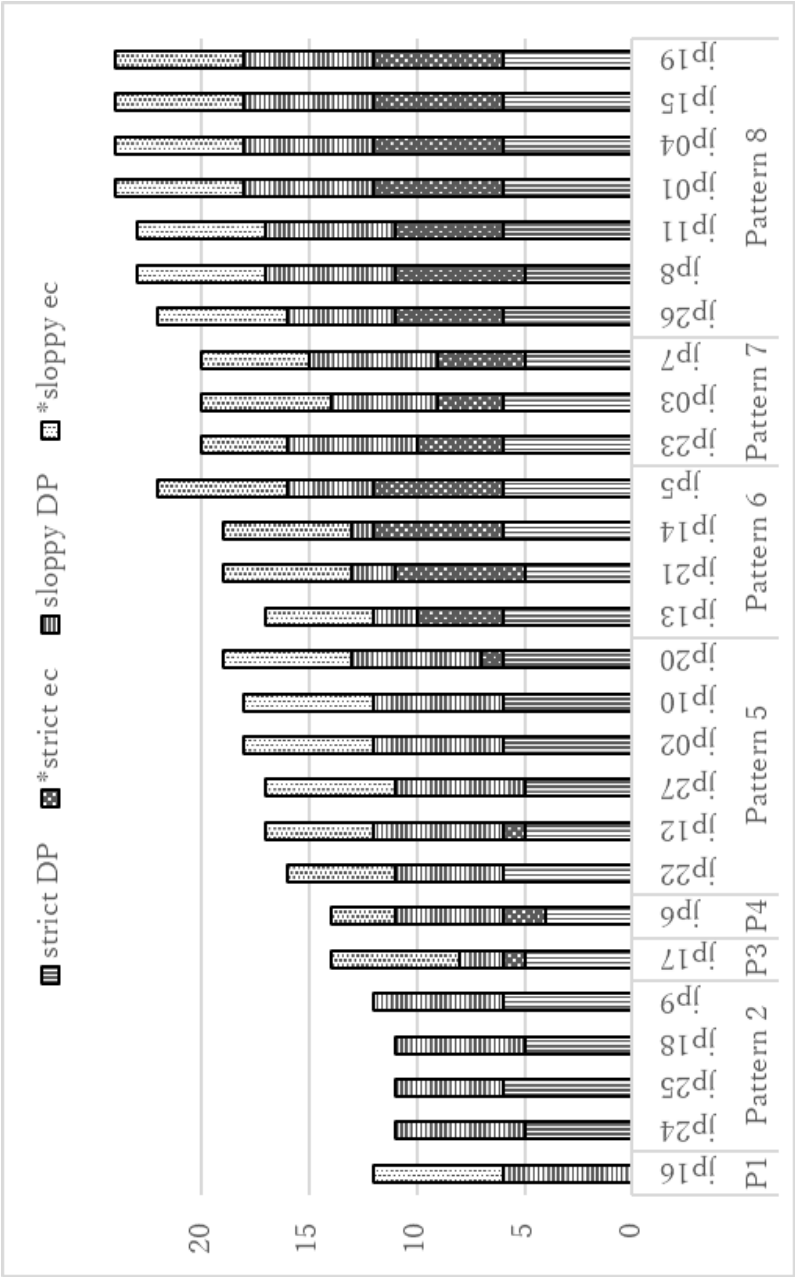


Figure 11. JLEs' target-like responses to Types 1-4 in AJT

The important findings of those analyses are that (1) learners' responses were largely consistent with one type of sentence; (2) only eight patterns, which is much smaller than the possible 16 patterns ($=2^4$), emerged from the data; and (3) no learner showed a pattern with accepting sloppy reading but rejecting strict reading for the empty category. The mirror pattern—that is, Type 2 incorrectly accepted but Type 4 correctly rejected—was found in Pattern 5. That finding is important because it strongly suggests that L2 acquisition is rule-based.

Regarding the data with pseudo-TCs, we added target-like responses to Type 5 (i.e., *pseudo-TC) on top of each participant's vertical bar in Figure 11, and the result appears in Figure 12. If rejecting a sentence with object drops implies that the learner's grammar is native-like, then the learners who accepted neither Type 2 (i.e., *strict *ec*) nor Type 4 (i.e., *sloppy *ec*) should reject a sentence such as (24) because the complement of infinitival verbs in English TCs has to be empty:

(24) *Edward is easy to buy expensive rings. (Type 5: *pseudo-TC)(=23)

Of the seven learners who showed target-like responses to Type 1–4 sentences (i.e., Pattern 8), only three learners showed consistent target-like responses to Type 5. We thus concluded that rejecting object drops and rejecting pseudo-TCs were unrelated.

As for the relationships between data from the AJT and ETVJT, in the same way as the pseudo-TC data were added, ETVJT data were placed on top of the vertical bars in Figure 12. The result appears in Figure 13.

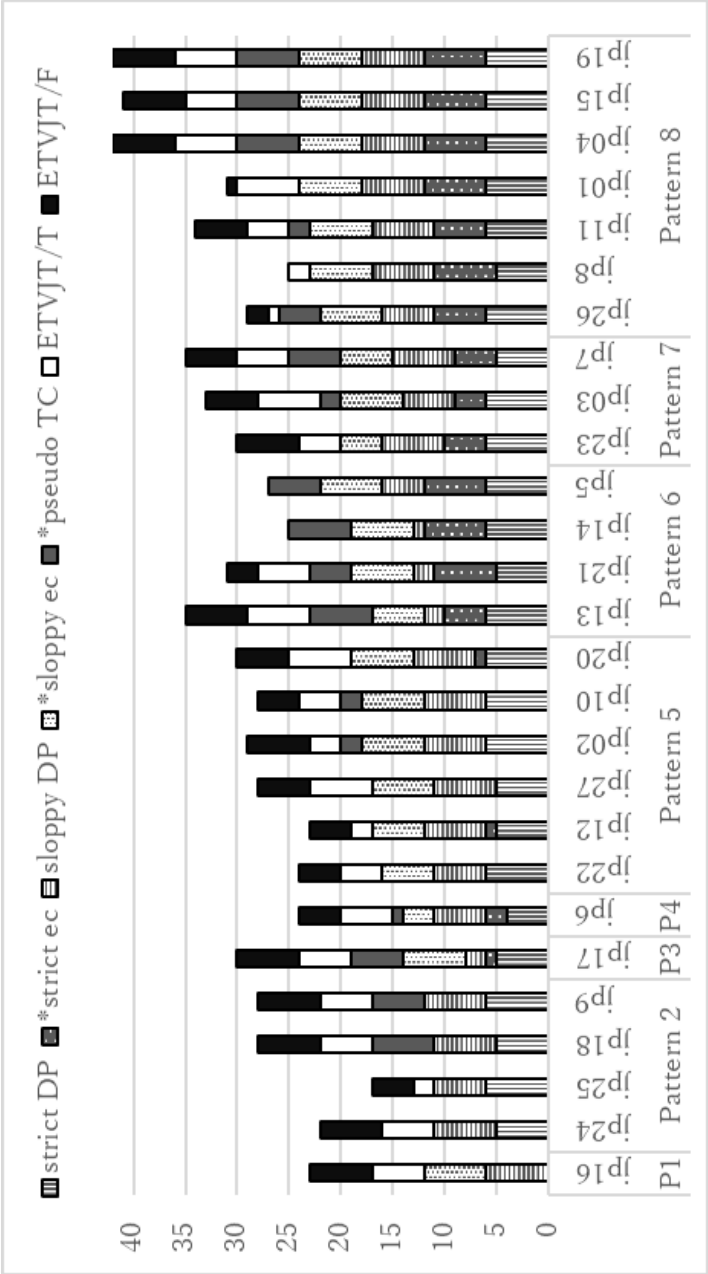


Figure 12. JLEs' target-like responses to Types 1–5 in the acceptability judgment task

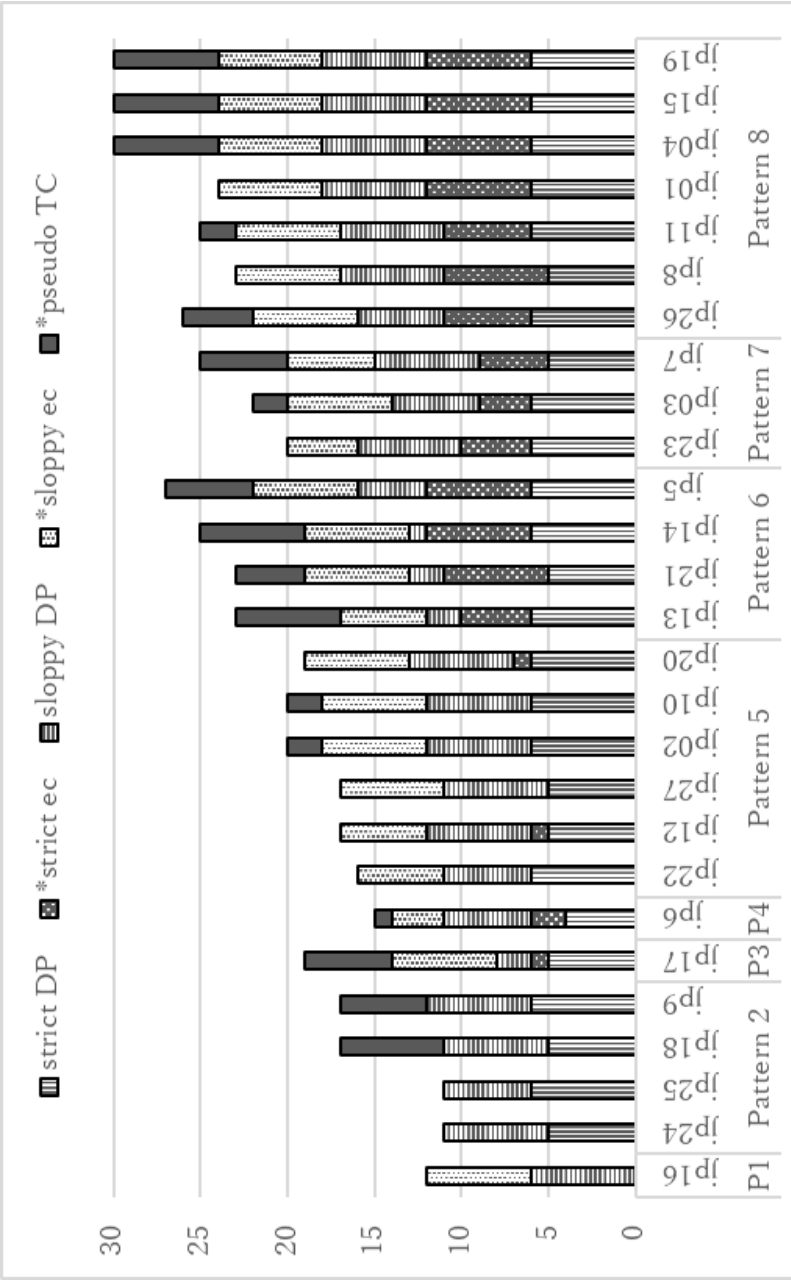


Figure 13. Individual target-like responses of Japanese learners of English to the acceptability judgment task and English truth-value judgment task

Among learners representing Pattern 8, jp04, jp15, and jp19 responded similar to adult native speakers. They may have had target-like grammar, but this may not have been the case. We will discuss this possibility in Section 7. On the other hand, many participants representing Patterns 1–5 also answered in target-like ways to the ETVJT even though they accepted object drops with strict reading. Thus, the results of the AJT and those of the ETVJT appear to be unrelated.

The results of data analysis showed that JLEs' grammar appears to vary far more widely than Kim and Schwartz (2022) determined. At least eight patterns appeared in the data from the AJT with object drops and the interpretation of empty categories, and those patterns do not fit with the response patterns in AJT with pseudo-TCs or with those in the ETVJT.²⁴

As for why L2 learners showed such variation, we answer that question in the next section.

²⁴ Native speakers' responses are not presented in this paper, because they consistently match what theory predicts.

7. DISCUSSION

We developed four research questions, as shown in Section 5. The first question concerned how JLEs interpret English TCs. The results of our experiment with the ETVJT showed that some learners interpreted the subjects of English TCs as being the agent or actor of the infinitival verb, whereas others interpreted them as being the patient or theme. Although Japanese allows both readings, some JLEs did not interpret the subject as the agent or actor (see Figures 8–13). Therefore, it is plausible that their grammar was more restrictive than their L1 grammar.

In short, some JLEs showed a target-like interpretation of English TCs. Whatever grammar guided that performance, the system was narrower than the system of their L1. If we assume that the L1 grammar was the initial learner's grammar, then a learnability problem arises, for learners' systems become more restrictive during L2 acquisition. Kim and Schwartz (2022) have suggested that UG operates in the L2 acquisition of English TCs and that advanced learners acquire a target-like grammar thanks to it. However, they do not discuss how UG operates to make a learner's grammar more restrictive than in an earlier stage, which counters the Subset Principle (Berwick 1985).

We suggest that although UG may guide L2 acquisition, its operation requires certain conditions such as consistency in input (Kimura and Wakabayashi 2024). Most of the L2 learners in our study likely had not acquired the target-like grammar considering that all learners but three (i.e., jp04, jp15, and jp19) did not respond to the AJT in target-like ways, as shown in Figure 14. The fact that many variations emerged in responses to the AJT also suggests that UG operates differently from its operation in L1 acquisition by not constraining a learner's grammar from the very outset of L2 acquisition.

We propose that learners' L1 grammar was used to interpret English TCs. There was also a lack of positive input for the agent or actor interpretation of English TCs and positive input for the patient or theme interpretation. Consequently, JLEs came to disallow syntactic and semantic operations underlying the former interpretation. That learning process may resemble the one suggested in Yuan (2014; cf. Zhang 2023), such that when L2 learners receive enough L2 input to establish a

grammar in which a formal feature used in an earlier stage of development is not used in the target grammar, the feature becomes dormant in a later stage. What feature, or set of syntactic and semantic operations in the case of English TCs, is relevant to (dis)allow the agent or actor interpretation is an important question. However, we need to leave it to future research.

The second question addressed whether JLEs allow object drops in sentences with an obligatory transitive verb such as (20b) and, if so, how they interpret the empty category following the verb. As shown in Figure 12, some learners allowed object drops, whereas others did not. Among ones who allowed them, some learners allowed strict and sloppy readings of empty categories, whereas others only allowed strict readings. No JLE allowed sloppy readings only. On that whole, that asymmetry clearly shows that L2 grammar is rule governed.

Four JLEs in Pattern 6 (see Figure 12) did not consistently accept Type 3 (i.e., sloppy DP) sentences even though the type was grammatical in both English and Japanese, thereby implying that a sloppy reading is generally difficult. Thus, JLEs may not have allowed object drops with sloppy readings more consistently than ones with strict readings. Another possibility is that UG makes a type of language subject to both types of reading of empty categories (e.g., Japanese) and another to allow only strict readings but not sloppy readings (e.g., Spanish); however, UG does not allow empty categories with only sloppy reading, and that constraint from UG may operate in L2 acquisition. Those two explanations are both speculative, but the latter is more plausible. However, we could not find relevant studies on that question.

The third question was whether JLEs accept sentences with the pseudo-TCs. The data showed that some did, but others did not. Importantly, that variation among learners did not correspond to the allowance of object drops, as shown in Figure 13. Beyond that, it did not correspond to the interpretation of the semantic role of the subject of English TCs, either. Figure 14 shows that some JLEs (e.g., jp18 and jp9 representing Pattern 2 and jp17 representing Pattern 3) rejected pseudo-TCs and correctly interpreted English TCs. Even so, other JLEs (e.g., jp14 and jp5 representing Pattern 6) rejected the pseudo-TCs but interpreted English TCs in a non-target-like way. The lack of

correspondence among the (non-)acceptance of object drops, that of pseudo-TCs, and the interpretation of the subject of English TCs suggest that the rules related to those three types of sentences are independent of each other, which is probably also valid with the three JLEs (i.e., jp04, jp15, and jp19 representing Pattern 8) even though they showed native-like responses to all sentence types.

The fourth and most important research question was whether JLEs behave similar to native speakers in response to all three items. We found that three JLEs did. However, as suggested, responses equivalent to those from NSEs were necessary but insufficient to argue that they have acquired native-like grammar. It is more likely that those learners behaved similar to native speakers based on a rule (i.e., a set of syntactic and semantic operations) specific to one item each. In any case, additional studies are needed to investigate whether and, if so, how UG operates in L2 acquisition of English TCs, if not others as well.

Another question we need to ask is why L2 grammar was likely to differ from the target grammar even if L2 learners behaved similarly to native speakers in responding to questions in the experiment. The key to answering that question lies in the answer to the question of how native speakers acquire the system in which English TCs are derived with operations described in Section 2.1. Wexler (2013) has suggested that children's early grammar does not allow English TCs because it lacks a lexical item (i.e., small *v*) that adult grammar has. Later, once it is acquired, children's grammar can deal with several complex syntactic structures, including English TCs. Because the item (i.e., small *v*) is available to L2 learners from an early stage of development, the acquisition of those complex structures may not be acquired at approximately the same time in L2 acquisition. L2 learners rely on the available L2 grammar at the developmental stage and adjust it to connect sounds, spellings, and signs with meanings. That adjustment may be guided by UG (Kimura and Wakabayashi 2024) and under principles of economy specific to L2 acquisition, including the principle of doing everything in narrow syntax (Wakabayashi 2021), which are specific to L2 (i.e., non-L1) acquisition. In short, although both L1 and L2 acquisitions are constrained by UG, the available knowledge and operative principles of language economy differ for acquiring complex

structures such as English TCs, which yields the difference in native child or adult speaker grammar and L2 grammar. Whether that consideration is on point is open to further research.

If L2 grammar is identical to a native speaker's grammar, in which the same operator movement is involved in English TC, then L2 grammar should allow English TCs with the parasitic gap as in (25a) with a *wh*-question in (25b). Beyond that, it should disallow the subject to be interpreted as the complement of a verb in a complex noun phrase as in (25c) in the same way that a *wh*-word extraction from a complex noun phrase island is not allowed, as in (25d). Whether that prediction is borne out is also left to further study.

- (25) a. The book is hard to buy without reading _____. (Chomsky 1982:56)
b. Which report did you file without reading _____?
(Lasnik and Uriagereka 1988:72)
c. *John is easy for you to believe the claim that Mary saw _____.
d. *Who did John believe the claim that Mary saw _____?

Kim and Schwartz (2022) have suggested that sentences such as (26) might have triggered operator movement in TCs in L2 grammar. However, those sentences may only show that the complements of phrasal verbs, or of prepositions following verbs, have to be null in that type of sentences. They may not trigger the acquisition of syntactic derivations (i.e., A'- and A- movement) used for English TCs in native speakers' grammar.

- (26) a. John is easy to talk to _____.
b. Mary is hard to fight with _____.
c. Bill is fun to laugh at _____. (= (10) in Kim and Schwarz 2022:281)

If we follow Kim and Schwartz's (2022) criteria—that is, interpret English TCs in target-like ways and disallow Type 2 (i.e., *strict *ec*) in the AJT consistently (i.e., they responded to four or more of six tokens in one way)—then three JLEs (i.e., jp11, jp13, and jp23) should be considered to have native-like grammar. However, their L2 grammar differed from native speakers' because they accepted pseudo-TCs.

To repeat our argument, it is more likely that the three learners who responded to all items in both tasks (i.e., jp04, jp15, and jp19) acquired their grammar by adjusting their grammar with L2 input, and the null operator movement in a native speaker's English grammar is unlikely to be involved. Even though UG operates in SLA, acquiring a native speaker's grammar underlying English TCs appears impossible, at least for JLEs.

8. CONCLUSION

In this paper, we have examined the grammar of JLEs underlying their interpretation of English TCs, along with the judgment of the acceptability of object drops and pseudo-TCs. After describing the syntactic derivation underlying English and Japanese TCs, we reviewed three papers regarding the acquisition of English TCs. Afterward, we presented a linguistic description of null and overt objects in Japanese, which is relevant to discussing whether JLEs' grammar for (apparent) English TCs is identical to native speakers' grammar. Next, we articulated research questions and conducted an experiment with two tasks. The results revealed that the data showed far more variation between individual learners than in previous studies, which implies that any learner's grammar that exhibited native-like responses in the experiment likely differed from a native speaker's grammar. At the same time, the experimental data showed that the grammar of learners is much more restricted than logically possible. The fact that the interpretation of (ungrammatical) empty categories in the complement position is only sanctioned by UG strongly suggests that UG operates in L2 acquisition. We argue, however, that UG operates in L2 acquisition in ways different from its operation in L1 acquisition because certain lexical items and/or features (and associated sets of syntactic operations) are available from an early stage of L2 development but may not be in L1 acquisition. We also suggest that different principles of economy operate in those two types of language acquisition.

REFERENCES

- Berwick, Robert C. 1985. *The Acquisition of Syntactic Knowledge*. Cambridge: MIT Press.
- Chiba, Shuji. 2019. *Eigo tough kôbun-no kenkyû* [Study on Tough Constructions in English]. Tokyo: Kaitakusya.
- Chomsky, Carol. 1969. *The Acquisition of Syntax in Children from 5 to 10*. Cambridge: MIT Press.
- Chomsky, Noam. 1964. *Current Issues in Linguistics Theory*. The Hague: Mouton.
- Chomsky, Noam. 1977. On wh-movement. In *Formal Syntax*, eds. Peter Culicover, Thomas Wasow, and Akmajian Adrian, pp.71–132. New York: Academic Press.
- Chomsky, Noam. 1982. *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge: MIT Press.
- Collins, Chris. 2005a. A smuggling approach to raising in English. *Linguistic Inquiry* 36:289–298.
- Collins, Chris. 2005b. A smuggling approach to the passive in English. *Syntax* 8: 81–120.
- Connine, Cynthia, Fernanda Ferreira, Charlie Jones, Charles Clifton, and Lyn Frazier. 1984. Verb frame preferences: Descriptive norms. *Journal of Psycholinguistic Research* 13:307–319.
- Gahl, Suzanne, Dan Jurafsky, and Douglas Roland. 2004. Verb subcategorization frequencies: American English corpus data, methodological studies, and cross-corpus comparisons. *Behavior Research Methods, Instruments, and Computers* 36:432–443.
- Hicks, Glyn. 2009. Tough-constructions and their derivation. *Linguistics Inquiry* 40:535–566.
- Hoji, Hajime. 1985. *Logical Form Constraints and Configurational Structures in Japanese*. Seattle: University of Washington dissertation.
- Inoue, Kazuko. 1978. Tough sentences in Japanese. In *Problems in Japanese Syntax and Semantics*, eds. John Hinds, and Irwin Howard, pp.122–154. Tokyo: Kaitakusya.
- Inoue, Kazuko. 2004. Japanese “tough” sentences revisited. *Scientific Approaches to Language* 3:75–111.
- Ito, Akihiro. 2018. Nihonzin eigo gakusyûsya-niyoru *tough* kôbun-no rikai [Comprehension of *tough* sentences by Japanese learners of English as a foreign language]. *JACET Journal* 62:49–67.
- Kaneko, Yoshiaki. 1994. Some topics in *tough* constructions. In *Current Topics in English and Japanese*, ed. Mineharu Nakayama, pp.27–51. Tokyo: Hituzi Syobo.
- Kim, Kitaek. 2014. *Unveiling Linguistic Competence by Facilitating Performance*. Mānoa: University of Hawai‘i at Mānoa dissertation.
- Kim, Kitaek, and Bonnie D. Schwartz. 2022. Learnability in the acquisition of the English *tough* construction by L1-Korean adult and child L2 learners. *Second Language Research* 38:259–287.
- Kimura, Takayuki, and Shigenori Wakabayashi. 2024. UG-as-guide in selection and reassembly of an uninterpretable feature in L2 acquisition of wh-questions: Evidence from islands and scope. In *Current Perspectives on Generative SLA—Processing*,

- Influence, and Interfaces: Selected Proceedings of the 16th Generative Approaches to Second Language Acquisition Conference*, eds. Marta Velnić, Anne Dahl, and Kjersti Faldet Listhaug, pp.318–350. Amsterdam: John Benjamins.
- Kishimoto, Hideki. 2017. Case marking. In *Handbook of Japanese Syntax*, eds. Masayoshi Shibatani, Shigeru Miyagawa, and Hisashi Noda, pp.447–496. Berlin: De Gruyter Mouton.
- Koizumi, Masatoshi. 2008. Nominative object. In *Oxford Handbook of Japanese Linguistics*, eds. Shigeru Miyagawa and Mamoru Saito, pp.141–164. Oxford: Oxford University Press.
- Kuno, Susumu. 1973. *The Structure of the Japanese Language*. Cambridge: MIT Press.
- Kuroda, Shige-Yuki. 1965. *Generative Grammatical Studies in the Japanese Language*. Cambridge: MIT dissertation.
- Kuroda, Shige-Yuki. 1987. Movement of noun phrases in Japanese. In *Issues in Japanese Syntax*, eds. Takashi Imai and Mamoru Saito, pp.229–271. Dordrecht: Foris.
- Lasnik, Howard, and Robert Fiengo. 1974. Complement object deletion. *Linguistic Inquiry* 5:535–572.
- Lasnik, Howard, and Juan Uriagereka. 1988. *A Course in GB Syntax: Lectures on Binding and Empty Categories*. Cambridge: MIT Press.
- Nagamori, Takakazu. 2017. Multiple case valuation via merge: In view of Japanese *tough*- constructions. *Sophia Linguistica: Working Papers in Linguistics* 66:91–113.
- Nakagawa, Naoshi. 1997. On tough constructions in Japanese. *IVY* 30:73–88. The Society of English Literature and Linguistics, Nagoya University.
- Nguyen, Quyen. 2018. Various factors in the acquisition of English tough-construction by Vietnamese EFL learners. *Korean Journal of Applied Linguistics* 34:227–254.
- Ohkado, Kikuyo. 1993. *Tough Construction in Japanese*. Montreal: McGill University, MA thesis, Canada.
- Ohso, Miekko. 1976. *A Study of Zero Pronominalization in Japanese*. Columbus: Ohio State University dissertation.
- Okuma, Tokiko. 2018. Ambiguity resolution in English *tough*-constructions by L1 Japanese speakers: A preliminary study. *Bulletin of the College of Liberal Arts and Sciences, Mie University* 3:15–24.
- Otani, Kazuyo, and John Whitman. 1991. V-raising and VP-ellipsis. *Linguistic Inquiry* 22:45–358.
- Postal, Paul M. 1971. *Cross-Over Phenomena*. New York: Holt, Rinehart and Winston.
- Schwartz, Bonnie D., and Rex A. Sprouse. 1994. Word order and nominative case in nonnative language acquisition: A longitudinal study of (L1 Turkish) German interlanguage. In *Language Acquisition Studies in Generative Grammar*, eds. Tuen Hoekstra, and Bonnie D. Schwartz, pp.317–368. Amsterdam: John Benjamins.
- Schwartz, Bonnie D., and Rex A. Sprouse. 1996. L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research* 12:40–72.
- Shibatani, Masayoshi. 1999. Dative subject constructions 22 years later. *Studies in the Linguistic Sciences* 29:45–76.

- Taguchi, Shigeki, and Fumikazu Niimura. 2009. On the subject of the Japanese middle and tough constructions. In *Proceedings of International Congress of Linguistics (ICL 18)*, Seoul, pp.1875–1886.
- Takezawa, Koichi. 1987. *A Configurational Approach to Case-Marking in Japanese*. Seattle: University of Washington dissertation.
- University of Cambridge Local Examinations Syndicate. 2001. *Quick Placement Test*. Oxford: Oxford University Press.
- Wakabayashi, Shigenori. 2021. A principle of economy in derivation in L2 grammar: Do everything in narrow syntax. *Second Language Research* 37:521–545.
- Wexler, Ken. 2013. Tough-movement developmental delay: Another effect of phasal computation. In *Rich languages from Poor Inputs*, eds. Massimo Piattelli-Palmarini, and Robert C. Berwick, pp.146–167. Oxford: Oxford University Press.
- Yip, Virginia. 1995. *Interlanguage and Learnability from Chinese to English*. Amsterdam: John Benjamins.
- Yoshimura, Noriko, Mineharu Nakayama, and Atsushi Fujimori. 2017. The acquisition of control, raising, and tough constructions among Japanese learners of English. In *Proceedings of Pacific Second Language Research Forum 2016 (PacSLRF 2016)*, Tokyo, pp.247–252.
- Yuan, Boping. 2014. ‘Wh-on-earth’ in Chinese speakers’ L2 English: Evidence of dormant features. *Second Language Research* 30:515–549.
- Zhang, Lulu. 2023. Existence and consequences of dormant features: Evidence from L2 acquisition of Chinese object ellipsis. *Second Language Research* 39:649–666.

Shigenori Wakabayashi
Department of English Studies
Faculty of Letters
Chuo University
742-1 Higashinakano, Hachioji, Tokyo 192-0393 Japan
swkbys37@gmail.com

Yuki Takahashi
Wakabayashi Lab
Department of English Studies
Faculty of Letters
Chuo University
742-1 Higashinakano, Hachioji, Tokyo 192-0393 Japan
takayu0302.yt@gmail.com

Shigenori Wakabayashi, Yuki Takahashi

英語中貌似 *tough* 構式的句法結構變化

若林茂則、高橋侑希
中央大學

這篇論文探討了第二語言 (L2) 學習者的語法，重點關注日語母語者對英語 *tough* 結構的解釋、賓語省略和偽 *tough* 結構的可接受性。與過往研究不同，我們的調查揭示了 L2 學習者的語法表現比之前認為的更為多樣化。我們通過真值判斷任務和可接受性判斷任務進行了實驗。結果顯示，L2 學習者對英語 *tough* 結構的解釋既不相關於賓語省略的接受和解釋，也不相關於偽 *tough* 結構的可接受性判斷。我們的發現表明，一些學習者採用一套操作來構建看起來是 *tough* 結構，而另一些學習者則採用另一種方法。此外，學習者對每種類型的句子的反應是一致的，其模式顯示出普遍語法限制了他們的 L2 語法。我們提出，在第一語言和第二語言習得中，普遍語法的運作方式不同。在 L2 語法中，詞彙項目的習得，包括特徵和操作集，受到發展階段學習者語法的影響，並且 L2 語法特有的經濟原則可能在其中起作用。

關鍵字：*tough* 結構、英語、日語、普遍語法、變異

IMPACTS OF EXTENSIVE ENGLISH EXPOSURE ON JAPANESE CHILDREN: LANGUAGE SKILLS AND BRAIN ACTIVITY*

Atsushi Miura, Natsumi Shibata, Shiro Ojima
Waseda University

ABSTRACT

This preliminary study investigates the potential effects of massive exposure to a second language (L2) not only on children's L2 abilities but also on their first language (L1) abilities. We conducted three different experiments. Experiment 1 compared the L2 speaking performance of children to that of adults, showing that children can attain high L2 speaking abilities even in an EFL context if they receive massive L2 input. In Experiment 2, we examined children's academic L1 abilities and speculated that the setting of L2 learning, rather than the amount of L2 exposure, might significantly influence academic L1 abilities. Using a neurolinguistic technique (event-related potentials or ERPs), Experiment 3 explored children's L1 lexical processing and suggested that children exposed to extensive L2 input may exhibit distinct neural lexical processing patterns. We discuss some implications of early L2 learning in a foreign language context.

Keywords: Child SLA, EFL, Speaking, Academic L1 abilities, Lexical processing

* This study was supported by Grant-in-Aid for Scientific Research (KAKENHI) from Japan Society for the Promotion of Science (JP19H01280) to the last author. We would like to express our gratitude to Mizuki Aoshima, Ayaka Yokoyama, Shohei Shimoyama, Misaki Takahashi, Sota Mitsumori, and Minori Shiozaki for their dedicated assistance in data collection.

1. INTRODUCTION

It has been debated whether earlier is really better in second language acquisition (SLA) research. The discussion on this issue is based on the hypothesis called the critical period or sensitive period hypothesis, which states that there is a certain period that is best for language acquisition (Lenneberg 1967). From this account, numerous researchers have made their proposals on age effects in SLA. One of the important generalizations is a distinction between ultimate attainment and the rate of second language (L2) acquisition (Krashen, Long, and Scarcella 1979). Previous studies provided mostly consistent results that young learners have advantages for ultimate attainment (e.g., Long 2013), but older learners for learning rate (e.g., Muñoz 2006). One more important discussion on the age issue is the learning condition. That is, learners in a second language context (e. g., English as a second language context: ESL) and those in a foreign language context (e. g., English as a foreign language context: EFL) usually greatly differ in terms of total language exposure they can receive. Since learners in a foreign language condition tend to get a limited quantity and quality of target-language exposure (Muñoz 2014), it is often difficult for them to acquire a second language.

The debate on "earlier is better" is also applied to English education policies. For example, in Japan, English was introduced as a subject in elementary school in 2020. This political decision led to an early English learning boom (see Terasawa 2018 for a detailed discussion). This is evident in self-help books for Japanese parents as reviewed by Nakamura (2021). In current Japan, some parents create an English-learning environment at home and guide their children's English learning (often without directly teaching them). These parents provide their children with English-learning materials such as CDs, DVDs, and books to familiarize them with target-language sounds and expressions. This type of learning, called Home-English or Ouchi Eigo in Japanese, is also found around the world (e.g., Prosic-Santovac 2016; Thomai 2015; see also Scheffler, Jones, and Domńska 2020). In Japan, there seem to be some cases where children learning English in their home attain a high level of English command. In addition to that type of learning, some parents have their children go to an international school where

instructions are given in L2 (i.e., English) rather than their native language (i.e., Japanese). Most of these schools were classified as Kakusyū Gakko, rather than Ichijō-ko defined by Japan's School Education Act, meaning that their curricula do not follow the "Course of Study" by the Ministry of Education, Culture, Sports, Science and Technology (MEXT). As with Home-English children, these children attending an international school are usually exposed to a massive amount of English input. Investigating these children using a variety of research methods could provide new insights not only into age-related issues in SLA but also into English language education in Asia and elsewhere.

It has also been discussed by researchers and the general public whether and to what extent early L2 learning affects native language development. In this debate, the learning conditions (second language or foreign language context) discussed earlier are also a key to answering the question. Previous studies provided results that children who got intensive L2 exposure in their daily life (i.e., second language context) could lose some of their L1 skills and knowledge (e.g., Pallier, Dehaene, Poline, LeBihan, Argenti, Dupoux, and Mehler 2003). However, children in a foreign language context usually do not lose their L1 (Ferjan Ramirez, and Kuhl 2017, 2020). It is still not clear to what extent children who have received massive L2 exposure through the Home-English context (i.e., massive exposure in an EFL context) have retained their L1 abilities.

In this study we explored the L1 and L2 performance of Japanese children who have been exposed to massive English input in their home or international school. These children can be regarded as different from typical EFL learners, who usually have limited English exposure. Thus, it is important to investigate these children's L1 and L2 abilities. More specifically, we investigated these Japanese children's English productive abilities and academic Japanese abilities as well as brain activity involved in L1 vocabulary processing.

2. BACKGROUND INFORMATION

2.1 Age Issue in Second Language Acquisition

Since the Critical (Sensitive) Period Hypothesis was proposed, several researchers have investigated the effect of age on L2 learning and acquisition. Some previous research has reported the child advantage for sound learning. This is known around the world as young children or babies are highly sensitive to linguistic sounds (Werker, Gilbert, Humphrey, and Tees 1981; Werker and Tees 1984). In terms of grammar, Johnson and Newport (1989) investigated the relationship between age of first arrival at the target-language context and L2 grammatical ability. The participants aged 3 to 39 took an auditory grammaticality judgement test (GJT). The result showed that the earlier they arrived at the target-language context, the higher scores they showed. The authors concluded that the critical period hypothesis originally proposed for L1 acquisition could be applied to L2 acquisition. Since then, several researchers tried to replicate this study (e.g., Dekeyser, Alfi-Shabtay, and Ravid 2010; Granena and Long 2013; see also Johnson 1992 for written GJT). This result was also supported by several researchers investigating the relationship between age and other aspects of language abilities such as production (e.g., Patkowski 1980) and phonology (e.g., Yeni-Komashian, Flege, and Liu 2000).

When discussing the effect of age on language acquisition, it is important to consider the language acquisition context (second language vs. foreign language) where the study was conducted, because the total amount of language exposure is important for language acquisition (e.g., Perani, Abutalebi, Paulesu, Brambati, Scifo, Cappa, and Fazio 2003). Typically, learners in a foreign language context receive limited target-language exposure, usually only in classrooms. Thus, one of the keys to the attainment of high L2 proficiency in a foreign language context is the total amount of exposure they receive. For example, they should get supplemental exposure out-of-school (e.g., Lindgren and Muñoz 2013; Peters, Noreille, Heylen, Bulte, and Desmet 2019). Ojima, Matsuda-Kurita, Nakamura, Hoshino, and Hagiwara (2011b) and Ojima, Nakamura, Matsuba-Kurita, Hoshino, and Hagiwara (2011c) conducted

large scale studies, investigating the effects of age and amount of exposure on second language neural processing. They used regression analyses to verify the effects of the age of first exposure and amount of exposure differently because typically those with an early age of first exposure tend to get a longer amount of exposure. After controlling for the age of first exposure, the results showed that the amount of exposure is a significant factor for target-language proficiency. They also stated the importance of the amount of exposure for foreign language learning in an EFL context, rather than the earlier age of first exposure.

One more important point for the age-related issue is the rate and ultimate attainment of learning. Muñoz (2006) conducted a large-scale longitudinal study named BAF (Barcelona Age Factor) project from 1996 to 2002. In this project, the Spanish-speaking participants were divided into different groups in terms of age of onset (AO) of English learning. Those who began English instruction at the age of 8 (Group 1) and 11 (Group 2) are the central groups investigated. Additionally, participants whose AO was 14, and those whose AO was 18 or older were investigated. The researchers investigated these learners' English performance three times, after 200 hours of instruction, after 416 hours, and after 726 hours. The project provided several key findings. First, in all tests measured, learners with AO of 11 showed higher performance than those with AO of 8. Second, from time 1 to time 2, while older learners grew up well, younger learners showed a little improvement, resulting in a huge difference between their performances. Third, even though the younger group (AO = 8) showed greater improvements, the older group (AO = 11) still showed higher performance at the third testing time (after 726 hours of instruction). These results suggest that older learners have an advantage in the rate of foreign language learning compared to the younger learners, because the older learners progressed faster at least initially. The researchers concluded that older learners may be able to learn a second language quicker than the younger learners, and that younger learners presumably lose their potential advantage when they could not get enough exposure with the target language. Similar results were reported in different countries such as Holland (e.g., Cadierno, Hansen, Lauridsen, Eskildsen, Fenyvesi, and Hannibal Jensen, 2022), Germany (e.g., Jaekel, Schurig, Florian, and Ritter 2017; Baumert,

Fleckenstein, Leucht, Koller, and Moller 2020) and Switzerland (e.g., Pfenninger and Singleton 2016). Thus, children could potentially reach higher levels of L2 proficiency in the long run, but older learners could learn an L2 faster than children at least initially.

Learning styles may also be different for child and adult learners. Basically, adult EFL learners tend to receive L2 input in instructional settings (explicit instruction or explicit learning). They have some advantage for explicit instruction or learning because their cognitive abilities have already matured. In contrast, children are good at learning languages implicitly (DeKeyser 2000). This view is also supported by an experimental study (e.g., Lichtman 2020; but see also Lichtman 2016). One important note for this comparison is that implicit learning requires a substantial amount of time (DeKeyser 2000). Therefore, it is also concluded that children, who engage in implicit learning show slower developmental patterns than adults, who engage explicit learning.

Based on the previous research in the second language and foreign language condition, the consensus is that children have some advantage for learning a second language if they could get a significant amount of exposure to the target language. If they are in the limited exposure condition, research has shown that the child advantage could not be demonstrated.

2.2 Effects of Early Second Language Acquisition on Native Language Knowledge

It has been studied whether L2 acquisition or living in an L2 environment may affect L1 abilities. In fact, previous literature on adult L1 attrition provided evidence that L1 attrition does occur when adult speakers get only L2 exposure but do not have L1 exposure anymore, as in the case of immigrants (see Gallo, Bermudez-Margaretto, Shtyrov, Abutalebi, Kreiner, Chitaya, Petrova, and Myachykov 2021. for more details). Even adults whose L1 has been once completely acquired could experience L1 attrition. Children could also undergo L1 attrition. As with the adult pattern, children potentially lose their L1 partially or even entirely. Pallier et al. (2003) investigated Korean participants who were adopted by French families in their childhood (age 3-8). The researchers

tested their Korean knowledge by behavioral tasks and fMRI (functional magnetic resonance imaging) tasks. The results showed that the participants' performance was similar to the control group consisting of native French speakers without any knowledge of Korean, indicating that L1 attrition occurred for the participants who were exposed only to their L2 in their daily life for many years.

When L1 exposure remains available to L2 learners, however, we can draw a different conclusion. Ferjan Ramirez and Kuhl (2017, 2020) investigated Spanish babies (age; 7 months to 33.5 month) who participated in 18-week (60 minutes for each instruction) or 36-week (45 minutes for each instruction) English instruction. Their results showed that the longer L2 exposure the babies received, the higher L2 scores they showed. Surprisingly, their L1 development was equal regardless of the length of L2 exposure. Synthesizing the results of Pallier et al. and Ferjan Ramirez & Kuhl studies, we could say that similar patterns may hold for both L2 acquisition and L1 maintenance. That is, it is difficult to acquire an L2 in a foreign language environment where exposure to the target language tends to be limited. Similarly, it is difficult to sustain one's L1 when exposure to it is limited in the daily life, but it could be sustained if it is used on a daily basis.

2.3 Motivation for the Current Study

Basically, learners in a foreign language context learn the target L2 in school settings. In the case of east Asian countries like Japan, China, South Korea, and Taiwan, the starting age of instruction has been lowered (see Butler 2015 for a review). Other than public education, some parents think that it is important for their children to start to learn English earlier than they did. Thus, they have their children engage in so-called Home-English activities such as watching DVDs in English in their home and provide them with a substantial amount of English input. Recently, it has been noted anecdotally that Japanese children with Home-English experience are able to use English at an extremely high level. Children who are able to obtain a large amount of input despite being in an EFL environment have not been investigated in previous

SLA studies, and investigating these children could provide many fruitful implications for the theory of SLA and foreign language teaching.

3. CURRENT STUDY

Because of the exploratory nature of our current project, we conducted three experiments from different perspectives using different methods to explore Japanese children's L1 (Japanese) and L2 (English) abilities. Thus, this section is divided into three parts in terms of these three experiments. In Experiment 1, we investigated the children's L2 speaking abilities, comparing it to typical EFL adult learners' performance. In Experiment 2, we tried to shed light on the effects of extensive L2 exposure on academic L1 abilities, investigating the scores of a traditional pencil-and-paper test. In Experiment 3, we conducted a neurolinguistic experiment (more specifically, an event-related potential or ERP experiment), where children's brain activity involved in L1 lexical processing was investigated.

3.1 Experiment 1

In Experiment 1, we investigated children's L2 speaking abilities. Although L2 speaking abilities are considered as the most difficult aspect to improve especially in an EFL context, the importance of fluent and intelligible speech has increased in the current globalized world. There are also anecdotal reports that some Japanese children engaging in Home-English have a very high command of spoken English. Thus, it is beneficial to investigate this aspect (i.e., speaking) in Japanese children's L2 English acquisition. In this experiment, we decided to analyze two different measurements (i.e., speech quantity and speech fluency) as the general speaking abilities.

3.1.1 Methods

3.1.1.1 Participants

Forty-five Japanese children were recruited in Experiment 1. Nine of them were students at international schools (from four different schools), while the others went to public schools in Japan, including the “Home-English” children and the children who have not been in an English-learning environment at home. All of these participants lived and went to school in Kanto (around Tokyo) and Kansai (around Osaka) areas in Japan. In addition to these children, 28 adult university students were included in the analysis to compare the children’s performance with that of young adults studied by a previous study where the same experimental procedure was used (Nambu 2015). Those university students were typical Japanese learners of English, who received a normal amount of English instruction in junior high and senior high school. The mean age of Home-English, international school, and university students was 9.00 (range 6-12), 9.33 (range 8-11), and 20.90 (range 18-24), respectively. We recorded their total hours of English exposure based on a detailed questionnaire. In the questionnaire, we ask caregivers to provide detailed information about their children’s English learning profile from their birth to the time of the experiment. For example, they were asked to write the period of an activity (e.g., 1-year-3-month-old to 2-year-5-month-old), average hours per week, and type of activity (e.g., listening to music, English conversation class, etc.). Then we calculate the estimate of total exposure to the target language by multiplying the hours per week, number of weeks per year ($365/7=52.14$), and period (i.e., total exposure = hours per week $\times$ 52.14 $\times$ period).

According to the questionnaire, the mean total hours of English exposure of them were 5,297 (364-17,472) for Home-English children, 13,845 (7,046-21,008) for international school children, and 1,746 (810-4379) for the university students. The descriptive information of all participants was summarized in Table 1.

Table 1. Information of participants by category in Experiment 1.

Category (<i>n</i>)	Mean total exposure (min-max)	<i>SD</i>
Home-English children (36)	5297.727 (306-17472)	4072.109
International school students (9)	13845.867 (7046-21008)	4262.784
University students (28)	1746.415 (810-4379.150)	802.306

Note. Most but not all the participants were overlapping among the three experiments.

Although we had two categories of children (Home-English and international school), we analyzed their performance not according to these categories but by dividing the children into two groups based on their total amount of exposure to English (low-to-high exposure group and extremely high exposure group). For more details, see Section 3.1.2.

3.1.1.2 Procedure

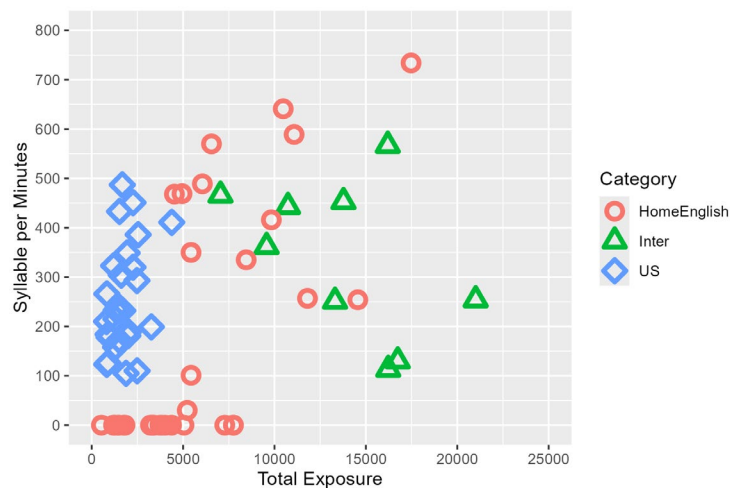
The participants' English-speaking abilities are assessed through a retelling task, in which they were asked to reproduce the content of a movie clip that they watched. The material and procedure in the current study are based on a previous study which studied L2 English speaking of Spanish university students (Muñoz 2014). A short video clip of *Modern Times*, which is about 8 minutes long, is used. The video was divided into two parts. The participants watched the whole video first. Next, they watched the first half of the video again. They were then asked to reproduce the content of the first half of the video in English. After that, they watched the latter half of the video and described it in English. Their retelling speech was recorded individually in a sound-attenuated room and later analyzed in terms of speech quantity and fluency.

3.1.1.3 Analysis

First, the participants' production was transcribed manually. Then any repeated production and utterance repair were excluded. Since there were two recording sessions for each participant (i.e., the former part and the latter), the two sets of data were summed for the final analysis. We choose two measurements reflecting different aspect of speaking abilities, namely, total number of produced words representing the speech quantity and syllables per minute representing speech fluency. The former is obtained as the total number of words each participant produced during the retelling task. The latter is calculated as the total number of syllables per minute of speech, that is, the total syllable number divided by the total time of speech measured in minutes.

3.1.2 Results

Figure 1 shows the results of the speech quantity (i.e., total number of words produced) and Figure 2 shows the results of fluency (i.e., syllable per minute of speech).



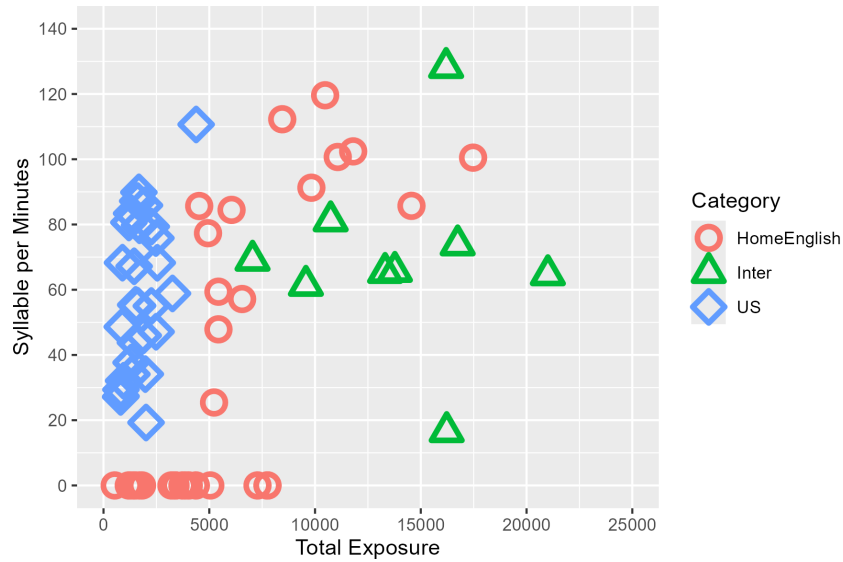


Figure 2. Descriptive plots of the results of syllable per minutes by three groups. US (blue-diamond shape) stands for University Students.

To identify the differences in their performance clearly, we divided the children into two groups based on their total exposure rather than their category (i.e., either Home-English children or international school children), 1) Low-to-high group (0 to 5000 hours of total exposure) and 2) Extreme group (more than 5000 hours of total exposure). Note that in this grouping analysis, we do not analyze the category of children (i.e., international school children and Home-English children). Table 2 summarizes the descriptive statistics of the three groups.

Table 2. Information of participants by group.

Group (n)	Mean total exposure (min-max)	SD
Extreme group (25)	10592.443 (5048.238-21008)	4506.358
Low-High group (20)	2525.995 (364-4940)	1477.992
University students (28)	1746.415 (810-4379.150)	802.306

Descriptive results of total number of words produced and syllables per minute are shown as Table 3 and 4, respectively.

We conducted a one-way analysis of variance (ANOVA) to analyze the differences in two measures, total number of words produced and syllable per minute, among the university students, low-to-high children, and extreme children. As for the total number of words (i.e., speech quantity), the homogeneity test was performed first, providing a significant result ($p < .001$). Thus, the Brown-Forsythe correction was applied for the analysis. The results of a one-way ANOVA revealed a significant difference among the three groups ($F(2, 53.289) = 14.664$, $p < .001$, $\eta_p^2 = .291$), so we further conducted post-hoc tests with Tukey correction. The result reveals that there is a significant difference between Extreme children and Low-High children ($t = 5.165$, $p < .001$, $d = 1.536$), and University students and Low-High children ($t = 4.249$, $p < .001$, $d = 1.244$). However, there is no significant difference between Extreme children and University students ($t = .292$, $p = .534$, $d = .292$).

For the fluency results (i.e., syllable per minute), we again ran a one-way ANOVA with syllable per minute as a dependent variable and Group as an independent variable. Because the homogeneity test resulted in statistical significance ($p = .014$), the Brown-Forsythe correction was again applied to the analysis. The result of ANOVA showed a significant effect of Group ($F(2, 59.244) = 22.554$, $p < .001$, $\eta_p^2 = .383$). The result of post-hoc tests with Turkey correction reveals that there is a significant difference between Extreme children and Low-High children ($t = 6.106$, $p < .001$, $d = 1.816$), and between University students and Low-High children ($t = 5.651$, $p < .001$, $d = 1.654$). However, there is no significant difference between Extreme children and University students ($t = .594$, $p = .524$, $d = .162$).

Table 3. Descriptive results of syllable per minute by three groups.

Group (n)	Mean (min-max)	SD
Extreme (26)	63.002 (0-128.158)	38.805
Low-High (20)	8.15 (0-85.663)	25.121
University students (28)	58.117 (19.285-110.659)	23.672

Table 4. Descriptive results of total number of words produced by three groups.

Group (n)	Mean (min-max)	SD
Extreme (26)	302.577 (0-734)	224.662
Low-High (20)	46.850(0-469)	144.201
University students (28)	253.929(105-487)	107.199

3.1.3 Summary

As for the English-speaking performance, the statistical result comparing the performance from low-to-high children and university students shows significant differences. As indicated by Figure 1, although some children could not produce any English sentence at all, all of the university students enrolled in the task managed to produce some speech. Regarding the comparison between the extreme group and university student group, the statistical test did not show any significant difference. It is important to note, however, that the extreme group shows huge standard deviation. This is confirmed by a visual inspection of the results, showing that there is huge individual variation. This presumably indicates that some child learners managed to take advantage of massive input, but others did not.

While all university students were able to perform the task, some of the children could not speak at all. When the total English exposure is short, older learners (late starters) show higher performance. This result is consistent with the previous study showing that older learners can learn L2 faster than younger learners at least initially (e.g., Muñoz 2006).

3.2 Experiment 2

In Experiment 2, we investigated the children's academic L1 abilities. It is unclear to what extent children who have received massive L2 exposure have maintained their academic L1 abilities. To tackle this issue, we administered a criterion-referenced diagnostic test of academic Japanese to Japanese children with diverse English-learning experience.

3.2.1 Methods

3.2.1.1 Participants

Most, but not all, participants were identical to those in Experiment 1. Seventy-two children in total were enrolled in the experiment. All of these participants lived and went to school in Kanto area (around Tokyo) in Japan. Eleven were students at international schools (from four different schools), 51 were Home-English children (from 32 different schools), and 10 had very limited English exposure (from 10 different schools). Due to the nature of the test, the children took a version of the test according to their grades. There were one child in Grade 1, 23 in Grade 2, 16 in Grade 3, 17 in Grade 4, 6 in Grade 5, and 9 in Grade 6. The detailed information about the participants is summarized in Table 5. In this experiment, we would like to focus on their academic Japanese abilities based on the learning environment (i.e., Home-English children or international school students). We thus divided the child participants into three groups, little English ($n = 10$; total exposure < 1000 hours), Home English ($n = 41$; total exposure > 1000 hours), and international school students ($n = 11$).

Table 5. Participants' information in Experiment 2.

Group (n)	Mean total exposure (min-max)	<i>SD</i>
Little English (10)	419.9 (104-988)	303.155
Home-English (41)	4597.716 (1040-17181)	3691.359
International school students (11)	13382.455 (5720-22126)	5315.414

3.2.1.2 Material

We used a diagnostic test of academic achievement referenced to the goals of the MEXT's "Course of Study". The test used is the academic Japanese part of Hyokenshiki CDT (標研式 CDT) test developed by Nihon Hyojun (日本標準). This test is supposed to be taken towards the end of an academic year (i.e., during January to March) and different versions are available for different grades. The test contains Kanji writing questions, Kanji reading questions, reading comprehension questions, and questions about attitude toward learning Japanese. For only Grade 1 to 4, it includes a listening test, where the test takers listen to a passage in Japanese.

3.2.1.3 Procedure

Referring to the criteria by the test, the test was administrated around the end of the academic year (i.e., January or February). Due to the COVID-19 pandemic, we decided to conduct the test online using the Zoom platform (<https://zoom.us/>). Participants entered a Zoom room from a quiet place in their home and were instructed to be ready for the test (for example, they were asked to turn on the camera, remove from the desk anything unrelated to the test (e.g., smartphone, textbook, etc.)).

3.2.1.4 Scoring Procedure

The test contains several question types such as multiple-choice questions and writing Kanji characters. As for Kanji questions, scoring discrepancies are often observed. For Grade 3-6, there is a written summary question. These questions could also result in disagreement by coders. Thus, intercoder reliability for written Kanji questions and written summary questions were checked using 25% of the data coded by two coders (18 out of 72 participants' data). The reliability was confirmed to be high enough (Kappa coefficient = .902).

3.2.2 Results

Figure 3 shows the descriptive plots showing the relationship between accuracy rate in the academic Japanese test and total hours of English exposure, differentiating the three groups (i.e., Little English, Home-English, and International School children) by the colors and shapes of the plots. First, we should note that all children below the accuracy of 70% are those attending an international school. Children in the Little English group and the Home-English group all scored 70% or higher. In other words, the scores of all Home-English children fell in the normal range shown by the Little-English group, whereas some of the children attending an international school were clearly below this range. The statistical test (i.e., ANOVA) confirms the differences among the three groups ($F(2, 59) = 10.963$, $p < .001$, $\eta_p^2 = .271$). According to the post hoc tests with Tukey correction, there is no statistically significant difference between the Little English and Home-English group ($t = .302$, $p = .951$, $d = .106$), but the international school students significantly underperformed children with little English experience ($t = 3.360$, $p = .004$, $d = 1.468$).

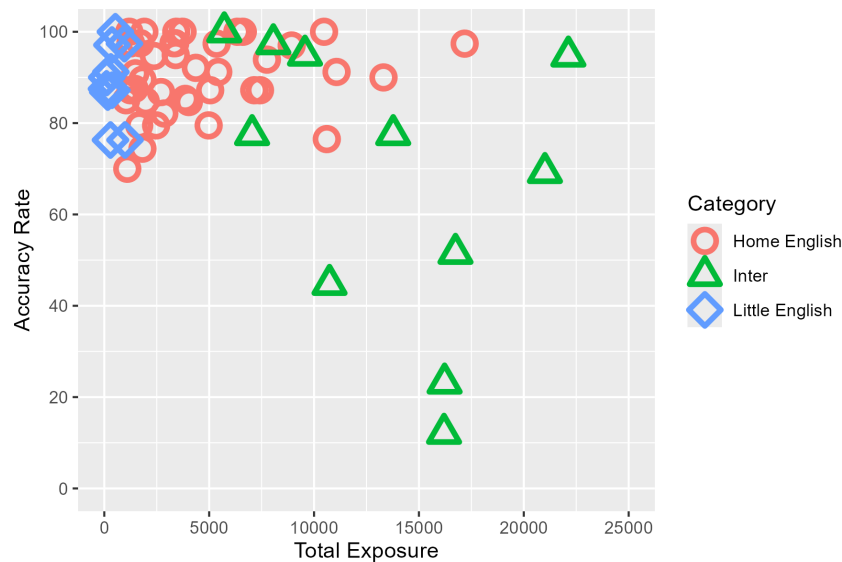


Figure 3. Descriptive plots of the results of accuracy rate by three groups

To further investigate the effects of L2 exposure on academic L1 abilities, we ran a simple correlation analysis separately for the three groups. We found no statistically significant correlation between accuracy rate and total hours of exposure in any of the three groups (all $ps > .2$). However, this may not mean that there is no relation between these two factors. Looking at the distribution of the data, it may be the case that individual variation in academic Japanese among international-school students increases as their total amount of English exposure increases (their data spread more widely towards the right-hand side in Figure 3). This tendency cannot be captured by a simple correlation analysis.

3.2.3 Summary

As for the children attending Japanese school (little English and Home English groups), there is no evidence to suggest that the amount of English exposure has an impact on their academic Japanese abilities. However, the results from children attending international schools show a large individual variation, with some showing lower accuracy rates. One possible reason for this variation is that the international schools attended by those children tend not to have Japanese (“Kokugo” or “national language”) classes. In addition to that, they receive a large amount of L2 input (i.e., English) throughout their school life. This might mean that they could not receive sufficient exposure to native language (i.e., Japanese) input in the academic setting. It is also important to note that the administrated test is criterion-referenced to the Course of Study set by Japan’s Ministry of Education, which is typically not followed by international schools in Japan. Therefore, children in international schools may have difficulty achieving high scores on this particular type of academic test. Using a different type of test that taps into more basic communicative aspects of native language proficiency could yield different results from those observed in the current study. Further investigation with a larger sample size is needed to obtain conclusive evidence regarding the effect of early L2 learning on academic L1 abilities.

3.3 Experiment 3

In Experiment 3, we investigated children's brain activity involved in L1 lexical processing. Electroencephalograms (EEGs) can provide rich information about both L1 and L2 processing in the brain. Event-related potentials (ERPs) are averaged segments of EEGs time-locked to a particular event such as the presentation of written or spoken words to participants. ERPs provide rich information on both L1 and L2 processing since they have high temporal resolution required to track rapid language processing in the brain.

Previous studies have shown that specific ERP components reflect language users' neural processing. For example, for lexical-semantic processing, it is well known that the ERP component called the N400 can be elicited by a mismatch condition where the semantic context does not allow the brain to predict an incoming word. The N400 peaks around 400 ms after the onset of the mismatch stimulus and has been supported by numerous researchers (e.g., Kutas and Federmeier 2000, 2011). Additionally, the Late Positive Component (LPC) sometimes appears after the N400, with a tendency to be missing in young children's L1 sentence processing (Juottonen, Revonsuo, and Lang 1996).

As for L2 acquisition research, the amplitude (e.g., to what extent the mismatch condition differs from the match condition in a picture-word mismatch paradigm) and the timing (when the mismatch effect appears) had been considered as an index of L2 lexical processing competence. Ojima et al. (2011c) conducted a three-years longitudinal study where child L2 learners were enrolled. Their results show that children with higher L2 proficiency show a clearer N400 effect. Moreover, as their proficiency increases, their N400 effect becomes comparable to that of more proficient learners (see also McLaughlin, Osterhout, and Kim 2004).

3.3.1 Methods

3.3.1.1 Participants

Most, but not all, participants were identical to those in Experiments 1 and 2. Forty-three Japanese children in total were enrolled in this experiment. All of these participants lived and went to school in Kanto area (around Tokyo) in Japan. We divided them into three groups according to their total hours of exposure to English: Low exposure group (below 1000 hours), High exposure group (between 1700 and 5000), and Extreme exposure group (more than 5000 hours). Detailed information about the participants is summarized in Table 6. Due to the sample size, we decided not to distinguish Home-English children and international school students.

Table 6. Participants' information in Experiment 3.

Group (<i>n</i>)	Mean total exposure (min-max)	SD
Group 1: Little exposure (14)	382.222 (26-988)	313.558
Group 2: High exposure (15)	3150 (1716- 4975.36)	938.35
Group 3: Extreme exposure (14)	9099 (5030.22- 16244)	3842.466

3.3.1.2 Material

The materials used in the current study are adopted from a previous study (Ojima et al. 2011c). One professional female bilingual narrator was recruited to record the experimental stimuli, which were 80 basic-level Japanese words. Colored pictures representing these words were also created. Each spoken word was presented twice during the experiment, one in the match condition, and the other time in the mismatch condition. The participants completed 160 trials (see Ojima et al. 2011c for more detailed information). They also completed another

set of 160 trials with English words, but we will report the Japanese data only due to the limited space of the current manuscript.

3.3.1.3 Procedure

The 160 Japanese words were presented in two sessions to insert a short rest in the middle. Each trial presenting one word consisted of the following events: fixation cross for 1000 ms; appearance of a picture; 1000 ms after the onset of the picture presentation, a word was played through loudspeakers; 750 ms after the offset of the word, the picture disappeared. Each trial took 4000 ms in total (see Figure 4). The experimental procedure was processed through E-Prime (Psychology Software Tools 2011). The total length of experiment was around 25 minutes on average.

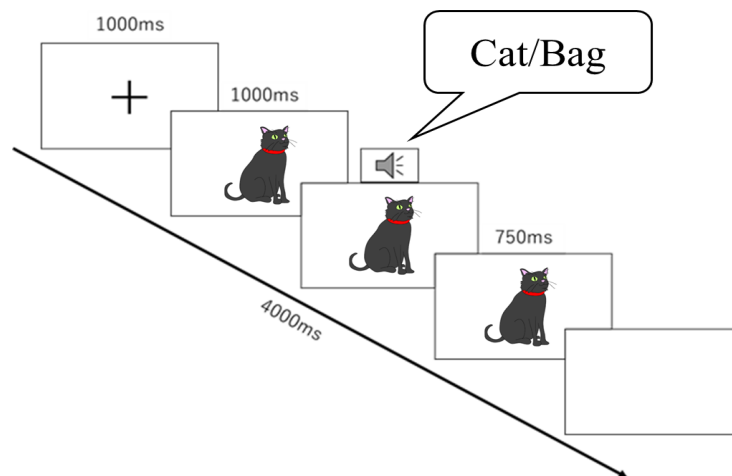


Figure 4. Trial sequence of the ERP experiment.

3.3.1.4 EEG recording

EEGs were recorded from 25 active electrodes attached to an elastic cap (ActiCap, Brain Products, Germany) according to the extended 10-20 system (Jasper 1958) (see Figure 5). Additional electrodes were

placed on each earlobe as references and between FPz and Fz as the ground. Four electrodes were placed at the outer canthus of each eye and above and below the left eye. Scalp electrodes were referenced online to the electrode on the left earlobe. All electrodes were re-referenced online to the averaged value from left and right earlobes. Impedance at all sites were kept at less than 10 k Ω throughout the experiment. The EEGs were amplified using Brain Amp (Brain Products, Germany) with a bandpass of 0.016 to 100 Hz, digitized at a 1000-Hz sampling rate.

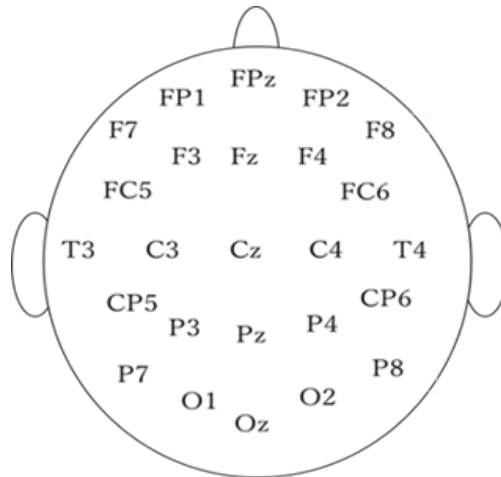


Figure 5. Electrode positions displayed on a schematic head seen from above.

3.3.1.5 EEG data analysis

The data were filtered offline with a 0.1-30Hz bandpass filter. Epochs of 1600 ms (consisting of 200 ms before the stimulus onset and 1400 ms after the stimulus onset) were extracted, and a pre-stimulus period of 200 ms was used in baseline adjustment. The ERPs of interest were quantified as mean amplitudes measured in specific time windows. According to the previous reports and a visual inspection of the data, two time windows were set: 300-600 ms as the N400 time window and 700-1100 ms as the LPC time window. As for N400 time window, repeated-measures analysis of variance (rmANOVA) were run with Condition

(match/mismatch), Hemisphere (left/right), Region (Anterior: FP1, FP2, F7, F8, F3, F4, Central: FC5, FC6, C3, C4, CP5, CP6, Posterior: O1, O2, P7, P8, P3, P4) as within-subject factors, and Group (Group 1: low exposure, 2: high, 3: extreme) as a between-subject factor. As for the LPC time window, we focused on the posterior region (O1, O2, P7, P8, P3, P4) only, because the LPC was limited to this region (the more anterior sites had a negative ERP in all groups, which is irrelevant to our interests here). Along with the N400 time window, we ran a rmANOVA with Condition (match/mismatch), Hemisphere (left/right), Region (O1-O2, P7-P8, P3-P4) as within-subject factors, and Group as a between-subject factor.

3.3.2 Results

3.3.2.1 Overview

The grand-average ERP waveforms of all participants from all three groups comparing the match and mismatch conditions in Japanese are shown in Figure 6. The grand-average waveforms for each group are shown in Figure 7 to 9. It is clear that the mismatch condition (blue/black line) evoked a large N400 at many sites and an LPC at posterior sites. These results are consistent with the previous literature (e.g., Juottonen et al. 1996; Ojima Matsuba-Kurita, Nakamura, and Hagiwara 2011a). As for the group differences, Figure 10 shows t-maps representing t-values for subtraction data (mismatch minus match condition) from 100 to 1400 ms of all three groups. In terms of the N400 component, although all groups show an N400 component around 300-800 ms, Group 1's N400 seems to be distributed in smaller areas than the other groups. The t-maps show that the negativity around 400-800 ms in Group 1 appeared at the central to anterior regions, whereas Group 2 and 3 show the negativity at broader areas. This presumably reflects the large LPC at the posterior region after the N400 in Group 1. It seems that the LPC in Group 1 is larger and longer than in Group 2 and Group 3. Qualitative differences among the three groups can be found in the LPC time window (700-1100 ms), which can be seen in Figure 10.

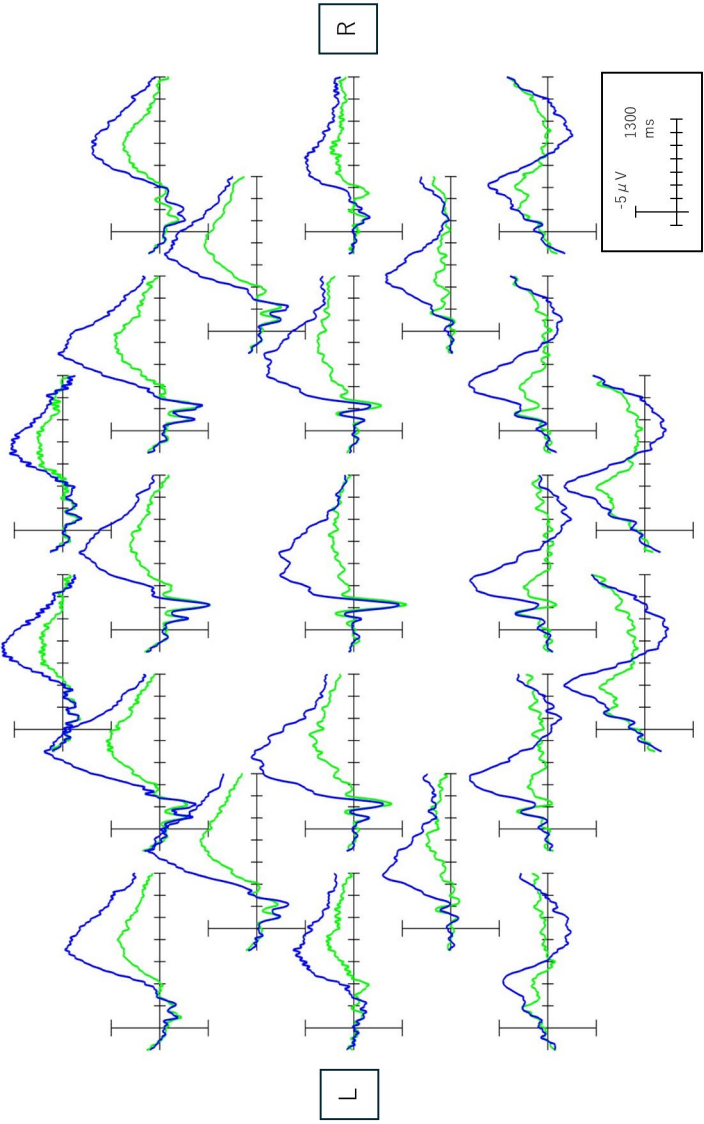


Figure 6. Grand-average ERP wave forms for semantic-mismatch (blue/black lines) and -match (green/gray lines) at all scalp sites by all three groups. L and R indicate left and right side of the head, respectively.

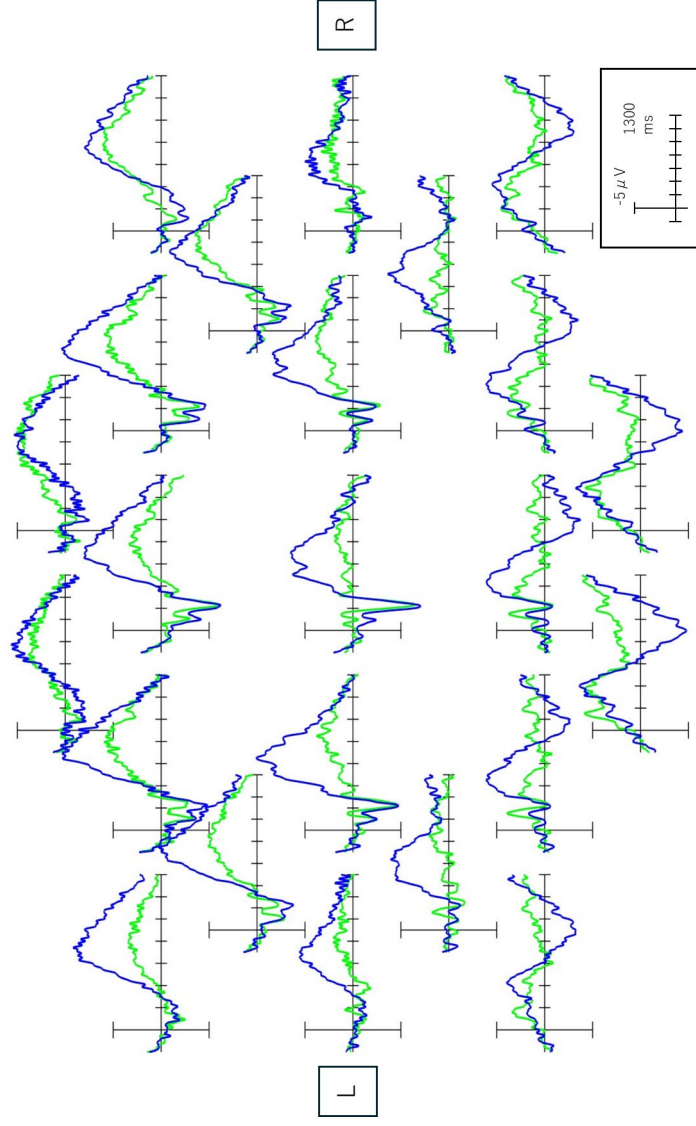


Figure 7. Grand-average ERP wave forms for semantic-mismatch (blue/black lines) and -match (green/gray lines). L and R indicate left and right side of the head, respectively.

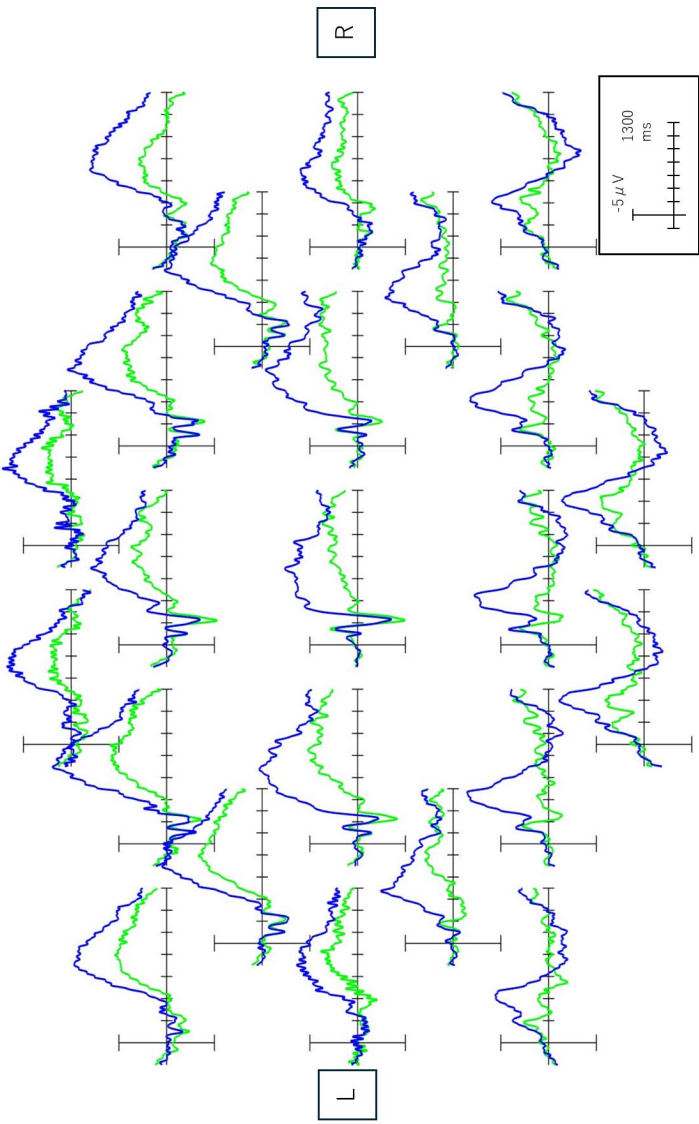


Figure 8. Grand-average ERP wave forms for semantic-mismatch (blue/black lines) and -match (green/gray lines) at all scalp sites by Group 2 (high exposure group). L and R indicate left and right side of the head, respectively.

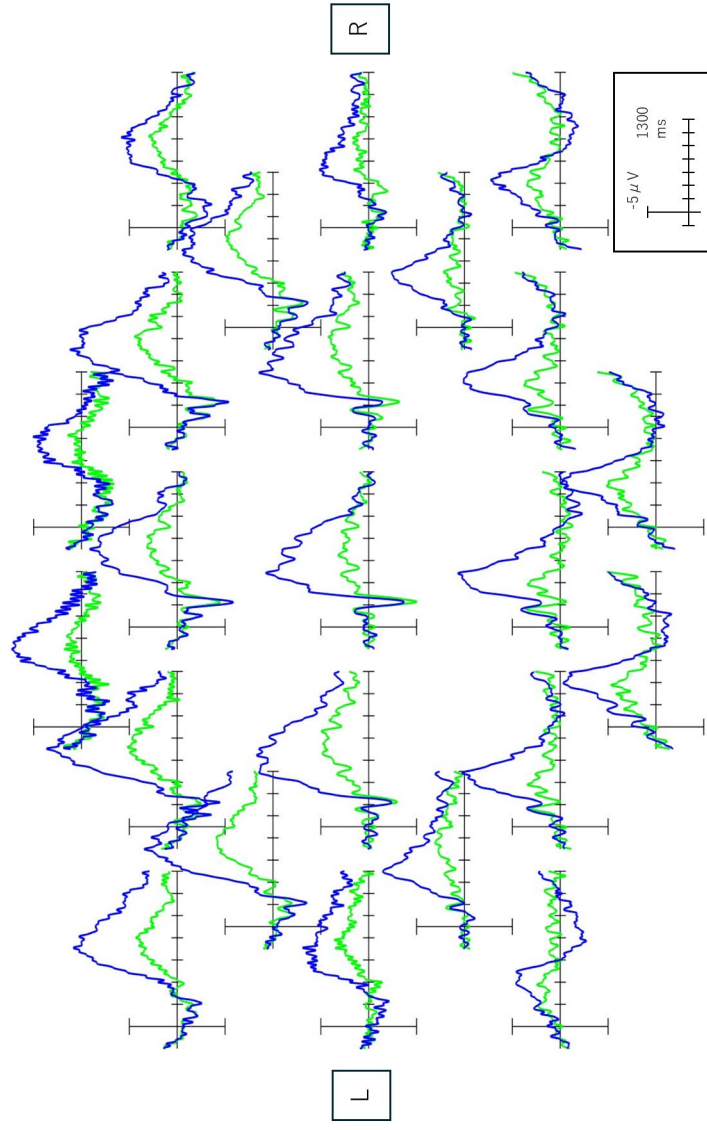


Figure 9. Grand-average ERP wave forms for semantic-mismatch (blue/black lines) and -match (green/gray lines) at all scalp sites by Group 2 (extreme exposure group). L and R indicate left and right side of the head, respectively.

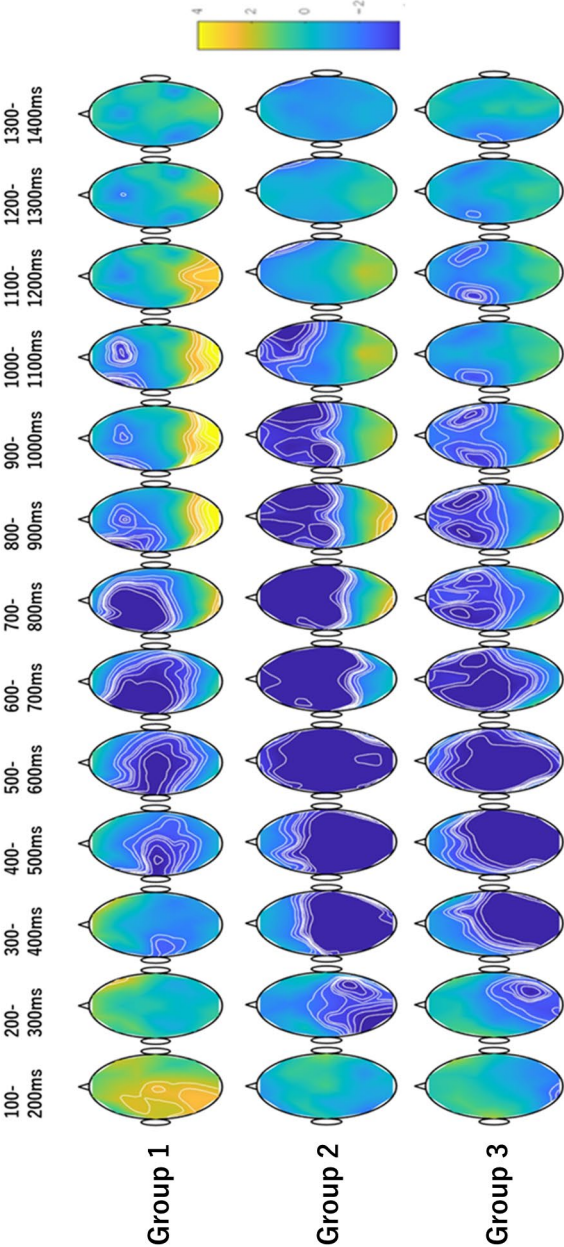


Figure 10. T -maps representing t -values obtained for subtraction data (mismatch mismatch condition) of the three groups. Mean amplitudes in 100-ms intervals from 100 ms up to 1400 ms were used to calculate the t -values. Thick colors inside the white lines indicate the regions where the match and mismatch conditions differed significantly.

3.3.2.2 N400 time window (300-600ms)

To investigate the N400 effect in the three groups, we ran a rmANOVA. The statistical results revealed a significant condition x group interaction ($F(2, 40) = 3.310, p = .047, \eta_p^2 = .142$), suggesting that the differences between two conditions could be different among the three groups. In addition, the Condition main effect ($F(2, 40) = 65.876, p < .001, \eta_p^2 = .622$) and the condition x region interaction ($F(2.6, 104.003) = 16.185, p < .001, \eta_p^2 = .288$) were found to be significant. None of the other interactions were significant (e.g., condition x region x group; $F(5.2, 104.003) = .674, p = .650, \eta_p^2 = .033$; condition x hemisphere x group; $F(2, 40) = .086, p = .917, \eta_p^2 = .004$).

We further analyzed the simple main effect of Condition in each group. The results show statistical significance in all groups, Group 1 ($F(1, 13) = 4.824, p < .047, \eta_p^2 = .273$), Group 2 ($F(1, 14) = 69.918, p < .001, \eta_p^2 = .833$), and Group 3 ($F(1, 13) = 28.831, p < .001, \eta_p^2 = .689$). This result supports that all groups exhibit a larger N400 component for the mismatch condition. A visual inspection of the subtraction data (mismatch-match) indicates that the amplitudes of N400 become larger in the order of Group 3, 2, and 1 (see Figure 11).

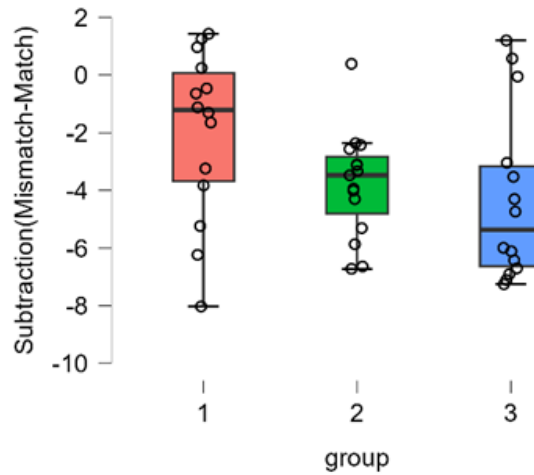


Figure 11. Box plots illustrating mean subtraction (mismatch-match) amplitudes in the N400 time window (300-600 ms).

3.3.2.3 LPC time window (600-1100ms)

As for the LPC time window, based on a visual inspection of the waveforms (Figures 6-9), we focused on the posterior area (O1-O2, P7-P8, P3-P4). The statistical results of rmANOVA show the significance of the condition x region x group interaction ($F(3.523, 70.450) = 2.527, p = .047, \eta_p^2 = .112$). This suggests that scalp distribution of LPC could be different among the groups. The statistical significance of the Condition main effect ($F(1, 40) = 13.618, p < .001, \eta_p^2 = .254$) and the condition x region interaction ($F(2, 80) = 8.836, p < .001, \eta_p^2 = .181$) were also shown.

Further analyses of the simple main effect of Condition in each group and in each region are summarized in Table 7. The results show that all regions are significant in Group 1 (low exposure), only O1/O2 is significant in Group 2 (high), and only P7/P3 is significant in Group 3 (extreme). This indicates that LPC is found in broader areas in Group 1 than the other groups, which is consistent with the visual inspection of the waveforms and the maps (Figures 6-107). Although no statistical tests were conducted to directly compare the groups, the amplitudes of LPC in Group 1 seem to be the largest of all based on the subtraction data in Figure 12.

Table 7. Results of simple main effects of condition.

Group	Region	<i>F</i>	<i>p</i>	η_p^2
1	O1/O2	23.949	<.001	.230
	P7/P8	11.808	.004	.129
	P3/P4	4.819	.047	.057
2	O1/O2	5.011	.042	.059
	P7/P8	2.823	.115	.034
	P3/P4	1.364	.262	.017
3	O1/O2	1.164	.300	.014
	P7/P8	5.034	.043	.059
	P3/P4	0.094	.764	.001

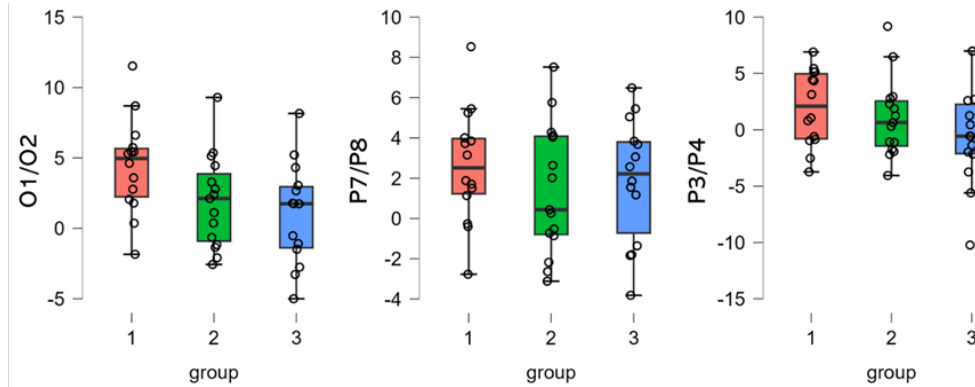


Figure 12. Box plots illustrating mean subtraction (mismatch-match) amplitudes in the LPC time window (600-1100 ms) separated by the region (O1/O2, P7/P8, and P3/P4).

3.3.3 Summary

All groups show the N400 component in the 300-600 ms time window. The results show that the N400 component appeared at broad areas of the scalp. From the results of the subtraction data (Figure 8), it seems that the N400 amplitude in Group 1 is smaller than that in the other groups. This might reflect the existence of a large LPC in Group 1 in the later time window.

In terms of the LPC, all three groups show the LPC in the 600-1100 ms time window in at least one region. One important observation is that the scalp distribution of the LPC in Group 1 is broader than that in Group 2 and 3. As with the N400 subtraction data, it seems that the mean amplitude of the LPC in Group 1 is larger than that in the other groups (Figure 12).

4. GENERAL DISCUSSION

Consistent with the three different experiments, it can be concluded that longer exposure seems to be a prerequisite for children's high language-learning ability to manifest itself regardless of their L2 or L1.

Based on Experiment 1, older learners seem to have an advantage for L2 learning in the short term, that is, up to 5000 hours of exposure. This fact has been previously discussed in many studies (e.g., Muñoz 2006). The children who participated in this project receive L2 input from viewing TV or DVD, listening to CDs, and natural conversations with peers and online teachers, rather than receiving explicit instruction of the rules. This type of learning is categorized as a kind of implicit learning which usually takes a longer time to produce advantages (DeKeyser 2000). Unlike this, the adult learners in Experiment 1 were typical EFL learners who received explicit instruction, which is potentially suitable for them. Thus, assuming that the adults and the low-exposure children received an almost equal amount of English exposure, the adult learners showed better performance; children whose exposure was below 5000 hours could not produce English utterances in the particular task that we employed. Though not showing statistical significance compared with adults, children who have received massive exposure (over 5000 hours) seem to have high English-speaking abilities. This might be one benefit of starting L2 acquisition in early childhood, which could lead to high ultimate attainment (e.g., Long 2013). However, it should be noted that large individual variation was also found in the children with massive input in the current study. One speculation for this variation is that the type of learning experienced by children might have affected the results. In the current study, the amount of exposure includes both production and perception. Children who could not speak in our task might not have had enough experience to produce English but have had enough time to perceive it. Children's aptitude for language learning could also explain this result, as Roehr-Brackin and Tellier (2019) suggested. Leaving these as our future investigations, research on the role of aptitude in child L2 acquisition is also warranted (see Beccia 2020). It can be concluded that even in an EFL context, if children have access to a massive amount of L2 input, it may be possible for them to attain high L2 speaking abilities.

So far, we have obtained no evidence that extensive "Home English" experience leads to lower academic Japanese abilities compared with normal children with no or little English contact. This might be because those learners receive enough L1 exposure when they attend school or play with their Japanese friends. Thus, exposure to L1 input is perhaps

maintained for them. However, children in international school seem to show a different pattern. They use English as much as possible in their school life, and presumably their classmates or close friends also speak English on a daily basis. As a result, they may not be able to receive enough L1 input. Of course, if children go to an international school which provides enough Japanese classes, they might retain their L1 knowledge. This cannot be decisively concluded because the sample size is limited in this preliminary study. Thus, a clearer conclusion is left for further investigations.

Finally, the neurolinguistic results in the current study highlighted the possibility that Japanese children with massive English exposure may differ from those with limited exposure in the neural processing of basic Japanese words. Although children with little English exposure clearly showed an LPC at all posterior sites, the LPC of the other two groups with high or extremely high exposure was more limited in terms of scalp distribution. However, we cannot conclude that this pattern was caused by children's L2 learning, because the data may simply reflect a correlational relationship, rather than a causal relationship. In other words, children with high or extremely high English exposure might have had distinct brain activity patterns even before they began to learn English. Further investigations are necessary to settle down this matter.

Potentially, our result might provide some implications on child L2 learning. However, current results do not support the "earlier is better" argument. An important point based on the result from Experiment 1 is that if children and adults received the same amount of input (around 2000 hours of exposure or less in Figure 1 and 2), the adult learners show better performance than the child learners. From this point, child learners cannot be regarded as having clear advantage over adults when they have received around 2,000 hours of exposure or less. Once children have received a massive amount of input, it is expected that they can reach higher target-language ability (i.e., fluent speech in Experiment 1 and nativelike brain pattern in Experiment 3). However, we could not clearly support the child advantage at least in this study. This is because it could also be possible that adult L2 learners with an equally massive amount of input could reach a similar stage. This possibility should be investigated further. What we can say according to the results from our

three experiments is that a massive amount of input is essential for both adult and child learners to attain high target-language ability.

5. CONCLUSION

In this project including three experiments, we investigated children who have received massive L2 input in their home or in a special case of school context. We drew implications for learning L2 in a foreign language context. To attain high L2 proficiency, it is important to maintain the target-language input on a daily basis. The longer exposure learners receive, the higher performance they could reach ultimately. We also obtained no evidence that extensive input in a Home-English environment leads to lower academic L1 abilities. However, neurolinguistic data from Experiment 3 revealed distinct brain activity patterns for L1 lexical processing among children with varying amounts of L2 exposure. The total amount of exposure to L1 might be a key factor to sustain and develop L1 abilities, but further psycholinguistic and neurolinguistic investigations are needed to interpret the results of the current experiments. We hope to further investigate Japanese children who have received massive English input in an EFL context, by employing a longitudinal design in the future.

REFERENCES

- Baumert, Jürgen, Johanna Fleckenstein, Michael Leucht, Olaf Köller, Jens Möller. 2020. The long-term proficiency of early, middle, and late starters learning English as a foreign language at school: A narrative review and empirical study. *Language Learning* 70:1091-1135.
- Beccia, Ashley. 2020. The role of language aptitude in children's rate and route of L2 acquisition: A direction for future research in Child SLA. *Working papers in Applied Linguistics & TESOL* 20:26-28.
- Butler, Yuko Goto. 2015. English language education among young learners in East Asia: A review of current research (2004-2014). *Language Teaching* 48:303-342.
- Cadierno, Teresa, Mikkel B. Hansen, Jørgen T. Lauridsen, Søren W. Eskildsen, Katalin Fenyvesi, and Signe Hannibal Jensen. 2022. A multi-factor study of the development of English receptive skills by young Danish children. *Bilingualism: Language and Cognition* 25:566-578.
- DeKeyser, Robert M. 2000. The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition* 22:499-533.
- DeKeyser, Robert M., Iris Alfi-Shabtay, and Dorit Ravid. 2010. Cross-linguistic evidence for the nature of age effects in second language acquisition. *Applied Psycholinguistic* 31:413-438.
- Granena, Gisela, and Michael H. Long. 2013. Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research* 29:311-343.
- Gallo, Federico, Beatriz Bermudez-Margaretto, Yury Shtyrov, Jubin Abutalebi, Hamutal Kreiner, Tamara Chitaya, Anna Petrova, and Andriy Myachykov. 2021. First language attrition: What it is, what it isn't, and what it can be. *Frontiers in Neuroscience* 15:686388.
- Ferjan Ramirez, Naja., and Patricia Kuhl. 2017. Bilingual baby: Foreign language intervention in Madrid's infant education centers. *Mind, Brain, and Education* 11:133-143.
- Ferjan Ramirez, Naja., and Patricia Kuhl. 2020. Early second language learning through SparkLingTM: Scaling up a language intervention in infant education centers. *Mind, Brain, and Education* 14:94-103.
- Jaekel, Nils, Michael Schurig, Merle Florian, Markus Ritter. 2017. From Early Starters to Late Finishers? A Longitudinal Study of Early Foreign Language Learning in School. *Language Learning* 67:631-664.
- Jasper, Herbert H. 1958. The ten-twenty system of the International Federation. *Electroencephalography and Clinical Neurophysiology* 10:371-375.
- Johnson, Jacqueline S. 1992. Critical period effects in second language acquisition: The effects of written versus auditory materials on the assessment of grammatical competence. *Language Learning* 42:217-248.

- Johnson, Jacqueline S., and Elissa L. Newport. 1989. Critical period effects in second language: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 21:60-99.
- Juottonen, Kirsi, Antti Revonsuo, and Heikki Lang. 1996. Dissimilar age influences on two ERP waveforms (LPC and N400) reflecting semantic context effect. *Cognitive Brain Research* 4:99-107.
- Krashen, Stephen D., Michael A. Long, and Robin C. Scarcella. 1979. Age, rate and eventual attainment in second language acquisition. *TESOL Quarterly* 9:573-582.
- Kutas, Marta, and Kara D. Federmeier. 2000. Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Science* 4:463-470.
- Kutas, Marta, and Kara D. Federmeier. 2011. Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology* 62:621-347.
- Lenneberg, Eric H. 1967. *Biological foundations of language*. New York: John Wiley and Sons.
- Lichtman, Karen. 2016. Age and learning environment: Are children implicit learners? *Journal of Child Language* 43:707-730.
- Lichtman, Karen. 2020. What about fluency? Implicit vs. explicit training affects artificial mini-language production. *Applied Linguistics* 42:668-691.
- Lindgren, Eva, and Carmen Muñoz. 2013. The influence of exposure, parents, and linguistic distance on young European learners' foreign language comprehension. *International Journal of Multilingualism* 10:105-129.
- Long, Michael H. 2013. Maturational constraints on child and adult SLA. In *Sensitive periods, language aptitude, and ultimate L2 attainment*, eds. Gisela Granena, and Michael H. Long, pp.3-42. Amsterdam: John Benjamin.
- McLaughlin, Judith, Lee Osterhout, and Albert Kim. 2004. Neural correlates of second-language word learning: minimal instruction produces rapid change. *Nature Neuroscience* 7:703-704.
- Muñoz, Carmen. 2006. The effect of age on foreign language learning: The BAF project. In *Age and the rate of foreign language learning*, ed. Carmen Muñoz, pp.1-40. Bristol: Multilingual Matters.
- Muñoz, Carmen. 2014. Contrasting effects of starting age and input on the oral performance of foreign language learners. *Applied Linguistics* 35:463-482.
- Nakamura, Janice. 2021. English parenting for Japanese parents: A critical review of advice in self-help books for raising bilingual children in Japan. *English Today* 39:47-52.
- Nambu, Kuki. 2015. *How are Japanese university students' English speaking abilities related to their learning experience?* Shiga: Shiga University B. Ed. thesis.
- Ojima, Shiro, Hiroko Matsuba-Kurita, Naoko Nakamura, and Hiroko Hagiwara. 2011a. The acceleration of spoken-word processing in children's native-language acquisition. An ERP cohort study. *Neuropsychologia* 49:790-799.
- Ojima, Shiro, Hiroko Matsuba-Kurita, Naoko Nakamura, Takahiro Hoshino, and Hiroko Hagiwara. 2011b. Age and amount of exposure to a foreign language during

- childhood: Behavioral and ERP data on the semantic comprehension of spoken English by Japanese children. *Neuroscience Research* 70:197-205.
- Ojima, Shiro, Naoko Nakamura, Hiroko Matsuba-Kurita, Takahiro Hoshino, and Hiroko Hagiwara. 2011c. Neural correlates of foreign-language learning in childhood: A 3-year longitudinal ERP study. *Journal of Cognitive Neuroscience* 23:183-199.
- Pallier, Christophe, Stanislas Dehaene, Jean-Baptiste Poline, Denis LeBihan, Anne-Marie Argenti, Emmanuel Dupoux, and Jacques Mehler. 2003. Brain imaging of language plasticity in adopted adults: Can a second language replace the first? *Cerebral Cortex* 13:155-161.
- Patkowski, Mark S. 1980. The sensitive period for the acquisition of syntax in a second language. *Language Learning* 30:449-472.
- Peters, Elke, Ann-Sophie Noreillie, Kris Heylen, Bram Bulté, and Piet Desmet. 2019. The impact of instruction and out-of-school exposure to foreign language input on learners' vocabulary knowledge in two languages. *Language Learning* 69:747-782.
- Perani, Daniela, Jubin Abutalebi, Eraldo Paulesu, Simona Brambati, Paola Scifo, Stefano F. Cappa, and Ferruccio Fazio. 2003. The role of age of acquisition and language usage in early, high-proficient bilinguals: An fMRI study during verbal fluency. *Human Brain Mapping* 19:170-182.
- Pfenninger, Simone E., and David Singleton. 2016. Affect trumps age: A person-in-context relational view of age and motivation in SLA. *Second Language Research* 32:311-345.
- Prosic-Santovac, Dnija. 2016. Popular video cartoons and associated branded toys in teaching English to very young learners: A case study. *Language Teaching Research* 21:568-588.
- Psychology Software Tools. 2011. *E-Prime™* [Computer software]. Pittsburgh, PA: Psychology Software Tools.
- Roehr-Brackin, Karen, and Angela Tellier. 2019. The role of language-analytic ability in children's instructed second language learning. *Studies in Second Language Acquisition* 41:1111-1131.
- Scheffler, Pawel, Christian Jones, and Anna Domńska. 2020. The Peppa Pig television series as input in pre-primary EFL instruction: A corpus-based study. *International Journal of Applied Linguistics* 31:3-17.
- Terasawa, Takunori. 2018. *Learning English in Japanese: Myths and realities*. Melbourne: Trans Pacific Press.
- Thomai, Alexiou. 2015. Vocabulary uptake from Peppa pig: A case study of preschool EFL learners in Greece. In *Current issues in second/foreign language teaching and teacher development: Research and Practice*, eds. Christina Gitsaki, and Thomai Alexiou, pp.285-301. Newcastle: Cambridge Scholars Publishing.
- Werker, Janet F., John H. Gilbert, Keith Humphrey, and Richard C. Tees. 1981. Developmental aspects of cross-language speech perception. *Child Development* 52:349-355.

Atsushi Miura, Natsumi Shibata, Shiro Ojima

- Werker, Janet F., and Richard C. Tees. 1984. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7:49-63.
- Yeni-Komashian, Grace H., James E. Flege, and Serena Liu. 2000. Pronunciation proficiency in the first and second language of Korean-English bilinguals. *Bilingualism: Language and Cognition* 3:131-149.

Atsushi Miura
Graduate School of Fundamental Science and Engineering
Waseda University
3-4-1, Okubo, Shinjuku-ku, Tokyo 169-8555, Japan
a.miura@asagi.waseda.jp

Natsumi Shibata
School of Business Administration
Senshu University
2-1-1, Tama-ku, Kawasaki-city, Kanagawa 214-0033, Japan
thz3761@senshu-u.jp

Shiro Ojima
Center for English Language Education, Faculty of Science and Engineering (CELESE)
Waseda University
3-4-1, Okubo, Shinjuku-ku, Tokyo 169-8555, Japan
ojima.s@waseda.jp

廣泛接觸英語對日本兒童的影響：語言能力和大腦功能

三浦篤史、柴田奈津美、尾島司郎
早稻田大學

這項研究初步調查了大量接觸第二語言（L2）對兒童 L2 能力以及對他們第一語言（L1）能力的潛在影響。我們進行了三個不同的實驗。實驗 1 比較了兒童和成人的 L2 口語表現，結果表明，如果兒童接受了大量的 L2 輸入，即使是在 EFL 環境下，早期 L2 學習依然可以成功。在實驗 2 中，我們考察了兒童的 L1 學習能力，結果表明與暴露量無關，並推測 L2 的學習安排可能會顯著影響 L1 的學習能力。實驗 3 採用神經語言學技術（事件相關電位或 ERPs），探討了兒童的 L1 詞彙加工，廣泛接受 L2 輸入的兒童可能表現出不同的神經詞彙加工模式。本文探討了在外語環境中早期 L2 學習的一些啟示。

關鍵字：兒童第二語言習得、英語作為外語、口語、一語學術能力、詞彙加工

**A FOLLOW-UP STUDY
ON THE POSITIVE CROSS-LINGUISTIC INFLUENCE
OF SENTENCE-FINAL PARTICLES ON ACQUISITION:
LEARNER, STRUCTURE,
AND LANGUAGE PROFICIENCY***

Shanshan Yan
Peking University

ABSTRACT

This study examines the positive cross-linguistic influence (CLI) in three related studies investigating the acquisition of two Mandarin Chinese sentence-final particles (the SFP *le* and the question SFP *ba*). Complementary and combined comparisons, which did not surface in the three studies reviewed, were conducted to obtain results in terms of the effects of CLI, particularly positive CLI demonstrated by learners (heritage vs. L2 learners, and learners vs. native speakers), language proficiency levels, as well as structures (across and within conditions). The findings reveal that heritage learners with early and richer exposure to the target language were not always advantaged in terms of acquisition compared to L2 learners. A transition from negative to positive CLI can be detected among L2 learners at advanced levels. Positive CLI exhibited contrasting effects on the acquisition of the different features of structures. Thus, a more holistic view of the manifestations of positive CLI and a refined classification of CLI are called for.

Keywords: positive cross-linguistic influence, sentence-final particle, heritage learner, L2 learner, structures, language proficiency

* This work was supported by the National Social Sciences Project “Acquisition of Modal Expressions by Preschool Mandarin-English Bilingual Children (24CYY081)” and Beijing Social Sciences Project “Bilingual Children’s Acquisition of Mandarin Sentence-final Particles (24DTR082)”.

1. INTRODUCTION

Cross-linguistic influence (CLI) has long been a subject of research interest, particularly with regard to second language acquisition (SLA) (Lago et al. 2021; Odlin 2022). CLI is used to demonstrate that the grammatical properties of prior or simultaneously acquired language(s) will influence subsequently acquired ones, a phenomenon that is also referred to as the effect of language transfer (e.g., L1 transfer). Two types of CLI are particularly prominent in SLA, namely negative CLI and positive CLI, with the former referring to situations in which L1 and L2 do not have structural equivalents or L1 has substantially different equivalents from L2, and the latter referring to scenarios where L1 and L2 do have similar structures. It has been argued that positive CLI facilitates L2 acquisition, while negative CLI poses difficulties for learners. Research in the area of SLA has hitherto largely focused on the negative aspects of CLI (Odlin 2001; Yuan 2004; Yuan and Zhao 2005; Jackson and Dussias 2009; among others), with little attention to the positive side (Odlin 1989; Yip et al. 2018; Yan et al. 2023). Therefore, to advance our understanding of specific manifestations of CLI, in this paper we review the following three studies and compare data regarding CLI that has not been explicitly discussed:

Study 1: Yan (2020). Syntactic and discourse features in Chinese **heritage grammars**: A case of acquiring features in the Chinese **sentence-final particle *ba***. *Languages*, 5(2), 26.

Study 2: Yan and Yuan (2022). Asymmetric form-meaning mappings in **L2 acquisition** of Chinese **sentence-final particles *ba***. *Journal of Chinese Linguistics*, 50(1), 1–38.

Study 3: Yan, Mai, and Zhao (2023). Positive cross-linguistic influence in representation and processing of **sentence-final particle *le*** by **L2 and heritage learners** of Chinese. *Frontiers in Psychology*, 14, 1145493.

The rationale for selecting these three studies mainly concerns their comparative flexibility across various levels. The studies were all part of the same larger project and therefore they share a participant pool, which

enables the comparison of learner language proficiency levels and learning backgrounds. Furthermore, a detailed examination of the differences and similarities concerning the target SFPs is possible since the same linguistic structures are assessed and similar tasks are performed across the three studies. Most importantly from a practical point of view, however, data were easily accessible from all three studies, thus enabling further analysis. In addition, two of the three studies (except Yan et al. 2023) adopted different theoretical frameworks and thus do not afford direct comparative findings related to CLI, which further motivates our current study.

In particular, we will scrutinize CLI in the acquisition of Mandarin Chinese sentence-final particles (SFPs) by English-speaking learners. Chinese SFPs, a non-existent functional category in the English language, provide us with a useful test ground for a close investigation of CLI. On the one hand, negative CLI can be tested on the basis that no L1 equivalents are available. The grammatical properties of some SFPs can be identified from various constructions that are substantially different from Mandarin Chinese SFPs (i.e., the SFP *ba* and English tag questions; see section 2.1), allowing us to assess the extent and ways in which CLI, particularly negative CLI, manifests in learners' output. On the other hand, several SFPs do have equivalents (i.e., the SFP *le* and the English adverb *already*; see section 2.2), which also allows us to examine the facilitating role of positive CLI.

In addition, the studies compare two types of learners, L2 and heritage learners, on a variety of levels (e.g., proficiency levels, performance in different tasks, and comparisons with native speakers). L2 learners, typically those who begin learning a second language after puberty, follow a sequence where their first language precedes the acquisition of their second. This allows us to examine CLI in terms of "influences from a previously acquired language." On the other hand, heritage learners have been exposed to two languages since a young age, allowing us to examine how "(nearly) simultaneously acquired languages" influence subsequently acquired languages. Details of the learners are summarized in section 4.1. With these comprehensive comparisons between L2 and heritage learners in terms of their acquisition of Mandarin Chinese SFPs, we expect to develop a more holistic and detailed understanding of CLI.

2. PROPERTIES OF THE SENTENCE-FINAL PARTICLES *LE* AND *BA*

Mandarin Chinese has a variety of SFPs instantiating syntactic, semantic, and discourse functions (Chao 1968; Li and Thompson 1981; Tang 2010; Paul 2015; Pan 2019; Paul and Yan 2022). Among these, the SFP *le* signals a semantic aspect associated with a “change of state,” requiring noun phrases (NPs, see e.g. (1a) and (1b)) or verb phrases (VPs, e.g. (2a) and (2b)) to co-occur when denoting higher/upward points:

“The change must take place from an earlier/lower point to a later/higher point along the associated scale, rather than vice versa” (Xiang 1998; Xing 2001; as summarized in Yan et al. 2023:3).

- (1) Higher/lower NP in SFP *le* sentences
 - a. Xiǎowáng shì **dà háizi le**.
Xiaowang be big child LE
‘Xiaowang is a big child already.’
 - b. */?Xiǎowáng shì **xiǎo háizi le**.
Xiaowang be small child LE
*/?‘Xiaowang is a small child already.’
- (2) Upward-/downward-entailing duration in SFP *le* sentences (slightly modified examples from Soh (2009:638–641))
 - a. Wǒ zài zhèr zhù le **zhìshǎo wǔ nián le**.
1SG at here live PERF at-least five CL LE
‘I have lived here for at least five years already (which was not the case before).’
 - b. */?Wǒ zài zhèr zhù le **búdào wǔ nián le**.
1SG at here live PERF less-than five CL LE
*/?‘I have lived here for less than five years already.’

On the other hand, the SFP *ba* used in question sentences, along with its use in suggestion sentences, encompasses both syntactic and discourse

functions (Yan 2020; Yan and Yuan 2022). The question SFP *ba* is characterized by syntactic features ([Q] and [-wh]) and plays a syntactic role in converting a declarative sentence (without the SFP *ba*) into a question (compare examples (3a) and (3b)), prohibiting the appearance of *wh*-phrases (see examples (3c) and (3d)). The type of question containing the SFP *ba* mainly aims to seek confirmation of assumptions that a speaker has made (i.e., the [confirmation seeking] discourse feature). Meanwhile, the suggestion SFP *ba* mainly carries discourse features so as to soften the speaker's tone (i.e., for the [suggestion] feature, compare examples (4a) and (4b)).

(3) question SFP *ba*: examples (a) and (b) from Yan (2020:4) and Yan and Yuan (2022:3)

- | | |
|--|---|
| a. Tā huì kāi chē.
3SG can drive car
'He(She) can drive
a car.' | b. Tā huì kāi chē ba ?
3SG can drive car BA
'He(She) can drive a car, can't
he(she)/right?' |
| c. Shéi huì kāi chē?
who can drive car
'Who can drive?' | d. *Shéi huì kāi chē ba ?
who can drive car BA
Intended: 'Who can drive?' |

(4) suggestion SFP *ba*: examples (a) and (b) from Chao (1968:807)

- | | |
|--|--|
| a. Kuài diǎnr zǒu!
hurry little go
'Hurry up!' | b. Kuài diǎnr zǒu ba !
hurry little go BA
'Hurry up, please!' |
|--|--|

The English language lacks a functional category equivalent to SFPs that serve various functions. Instead, it employs syntactic means, as exemplified in various constructions such as yes-no questions (employing auxiliary-subject inversions) and, *let's/we'd better/shall* suggestions (refer to the three studies for detailed syntactic comparisons). This applies to most Mandarin SFPs, one of which is the SFP *ba*. However, a few Mandarin SFPs can be identified to have English equivalents, for example the SFP *le* and the English adverb *already*. From a CLI perspective,

particular Mandarin SFPs and their English equivalents are summarized in Figure 1.

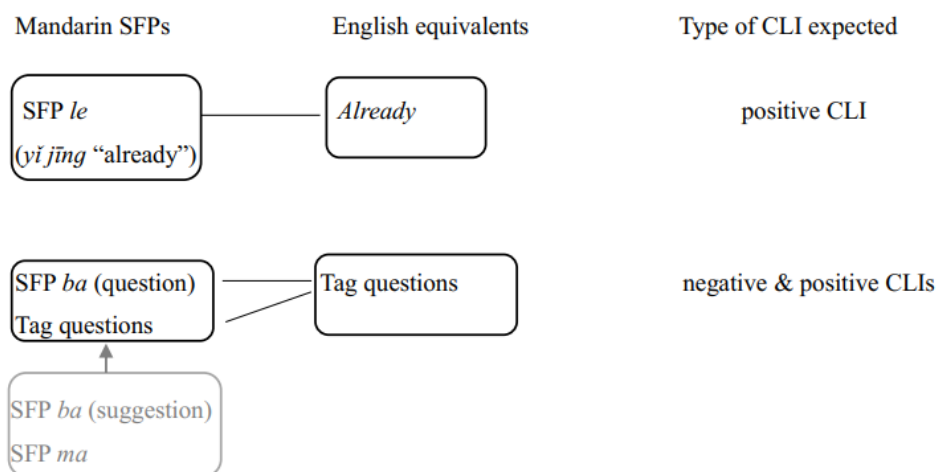


Figure 1. Types of CLI expected when a comparison is conducted between particular Mandarin Chinese SFPs and their English equivalents (“——” represents a mapping relationship, “→” represents an influence from A to B)¹

The direct correspondence between the Mandarin SFP *le* under study and the English *already* suggests a positive CLI, as learners’ L1 (English) contains an equivalent expression in L2 (Mandarin). Conversely, for the question SFP *ba* in Chinese, English-speaking learners might experience negative CLI due to the lack of a direct equivalent in their L1 (English). However, there is an English structure (i.e., English tag questions) with all same features as the question SFP *ba*, which could be seen as a type of positive CLI from a broader perspective. Mapping the question SFP *ba*

¹ While Mandarin Chinese has a lexical equivalent to the English *already* in *yǐ jīng*, this does not impact our current discussion on CLI, which centers on identifying equivalents to L2 Mandarin SFPs in L1 English. Furthermore, upon reviewing the literature, Yan et al. (2023) identified the English *already* as the L1 equivalent of the Mandarin SFP *le* (e.g., Szeto et al. 2017). Even though we cannot exclude the influence of the mapping between English *already* and *yǐ jīng*, our research on mapping issues will still be illuminative.

onto multiple comparable English structures therefore requires substantial effort (for a discussion on L1-L2 one-to-many mapping, see Yan and Yuan 2022), and depends significantly on recognition of the L2 input, as well as making distinctions between homomorphous SFPs *ba* (question and suggestion forms) and similar SFPs (SFP *ma* and SFP *ba* in questions, Yan and Yuan 2022; also see the gray square in Figure 1). In this sense, the complex L2 input might disguise the mapping and prevent positive CLI effects.

However, once the mapping is established the acquisition of the question SFP *ba* is predicted to be easy, since both the L1 and L2 correspondences show similar syntactic and discourse constraints. This suggests that exposure to L2 input is vital. Thus, L2 learners at higher proficiency levels, reflecting greater exposure and a stronger understanding of L2 grammars, along with heritage learners—who usually benefit from early and more extensive L2 exposure compared to other L2 learners—may exhibit differences in terms of CLI, especially positive CLI. This emphasizes the need for a detailed follow-up study on the differences between L2 and heritage learners as well as the potential benefit of situating our concerns in a wider context, in order to include the question SFP *ba* (Yan 2020; Yan and Yuan 2022) to complement findings from the SFP *le* (Yan et al. 2023).

3. RESEARCH QUESTIONS

As summarized in Figure 1, in terms of CLI, narrowly speaking, for an L2 target, positive CLI refers to a case where a close equivalent can be found in L1. The SFP *le* exemplifies “positive CLI,” since it can align with a lexical item such as a particle that also occurs at sentence ends in English. Conversely, the SFP *ba*, whether in question or suggestion form, must be expressed through various different grammatical structures in English due to the absence of SFPs in that language, and is therefore an instance of “negative CLI.” However, from a broader and more general perspective the identification of closely related structures in both L1 and L2, even if these are not in identical lexical forms, allows for the possibility of “positive CLI.” In this sense, features of the SFP *ba* could be found in

English, albeit realized in different ways, and L2 acquisition may in fact benefit from “positive CLI.”

CLI can manifest at a variety of different levels, such as the type of learner, type of task, and different grammatical structures (Yan et al. 2023). In terms of the main findings from the three studies under review, Yan (2020) and Yan and Yuan (2022) posit that the SFP *ba* in its suggestion form is easily acquired by both L2 and heritage learners. Therefore, in the current study we will mainly focus on the question form of the SFP *ba*, which might provide us with more evidence in discussing CLI. Also, bearing in mind the importance of comparison, tasks within the same modality yield more reliable results. The online task that was exclusively used for testing the SFP *le* by Yan et al. (2023) is not examined here due to its unique form. In the context of the expansion of research into the question SFP *ba*, here we are particularly interested in detailing what CLI manifestations exist and to what extent.

Our investigation is guided by three main research questions (RQs):

RQ1: Learner: Do heritage learners outperform their L2 peers?

RQ2: Language proficiency: At which language proficiency level does CLI exert the greatest influence on learners?

RQ3: Structure: Will linguistic structures make a difference to CLI?

Predictions: According to findings related to the SFP *le* in Yan et al. (2023), we predict that differences will be observed among learners and structures in this study. Specifically, heritage learners with early and richer exposure to Mandarin Chinese will outperform their L2 peers in all situations. Regarding structure, given that positive CLI is thought to be possible in the acquisition of Mandarin SFPs, we expect that its presence will facilitate this process, especially concerning the features of the question SFP *ba*. As for language proficiency, we propose that learners at more advanced levels will have a better grasp of the target language grammar than those at less proficient levels.

Table 1. Participant information (SD in brackets; combined information from Yan (2020), Yan and Yuan 2022, and Yan et al. 2023)

Groups	N	Average age	Onset age	Months studying Chinese	Months in China	Cloze test
L2 low-intermediate	20	24	19	48	24.9	16.6 (2.8)
L2 high-intermediate	20	22	19	39	11.8	24.6 (2.3)
L2 advanced	18	23	18	60	38.1	31.3 (1.8)
HL high-intermediate	16	24	—	199	13.5	25.3 (3.5)
HL advanced	19	22	—	187	23.2	33.2 (2.4)
Native	18	23	—	—	—	38.8 (1.1)

Table 2. Summary of the AJT and DCT tasks in the three studies

Task	Target SFP	Structure	Learner group	Language proficiency
AJT	question SFP <i>ba</i> ;	SFP <i>ba</i> questions;	L2; heritage	low-intermediate, high-intermediate, advanced
	SFP <i>le</i>	NP-SFP <i>le</i> ,		
		VP-SFP <i>le</i>		
DCT	question SFP <i>ba</i>	SFP <i>ba</i> questions	L2; heritage	low-intermediate, high-intermediate, advanced

4. THIS STUDY

4.1 Participants

Table 1 shows the details of the participants. Among them, the heritage speakers had been exposed to Mandarin Chinese ($n = 25$) or Cantonese and Min Chinese ($n = 10$) through everyday interactions with family members from birth, with at least one parent being a native speaker of Mandarin Chinese. They all reported English as their dominant (native) language (6 as immigrants, before the age of 5). L2 learners were all English native speakers, and their onset age of learning Chinese was after puberty (around 18 years old). Native speakers were all born and raised in northern China, were postgraduates or undergraduates studying in Beijing, and had not been to an English-speaking country at the time of data collection.

Results of a one-way ANOVA and *post hoc* Tukey tests indicated that all groups were significantly different from each other ($p < .001$) in terms of their proficiency as demonstrated in the Cloze test (refer to Yan et al. 2023 for details), except for the heritage and L2 learners at advanced levels ($p = .157$) and those at high-intermediate levels ($p = .967$).

4.2 Research methods

Data from the two main tasks in the three studies were gathered in the current study (see Table 2), namely an acceptability judgment task (AJT) and a discourse completion task (DCT), which will be introduced and compared below.

In the AJT the participants were asked to judge the acceptability of four types of sentences (see below) on a four-point Likert scale (with possible responses ‘completely unacceptable,’ ‘probably unacceptable,’ ‘probably acceptable,’ and ‘completely acceptable’), with an additional “I don’t know” option. This task was adopted to test both the question SFP *ba* and the SFP *le*.

Sentence types tested (each type has four tokens)

SFP *le*:

Type 1: higher NP + SFP *le* (e.g., (1a));

Type 2: */?lower NP + SFP *le* (e.g., (1b));

Type 3: upward VP + SFP *le* (e.g., (2a));

Type 4: */?downward VP + SFP *le* (e.g., (2b)).

Question SFP *ba*:

Type 5: declarative (e.g., (3a));

Type 6: declarative + question SFP *ba* (e.g., (3b));

Type 7: *wh*-phrase (e.g., (3c));

Type 8: */?*wh*-phrase + question SFP *ba* (e.g., (3d)).

In the DCT, which was mainly used to test the question SFP *ba*, participants were directed to complete a short dialogue (see e.g., (5)) by choosing an appropriate item from a short list of items including the target SFP *ba* and other distractors (such as *de* (a morpheme), *yòu* ('again'), SFP *ma*, SFP *ba*, SFP *ne*, etc.), after completely understanding the dialogue.

(5) the question SFP *ba* is expected (4 tokens)

Speaker A: Nǐ huì shuō Hànyǔ_____?

2SG can speak Chinese

'You can speak Chinese, right?'

Speaker B: Wǒ huì, nǐ zěnmē zhīdào?

1SG can 2SG how know

'I can. How do you know?'

5. RESULTS

5.1 Acquisition of the SFP *ba*

5.1.1 Mean ratings and main statistical findings from Yan (2020) and Yan and Yuan (2022)

AJT

The mean ratings (Table 3) and main statistical findings from Yan (2020) and Yan and Yuan (2022) are first summarized to sketch a general picture of the findings regarding the acquisition of the question SFP *ba* by L2 and heritage learners.

Table 3. Mean ratings in the AJT for sentences with/without the question SFP *ba*

Group	N	declarative (SD), (e.g., (3a))	declarative + SFP <i>ba</i> (SD), (e.g., (3b))	* <i>wh</i> -phrase + SFP <i>ba</i> (SD), (e.g., (3c))	+ <i>wh</i> -phrase (SD), (e.g., (3d))
L2 low-intermediate	18	3.91 (0.2)	2.50 (0.9)	2.06 (0.8)	3.69 (0.4)
L2 high-intermediate	20	3.98 (0.8)	2.40 (1.0)	1.61 (0.8)	3.91 (0.2)
L2 advanced	18	3.97 (0.9)	2.81 (1.1)	1.31 (0.5)	3.86 (0.3)
HL high-intermediate	16	3.91 (0.2)	2.69 (1.1)	1.72 (0.8)	3.92 (0.2)
HL advanced	19	4.00 (0.0)	3.55 (0.7)	1.22 (0.4)	3.93 (0.2)
Native	18	4.00 (0.0)	3.89 (0.2)	1.32 (0.5)	3.96 (0.1)

Note. “*” is standing for ungrammatical sentences.

The results show that all the heritage learners were indistinguishable from the native speakers in rating the declarative + the question SFP *ba* (Yan 2020). However, all the L2 groups were found to be statistically significantly different from the Chinese native group (Yan and Yuan 2022). Meanwhile, the high-intermediate heritage learners (Yan 2020) and the low-intermediate L2 learners (Yan and Yuan 2022) were found to be statistically significantly different from the native speakers in rating the **wh*-phrase + the question *ba*. All the results indicate incomplete acquisition of the syntactic features of the question SFP *ba* in L2 and heritage Chinese grammars to some extent.

DCT

Table 4. Mean scores of all groups in the DCT concerning the question SFP *ba*

Groups	N (after screening)	Use of <i>ba</i> in confirmation seeking (Max.=4, [SD])
L2 low-intermediate	6	1.50 (1.6)
L2 high-intermediate	9	1.11 (1.3)
L2 advanced	9	3.33 (0.9)
HL high-intermediate	16	0.94 (1.3)
HL advanced	19	1.95 (1.3)
Native	18	3.78 (0.4)

Note: The screening for numbers of participants is based on their acceptance of the question SFP *ba* (see section 5.1.3 for details).

Table 4 and the findings from Yan and Yuan (2022) show that the advanced L2 learners were statistically indistinguishable from the native speakers in using the question SFP *ba* in its favored contexts. However, all the intermediate L2 learners (Yan and Yuan 2022) and both the intermediate and advanced heritage learners (Yan 2020) were found to be statistically significantly different from the native speakers in these situations.

5.1.2 AJT – statistical learner group comparisons

Data from Yan (2020) and Yan and Yuan (2022) were combined in terms of the rating scores of L2 and HL learners in the AJT. This combined analysis allowed us to directly examine the differences between the L2 and HL learner groups (and among various levels of learners), and between learners and native speakers, thus providing us with detailed evidence to investigate the CLI from the learner's perspective (i.e., in terms of progress with proficiency, learner type). Furthermore, it also offered the opportunity to examine whether the structural differences between conditions were acquired by both learners and native speakers, in order to provide additional evidence in discussing CLI demonstrated on structures. Cumulative link mixed models (CLMMs, "ordinal" package; Christensen 2015) were performed in R (version 4.1.1; R Core Team 2014) to produce statistical results. Rating scores were the dependent variable, with Group, Condition, and interaction between Group and Condition as fixed effects. The intercepts for subjects and items, as well as by-subject and by-item random slopes, were treated as random factors. Factors were removed if they did not significantly improve the model according to maximum likelihood ratio tests. The results of the modeling are reported in Table 5.

The main effects of the CLMMs show that Mandarin Chinese native speakers could statistically distinguish between declarative sentences with and without the question SFP *ba* (i.e., the declarative vs. declarative + the question SFP *ba* conditions), indicating a sensitivity to the addition of the question SFP *ba* to the declaratives.

Further pairwise comparisons (using the *emmeans* package) indicate that all learner groups, similar to native speakers, were sensitive to the declaratives with or without the question SFP *ba* (all $p < .001$), as well as to sentences embedding a *wh*-phrase with or without the question SFP *ba* (i.e., *wh*-phrase vs. **wh*-phrase + question SFP *ba* conditions, all $p < .001$), demonstrating grammatical sensitivity to the addition of the SFP *ba* to both declaratives and sentences with *wh*-phrases. However, only the HL advanced learner group showed native-like ratings for both the declarative + the question SFP *ba* and the **wh*-phrase + the question SFP *ba* conditions ($p < .05$ for HL high-intermediate and L2 low-intermediate

groups, $p < .001$ for L2 advanced and high-intermediate groups as compared to the native group).

The analysis also found that all learner groups, irrespective of whether they were L2 or heritage learners, were not statistically significantly different from one another in judging both declaratives + the question SFP *ba* and **wh*-phrase + the question SFP *ba* conditions. These findings suggest that learners were sensitive to the grammatical [Q] and [-wh] features of the question SFP *ba*, with no learner difference identified from the perspective of either learner type or level of Chinese proficiency. Only the advanced heritage learners seemed to closely resemble the native speakers.

Table 5. Model outcome from CLMMs (reference level for group: native; reference level for condition: declarative)

Fixed effects	Estimate	Std. Error	z-value	p-value
Condition				
declarative vs. <i>wh</i> -phrase	-2.33	0.29	-7.95	0.00***
declarative vs. declarative + question SFP <i>ba</i>	-3.41	0.29	-11.77	0.00***
declarative vs. * <i>wh</i> -phrase + question SFP <i>ba</i>	-6.75	0.32	-21.17	0.00***
Group				
native vs. HL advanced	-0.50	0.34	-1.46	0.14
native vs. HL high-intermediate	-0.76	0.36	-2.11	0.03*
native vs. L2 advanced	-1.07	0.35	-3.09	0.00***
native vs. L2 high-intermediate	-1.05	0.34	-3.12	0.00***
native vs. L2 low-intermediate	-0.86	0.35	-2.49	0.01*

Notes. “*” = $p < .05$, “***” = $p < .001$. Model formula: ratings ~ condition + group + (1 | subject).

5.1.3 DCT – statistical learner group comparisons

An analysis was conducted of combined data from Yan (2020) and Yan and Yuan (2022) in terms of usage of the question SFP *ba* in its favored context. Since the data were screened according to learners' acceptance of the question SFP *ba* (i.e., allowing the SFP *ba* to occur in questions, high acceptance of declarative + the question SFP *ba* in at least three out of four tokens), the data is ready to allow us to directly measure the extent to which positive CLI affected different learners.

Generalized linear mixed models (GLMMs) were applied to the statistical analysis of the DCT data. Similar to the procedure for performing CLMM analysis discussed in section 5.1.2, scores were the dependent variable, with Group, Age, Length, and interactions between Group, Age, Length as fixed effects. The intercepts for subjects and items, as well as by-subject and by-item random slopes, were treated as random factors. Factors were removed if they did not significantly improve the model according to maximum likelihood ratio tests. The results are reported in Table 6.

Table 6. Model outcome from GLMMs (reference level for group: native)

Fixed effects	Estimate	Std. Error	z-value	p-value
(Intercept)	3.50	0.85	4.09	0.00***
Group	-3.60	0.65	-5.52	0.00 ***
native vs. HL advanced	-5.06	0.71	-7.14	0.00***
native vs. HL high intermediate	-1.40	0.73	-1.91	0.06
native vs. L2 advanced	-4.78	0.75	-6.37	0.00***
native vs. L2 high-intermediate	-6.45	1.01	-6.40	0.00***
native vs. L2 low-intermediate	3.50	0.85	4.09	0.00***

Note. “***” = $p < .001$. Model formula: ratings ~ group + (1 | item).

The results from the GLMMs analysis show that all the learner groups except the L2 advanced learners used the question SFP *ba* significantly less in its favored confirmation-seeking contexts. Further pairwise comparisons demonstrate that the heritage learners' use of the question SFP *ba* was statistically lower than that of the L2 learners at all proficiency levels (all $p < .05$), except the HL high-intermediate learners as compared to both high- ($p = .59$) and low-intermediate L2 learners ($p = .10$). Among the learner groups, both the heritage learners (high-intermediate vs. advanced levels, $p < .001$) and L2 learner groups (both high-intermediate and low-intermediate vs. advanced levels, all $p < .001$) demonstrated improvement in their proficiency regarding the use of the question SFP *ba*, but not the L2 learners at the intermediate levels (high- vs. low-intermediate, $p = .06$).

5.2 Acquisition of the SFP *le*

5.2.1 Mean ratings and main statistical findings in Yan et al. (2023)

The mean ratings (Figures 2 and 3, slightly modified) and main statistical findings from Yan et al. (2023) are given first to facilitate the understanding of results from the complementary analysis (introduced in section 5.2.2).

NP condition

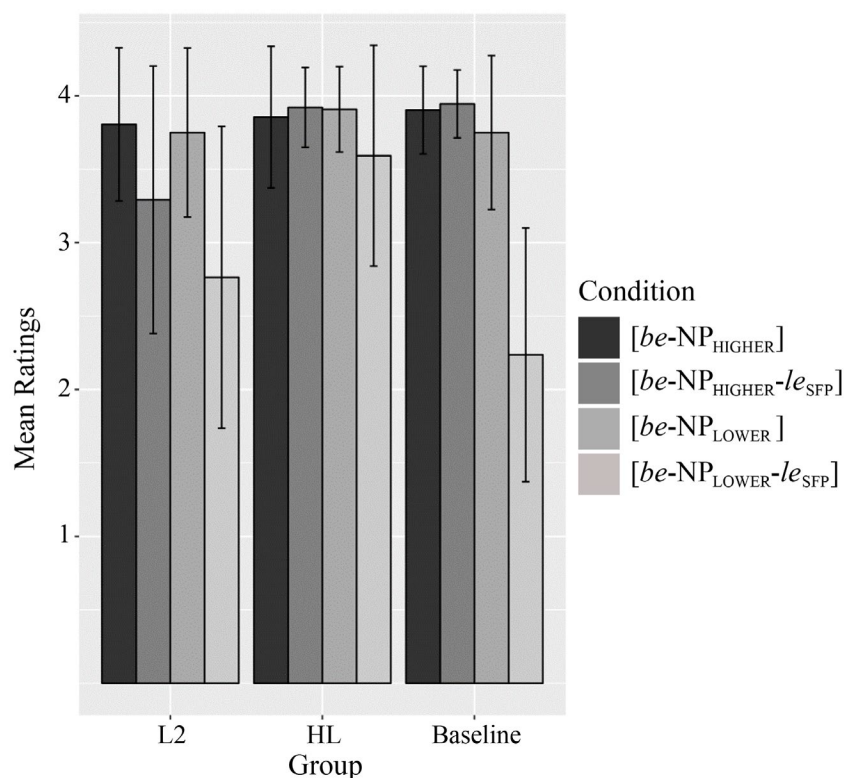


Figure 2. Mean ratings in the AJT in the SFP *le*-NP condition (Yan et al. 2023, slightly modified)

(conditions: [be-NP_{HIGHER}-le_{SFP}] and [be-NP_{LOWER}-le_{SFP}] are simplified as “higher NP + SFP *le*” and “*/?lower NP + SFP *le*” respectively in this study, with [be-NP_{HIGHER}] and [be-NP_{LOWER}] as control sentences for “higher NP + SFP *le*” and “*/?lower NP + SFP *le*” respectively; “Baseline group” = “native group”)

Yan et al. (2023) found that both the L2 and HL groups (at advanced levels) were able to statistically rate the grammatical condition with a higher NP + SFP *le* higher than that of ungrammatical conditions with a */?lower NP + SFP *le* ($p < .05$, results from LMMs).

VP condition

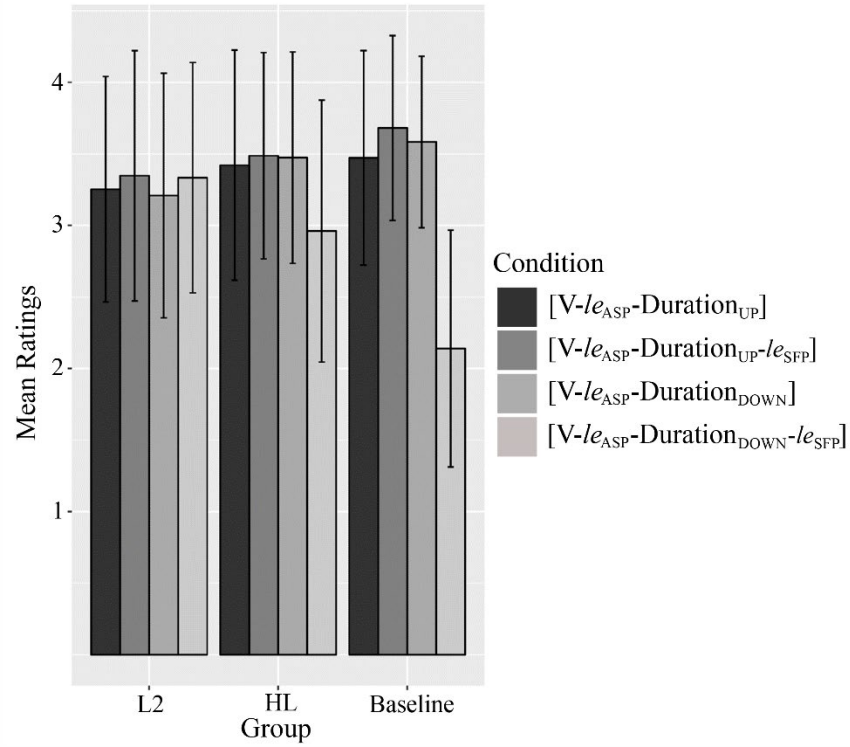


Figure 3. Mean ratings in the AJT in the SFP *le*-VP condition (Yan et al. 2023, slightly modified)

(conditions: [V-*le*_{ASP}-Duration_{UP}-*le*_{SFP}] and [V-*le*_{ASP}-Duration_{DOWN}-*le*_{SFP}] are simplified as “upward VP + SFP *le*” and “*/?downward VP + SFP *le*” respectively in this study, with [V-*le*_{ASP}-Duration_{UP}] and [V-*le*_{ASP}-Duration_{DOWN}] as control sentences for “upward VP + SFP *le*” and “*/?downward VP + SFP *le*” respectively; “Baseline group” = “native group”)

Yan et al. (2023) found that only the HL advanced learners rated the grammatical condition with upward VP + SFP *le* higher than the */?downward NP + SFP *le* condition ($p < .01$), on a par with the native learners.

5.2.2 AJT: complementary statistical analysis of learner and native groups

CLMMs were also applied (using the same procedure as described in section 5.1.1) to re-analyze the data concerning the SFP *le*, mainly to further examine the statistical differences between learner and native groups as well as those between learner groups. This was intended to complement the findings of Yan et al. (2023), and to provide detailed evidence for assessing the CLI from learner perspectives (i.e., according to proficiency levels and learner type). The results are summarized in Tables 7 and 8.

NP condition

Mandarin native speakers rated the */?lower NP + SFP *le* statistically lower than that of the higher NP condition ($p < .001$). Further pairwise comparisons suggest that both learner groups rated the ungrammatical condition with the SFP *le* + lower NP statistically higher than the native speakers ($p < .01$ for HL learners, $p < .05$ for L2 learners), while only the L2 learners were found to be statistically significantly different from the native speakers in accepting the grammatical condition with the SFP *le* + higher NP ($p < .01$). An examination of the performance of learner groups demonstrated that they were statistically significantly different from one another in both the grammatical condition (higher NP + SFP *le*) and the ungrammatical condition (*/?lower NP + SFP *le*, all $p < .001$), suggesting that in the grammatical condition the HL advanced learners outperformed the L2 learners with higher acceptance ratings, whereas the L2 learners outperformed the HL learners in judging the ungrammatical condition as unacceptable (refer to Figure 2).

Table 7. Model outcome for the SFP *le*-NP (reference level for group: native; for condition: higher NP; “*” = $p < .05$, “***” = $p < .001$. Model formula: ratings ~ condition + group + condition: group + (1 | subject) + (1 | item).)

Fixed effects	Estimate	Std. Error	z-value	p-value
Condition				
higher NP vs. lower NP	-1.01	0.55	-1.83	0.07
higher NP vs. higher NP + SFP <i>le</i>	0.58	0.70	0.84	0.40
higher NP vs. */?lower NP + SFP <i>le</i>	-5.02	0.55	-9.18	0.00***
Group				
native vs. HL advanced	0.04	0.70	0.05	0.96
native vs. L2 advanced	-0.43	0.67	-0.64	0.52
Condition : Group				
lower NP : HL advanced	1.28	0.77	1.66	0.10
higher NP + SFP <i>le</i> : HL advanced	-0.13	0.90	-0.14	0.89
*/?lower NP + SFP <i>le</i> : HL advanced	3.49	0.70	5.01	0.00***
lower NP : L2 advanced	0.69	0.70	0.99	0.32
higher NP + SFP <i>le</i> : L2 advanced	-2.62	0.80	-3.27	0.00***
/?lower NP + SFP <i>le</i> : L2 advanced	1.65	0.64	2.56	0.01

Table 8. Model outcome for the SFP *le*-VP (reference level for group: native; for condition: upward VP; “**” = $p < .05$, “***” = $p < .001$. Model formula: ratings ~ condition + group + condition: group + (1 | condition | subject) + (1 | item).)

Fixed effects	Estimate	Std. Error	z-value	p-value
Condition				
upward VP vs. downward VP	0.24	0.71	0.34	0.73
upward VP vs. upward VP + SFP <i>le</i>	1.19	0.81	1.47	0.14
upward VP vs. */?downward VP + SFP <i>le</i>	-4.05	0.85	-4.75	0.00***
Group				
native vs. HL advanced	-0.02	0.62	-0.03	0.98
native vs. L2 advanced	-0.83	0.62	-1.34	0.18
Condition : Group				
downward VP : HL advanced	-0.17	0.55	-0.31	0.76
upward VP + SFP <i>le</i> : HL advanced	-0.98	0.72	-1.36	0.17
/?downward VP + SFP <i>le</i> : HL advanced	2.29	0.86	2.67	0.01
downward VP : L2 advanced	-0.40	0.54	-0.74	0.46
upward VP + SFP <i>le</i> : L2 advanced	-0.74	0.73	-1.02	0.31
*/?downward VP + SFP <i>le</i> : L2 advanced	4.41	0.88	4.99	0.00***

VP condition

In the VP condition the main effects of the CLMMs also suggest that Mandarin native speakers rated the condition */?downward VP + the SFP *le* statistically different from the upward VP condition. Results from further pairwise comparisons show that the L2 learners rated both the grammatical condition with the upward VP + SFP *le* ($p < .05$) and the ungrammatical condition with */?downward VP + the SFP *le* ($p < .001$) statistically different from the native group. Meanwhile, for the HL group only the ungrammatical condition ratings were found to be statistically significantly different ($p < .001$) from the native group. This suggests that both the L2 and HL learners did not perform in a native-like way in rating ungrammatical conditions; while in contrast the HL learners exhibited native-like grammar in accepting the grammatical condition. A comparison between learner groups indicates that the HL group was found to be advantaged compared to the L2 group only in the lower rating of the ungrammatical condition (*/?downward VP + SFP *le*, $p < .05$). This further confirms that the HL learners outperformed the L2 learners in the ratings of ungrammatical conditions.

6. SUMMARY: ANSWERING THE RQS

As summarized in Table 9, in terms of the syntactic features of the question SFP *ba*, our findings suggest only heritage learners (at advanced levels) show a degree of native-like performance (no statistically significant difference WAS found between them and the native group). However, no significant difference was found between L2 and heritage learners in the sense that both were considered sensitive to cases with and without the question SFP *ba*.

Examination of the discourse feature of the question SFP *ba* seems to suggest L2 learners outperformed heritage learners, especially those at advanced levels who performed in a native-like manner. L2 learners also demonstrated improvement in their Chinese proficiency.

Table 9. Summary of findings

Target SFPs	Structures	Native-like performance	Sensitivity	L2 vs. HL
Question SFP <i>ba</i>	declarative + SFP <i>ba</i>	(Yan, 2020; Yan & Yuan 2022)	/	
	* <i>wh</i> -phrase + SFP <i>ba</i>	(Yan, 2020; Yan & Yuan 2022)	/	
	declarative vs. declarative + SFP <i>ba</i>	/	L2✓ HL✓	/
	<i>wh</i> -phrase vs. * <i>wh</i> -phrase + SFP <i>ba</i>	/	L2✓ HL✓	/
	usage of SFP <i>ba</i> in confirmation seeking contexts	L2✓	/	L2✓
(Yan 2020; Yan and Yuan 2022; this study)				
SFP <i>le</i>	higher NP + SFP <i>le</i>	HL✓	/	HL✓
	*/?lower NP + SFP <i>le</i>			L2✓
	higher NP + SFP <i>le</i> vs. */?lower NP + SFP <i>le</i>	/	L2✓ HL✓	/
	upward VP + SFP <i>le</i>	HL✓	/	
	*/?downward VP + SFP <i>le</i>			HL✓
(Yan et al. 2023)				
	upward VP + SFP <i>le</i> vs. */?downward VP + SFP <i>le</i>	/	HL✓	/
(Yan et al. 2023)				

Notes: “✓” indicates a learner effect was found; “/” means not applicable. “Sensitivity” suggests whether a difference is detectable. “Native-like performance” indicates the similarity of learners’ performance as compared to native speakers.

Findings from the acquisition of the syntactic constraints of the SFP *le* show a complex picture. For the NP condition, at advanced levels both groups of learners showed sensitivity to the syntactic constraints of the SFP *le*, whereas in the VP condition only the heritage learners exhibited such sensitivity. Heritage learners performed in a native-like way and outperformed their L2 peers in terms of grammatical conditions (both NP and VP conditions), as well as in the ungrammatical VP condition. In contrast, however, they performed worse than L2 learners in the ungrammatical NP condition.

In response to our RQs, in terms of acquiring the Mandarin SFPs *ba* and *le*, the difficulties of L2 learners lie in the syntactic features (i.e., [Q] and [-wh] of the question SFP *ba*) and more complex structures (i.e., VPs vs. NPs co-occurring with the SFP *le*). In these scenarios the heritage learners seem to outperform L2 learners; for the heritage learners, difficulties mainly lie in discourse features (i.e., [confirmation seeking] of the question SFP *ba*), suggesting that the L2 learners were disadvantaged. Proficiency seems to have a greater influence in terms of acquiring discourse features among L2 learners, but it is less important across different types of learners.

7. DISCUSSION

This study has conducted a follow-up complementary re-analysis of data from three related studies in the hope of better understanding CLI, particularly positive CLI. Findings from the complementary analysis of the data for the SFP *le* confirm the findings of Yan et al. (2023), and combined comparisons between learners' data from Yan (2020) and Yan and Yuan (2022) offer more evidence on how CLI and positive CLI are demonstrated among heritage and L2 learners at various proficiency levels with regard to a particular grammatical structure. In light of recent discussions on positive CLI (e.g., Yan et al. 2023; Yip et al. 2018) and in response to our RQs in this study, we will discuss the current findings regarding CLI and positive CLI in terms of learner type, language proficiency levels, and structures.

7.1 Learner type: Early and richer target language exposure might not always be advantageous

Heritage learners' superior performance compared to their L2 peers in syntactic aspects of the question SFP *ba* and the SFP *le*, along with their relative disadvantages in discourse aspects of the question SFP *ba*, suggest that syntax represents the strongest area, while discourse-related domains (like the semantics-discourse interface) are more vulnerable in heritage language grammars (Montrul 2012).

We have also identified the difficult asymmetry between L2 and heritage groups within the syntactic sphere. With regard to the complementary analysis in the current study, we have found that within the syntactic constraints of the SFP *le*, heritage learners outperformed their L2 peers concerning the ungrammatical VP condition, a complex structure since it is deeply embedded and thus linearly distant from the target element, as discussed in Yan et al. (2023). In contrast, L2 learners outperformed heritage learners in terms of the ungrammatical NP condition. This seems to further suggest that heritage learners are very sensitive to syntactic constraints, even for complex structures. It might be argued that this advantage is due to heritage learners' early and richer exposure to the heritage language. However, we suggest that this might account for only some of the acquisition.

Heritage learners' disadvantages in the discourse field indicate their unsuccessful (or incorrect) mapping of the question SFP *ba* as compared to L2 learners. As argued in Yan and Yuan (2022), the acquisition of the question SFP *ba* features crucially depends on successfully mapping it onto English tag questions. The fact that L2 learners are indistinguishable from native speakers in this regard strongly suggests that their establishment of this mapping is robust and impervious to complex distracting evidence from target language input (see Figure 1 for the complexity, and refer to Yan and Yuan 2022 for details). On the other hand, however, heritage learners might be significantly misled by the input and might hesitate in establishing this mapping relationship, due to the prediction that once the mapping is established the acquisition of the discourse feature can be readily achieved. This constitutes a positive CLI, as defined in the current study. In this sense, it might take longer for

heritage learners to review and digest complicated evidence from their input and establish robust mapping relations, thereby allowing positive CLI to come into play. Future research could look at near-native heritage learners to further investigate this issue.

7.2 Language proficiency levels: Where transition from a negative CLI to a positive CLI might be detected

The inclusion of the language proficiency variable allowed us to enhance our comprehension of CLI, especially its positive aspects and how they manifest in learners. The unique influence of language proficiency is observed among L2 learners during the acquisition of the discourse feature of the question SFP *ba*. This provided us with a window through which to examine the transition of CLI from negative to positive.

Recall that in section 2 we attempted to refine the classification of CLI. Instead of simply categorizing CLI as either negative or positive in the context of acquiring an L2 structure, we proposed that both types of CLI can be involved, albeit at different stages of learning. This distinction is particularly crucial for functional categories that do not exist in L1 and where direct lexical correspondence is absent, but where a closely related structure can be identified (as in the case of the question SFP “*ba*” and English tag questions). In such scenarios there may be an initial negative CLI, seemingly obstructing the accurate mapping of the L2 target onto the related L1 structure and thereby impeding the emergence of positive CLI that could facilitate acquisition. In this sense, only L2 learners at advanced levels could successfully acquire the discourse feature of the question SFP *ba*, suggesting successful mapping between the L1 and L2 counterparts and hence indicating that positive CLI has finally occurred. This underscores the crucial role of mapping, as previously discussed in Yan and Yuan (2022). Future research could include tasks designed to identify mapping challenges, particularly in terms of pinpointing precise mapping correspondences at the point when a negative CLI is operative.

7.3 Structures: Different features within the same structure could be insurmountable

The extent of positive CLI across structures governed by the same syntactic rule was demonstrated in Yan et al. (2023). This finding was corroborated by the present study through an additional analysis of the SFP *le*, where L2 learners encountered difficulties with VP conditions. Moreover, comparing the acquisition of the question SFP *ba* between heritage and L2 learners offers further evidence of the positive impact of CLI on structural issues.

This relates to a scenario where a single structure displays various features to fulfill its functions across different levels. For instance, the use of the question SFP *ba* primarily aims to confirm an assumption, functioning at the discourse level (a discourse feature) while also taking the form of a yes-no question which operates at the syntactic level (syntactic features). Learners in this study demonstrated a contrasting grasp of grammar regarding the question SFP *ba*: the L2 learners at advanced proficiency levels outperformed heritage learners in terms of the discourse feature, but were inferior at the syntactic level. At more advanced levels the L2 learners were proposed to have established a mapping between L1 and L2 correspondences, and a positive CLI was predicted to take effect. However, this finding seems to suggest that positive CLI is feature-sensitive, even within the same target structure. Thus, the findings enrich our understanding that positive CLI, when it is present, is not consistently active and facilitative, either across different structures or within a single structure.

8. CONCLUSION

With complementary and combined analyses of data regarding the question SFP *ba* from Yan (2020), Yan and Yuan (2022) and the SFP *le* from Yan et al. (2023), this study confirms the proposal offered by Yan et al. (2023) that the learner and the structure do indeed mediate positive CLI. In particular, we have demonstrated that early and richer target language exposure might not always be advantageous, and that positive CLI might

Shanshan Yan

be feature-sensitive, even within a single structure. Language proficiency as a newly added variable in examining CLI proves to be significant in regulating a transition from negative to positive CLI. This study calls for a more holistic view of various manifestations of CLI, and a delicate and refined analysis of different CLI effects in SLA.

REFERENCES

- Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Christensen, Rune Haubo Bojesen. 2015. Ordinal: Regression Models for Ordinal Data (R package version 2015), retrieved March 16, 2024, from <http://www.cran.r-project.org/package=ordinal/>
- Jackson, Carrie N., and Paola E. Dussias. 2009. Cross-linguistic differences and their impact on L2 sentence processing. *Bilingualism: Language and Cognition* 12(1):65–82.
- Lago, Sol., Michela Mosca, and Ana Stutter Garcia. 2021. The Role of Crosslinguistic Influence in Multilingual Processing: Lexicon Versus Syntax. *Language Learning* 71(S1):163–192.
- Li, Charles, and Sandra A. Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.
- Montrul, Silvina. 2012. Is the heritage language like a second language? In *Eurosla Yearbook 12(1)*, eds. Leah Roberts, Christina Lindqvist, Camilla Bardel, and Niclas Abrahamsson, pp.1–29. Amsterdam: John Benjamins Publishing Company.
- Odlin, Terence. 1989. *Language Transfer: Cross-linguistic Influence in Language Learning*. Cambridge: Cambridge University Press.
- Odlin, Terence. 2001. *Cross-linguistic Influence in Language Learning*. Shanghai: Shanghai Foreign Language Education Press.
- Odlin, Terence. 2022. *Explorations of Language Transfer*. Bristol: Multilingual Matters.
- Pan, Victor Junnan. 2019. *Architecture of the Periphery in Chinese: Cartography and Minimalism*. Oxon: Routledge.
- Paul, Waltraud. 2015. *New Perspectives on Chinese Syntax*. Berlin: De Gruyter Mouton.
- Paul, Waltraud, and Yan, Shanshan. 2022. Sentence-final particles in Mandarin Chinese: Syntax, semantics and acquisition. In *Discourse Particles: Syntactic, Semantic, Pragmatic and Historical Aspects, Linguistik Aktuell 276 (Linguistics Today)*, eds. X. Artiagoitia, A. Elordieta, and S. Monforte, pp.179–206. Amsterdam: John Benjamins Publishing Company.
- R Core Team. 2014. R: A language and environment for statistical computing (Version 4.1.1). Vienna, Austria: R Foundation for Statistical Computing, retrieved Aug 10, 2021, from <http://www.R-project.org/>
- Szeto, Pui Yiu, Stephen Matthews, and Virginia Yip. 2017. Multiple correspondence and typological convergence in contact-induced grammaticalization. *Journal of Language Contact* 10:485–518.
- Soh, Hooi Ling. 2009. Speaker presupposition and Mandarin Chinese sentence-final-*le*: A unified analysis of the ‘change of state’ and the ‘contrary to expectation’ reading. *Natural Language & Linguistic Theory* 27(3):623–657.
- Tang, Sze-Wing. 2010. *Formal Chinese Syntax*. Shanghai: Shanghai Educational Publishing House.

Shanshan Yan

- Xiang, Kaixi. 1998. On the prominence of entities and the marker *da*. *Hanyu xuexi* 1:12-16.
- Xing, Fuyi. 2001. *A Study of Chinese Complex Sentences*. Beijing: Shangwu yinshuguan.
- Yan, Shanshan. 2020. Syntactic and discourse features in Chinese heritage grammars: A case of acquiring features in the Chinese sentence-final particle *ba*. *Languages* 5(2):26.
- Yan, Shanshan, and Boping Yuan. 2022. Asymmetric form-meaning mappings in L2 acquisition of Chinese sentence-final particles *ba*. *Journal of Chinese Linguistics* 50(1): 1–38.
- Yan, Shanshan, Ziyin Mai, and Yang Zhao. 2023. Positive cross-linguistic influence in the representation and processing of sentence-final particle *le* by L2 and heritage learners of Chinese. *Frontiers in Psychology* 14:1145493.
- Yip, Virginia, Ziyin Mai, and Stephen Matthews. 2018. CHILDES for bilingualism. In *Bilingual Cognition and Language: The State of the Science Across its Subfields*, ed. D. Miller, F. Bayram, J. Rothman, and L. Serratrice, pp. 183–201. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Yuan, Boping. 2004. Negation in French–Chinese, German–Chinese and English–Chinese interlanguages. *Transactions of the Philological Society* 102(2):169–197.
- Yuan, Boping, and Yang, Zhao. 2005. Resumptive pronouns in English–Chinese and Arabic–Chinese interlanguages. *International Review of Applied Linguistics in Language Teaching* 43(3): 219–237.

Shanshan Yan

School of Chinese as a Second Language

Peking University

No.5 Yiheyuan Road, Haidian District, Beijing, 100871

shanshanyan@pku.edu.cn

APPENDIX: Samples of experimental sentences

AJT:

The SFP *le*:

Experimental sentences:

Type 1: higher NP + SFP *le*

Xiǎowáng shì dà hái zi le.

Xiaowang be big child LE

‘Xiaowang is a big child already.’

Type 2: */?lower NP + SFP *le*

*/?Xiǎowáng shì xiǎo hái zi le.

Xiaowang be small child LE

*/?‘Xiaowang is a small child already.’

Type 3: upward VP + SFP *le*

Wǒ zài zhèr zhù le zhìshǎo wǔ nián le.

1SG at here live PERF at-least five CL LE

‘I have lived here for at least 5 years already (which was not the case before).’

Type 4: */?downward VP + SFP *le*

*/?Wǒ zài zhèr zhù le búdào wǔ nián le.

1SG at here live PERF less-than five CL LE

*/?‘I have lived here for less than 5 years already.’

Control sentences:

Type 1: higher NP

Xiǎowáng shì dà hái zi.

Xiaowang be big child

‘Xiaowang is a big child.’

Type 2: lower NP

Xiǎowáng shì xiǎo hái zi.

Xiaowang be small child

‘Xiaowang is a small child.’

Type 3: upward VP

Shanshan Yan

Wǒ zài zhèr zhù le zhìshǎo wǔ nián.
1SG at here live PERF at-least five CL
'I have lived here for at least 5 years.'

Type 4: downward VP

Wǒ zài zhèr zhù le búdào wǔ nián.
1SG at here live PERF less-than five CL
'I have lived here for less than 5 years.'

The question SFP *ba*:

Type 5: declarative

Tā huì kāi chē.
3SG can drive car
'He(She) can drive a car.'

Type 6: declarative + question SFP *ba*

Tā huì kāi chē ba?
3SG can drive car BA
'He(She) can drive a car, can't he(she)/right?'

Type 7: *wh*-phrase

Shéi huì kāi chē?
who can drive car
'who can drive?'

Type 8: **wh*-phrase + question SFP *ba*

*Shéi huì kāi chē ba?
who can drive car BA
Intended: 'who can drive?'

DCT:

The question SFP *ba* (instruction as selecting the most appropriate particle with complete reading of the dialogue):

the question SFP *ba* is expected (4 tokens)

Speaker A: Nǐ huì shuō Hànyǔ _____?
2SG can speak Chinese
'You can speak Chinese, right?'

Positive Cross-linguistic Influence in L2 and Heritage Chinese Acquisition

Speaker B: Wǒ huì, nǐ zěnmē zhīdào?
1SG can 2SG how know
'I can. How do you know?'

Shanshan Yan

句末助詞習得中正向跨語言影響的追蹤研究：
學習者、結構和語言水平再探

閔姍姍
北京大學

本研究對三個漢語普通話句末助詞研究進行了深入分析。通過補充和結合的對比分析，本研究考察了跨語言正向影響在學習者、語言水平以及同一結構中的差異。研究發現，繼承語學習者並不總是比二語學習者更有優勢。在高級水平的二語學習者中，從負向到正向的跨語言影響可以被監測到。跨語言的正向影響在同一句法結構的不同特徵上表現不同。未來研究要從更宏觀的角度觀察跨語言的正向影響。

關鍵字：跨語言正向影響、句末助詞、繼承語學習者、二語學習者、句法結構、語言水平

**FOR WHOM IS L1 SPEAKER MODIFICATION
IN L1-L2 CONVERSATION?:
RECONCEPTUALIZING FOREIGNER TALK FROM
THE PERSPECTIVE OF INTERACTION ANALYSIS***

Tomoharu Yanagimachi
Hokusei Gakuen University

ABSTRACT

Previous research on linguistic modification by L1 speakers in L1-L2 conversation has primarily relied on data collected in ESL/EFL classrooms or experimental settings. However, the majority of L1-L2 interactions occur not in controlled environments but in everyday interactional situations. In this study, the author employed the framework of interaction analysis to investigate how L1 English speakers modify their utterances, or ‘foreigner talk,’ in a ski shop in Japan. The analysis was conducted using video recordings capturing interactions where English-speaking customers asked Japanese staff to repair skis. A recurring phenomenon in the data was that the customers modified their utterances to facilitate mutual understanding and to elicit the staff’s professional assessment and advice regarding the condition of their skis and the necessary repair work. Both employees and customers oriented themselves toward each other’s participatory roles, knowledge, and skills, rather than focusing on the linguistic asymmetry between them. The findings suggest some pedagogical implications, and can raise awareness among teachers, learners, native speakers, and researchers.

Keywords: speech modification, foreigner talk, authenticity, social context, interaction analysis, video data

*I would like to express my sincere appreciation to Professor Stano Kong for his support and guidance throughout the completion of this paper. I would also like to thank the two anonymous reviewers who provided insightful comments on earlier versions of the manuscript. However, all remaining errors are mine.

1. BACKGROUND OF THE STUDY

It is well known that speech from L1 speakers to L2 speakers is often subject to modification in linguistic forms and interaction. According to Ferguson (1975), “speech communities have conventional varieties of ‘simplified’ speech which are regarded by the speakers as appropriate for use when the hearers do not have full understanding of the language” (p. 1). Ferguson (1971, 1975) claims that ‘foreigner talk’ employed by L1 speakers when they interact with L2 speakers is one of these special varieties. Foreigner talk is characterized by linguistic modifications and simplifications in L1 speaker speech, including slower speech rate, careful articulation, pauses, emphatic stress, avoidance of contractions, paraphrasing, use of high-frequency words and expressions, tag questions, topicalization (Ferguson 1971, 1975; Gass and Selinker 2008; Saville-Troike 2012).

In the 1980, in the field of second language learning and teaching two hypotheses which were closely related to Ferguson’s studies of simplified speech forms in foreigner talk were proposed: Krashen’s (1981, 1982) Input Hypothesis and Long’s (1985) Interaction Hypothesis. Krashen referred to the type of input, spoken or written, that is slightly more advanced than their current level as “comprehensible input.” He argued that learners can develop their language system naturally when they receive the messages that contain comprehensible input and process them for meaning. Long, whose hypothesis focuses on the importance of face-to-face interaction and negotiation of meaning, claimed that linguistic and interactional modification in interaction promotes second language learning. These hypotheses proposed by Krashen and Long shared common ground in claiming that the input made comprehensible for the learner is essential for second language acquisition.

When Ferguson (1971, 1975) developed his argument about simplified speech in foreigner talk, however, his academic inquiry was primarily focused on sociolinguistics and anthropological linguistics, exploring (1) the relationship between the forms of the special variety and observable behaviors of foreigners; and (2) the cognitive and historical origin of such varieties. In contrast, within the research communities of second language learning and teaching, studies on L1 speakers’ modifications in linguistic

forms during speech and interaction, including Ferguson's studies on foreigner talk, have been approached from the perspective of how modified input affects L2 speakers' acquisition of the target language. While there is merit in this objective, this research tradition and paradigm have focused on the analysis and discussion of spoken data collected predominantly in educational settings, such as language classrooms or experimental settings. In these data, L1 speakers are typically language teachers or native speakers recruited by researchers, and the L2 speakers are language learners.

The issue with the above research paradigm and its assumptions is presupposition that L1 speakers always act as benefactors who assist L2 speakers' language learning, and L2 speakers always benefit from simplified input. As I will discuss below, however, if we observe real-life conversations, we find that this approach diverges significantly from actual language use. This bias partially stems from the academic field's narrow focus on elucidating mechanisms of second language acquisition and developing teaching methods and materials aimed at optimizing it.

L1 speakers are not always language teachers and L2 speakers are not always learners. In reality, most L1-L2 interactions occur in everyday settings such as shops, restaurants, train stations, hotels, hospitals, and workplace. In these natural conversational contexts, L1 and L2 speakers assume roles like clerk or customer, station attendant or passenger, doctor and patient, or colleagues. For instance, an L2-speaking station attendant, with limited language proficiency, may provide important information to an L1-speaking passenger, or an L2-speaking doctor may examine an L1-speaking patient. When L1 speakers modify their language in these contexts, it is typically to facilitate the ongoing authentic activity they are engaged in, rather than to aid L2 speakers in acquiring the target language. L1 speakers typically do not have any pedagogical intent, and L2 speakers do not expect it from them. The research tradition focused on simplified input and linguistic modification has overlooked describing how such simplification and modification between participants who do not share the same language contributes to organizing activities in authentic daily interactions.

2. PURPOSE OF THE STUDY

This paper aims to elucidate, based on conversational data videotaped in a ski shop at a winter resort in Japan, how L1 English-speaking customers adapt their speech and interaction when communicating with L2 English-speaking Japanese staff, primarily for their own practical benefit, rather than for the staff's benefit. In the data the L1 English-speaking customers communicate with Japanese staff in English to request ski repairs. A recurring phenomenon observed in the data is that these English-speaking customers modify their speech in order to enhance comprehensibility, aiming to achieve mutual understanding and solicit the staff's assessment and advice on ski conditions and necessary repairs. Despite their linguistic advantage, these English-speaking customers lack the ski repair knowledge possessed by the Japanese staff.

Tarone (2000), while acknowledging second language acquisition as a psycholinguistic process, argues that it is only part, not all, of the process and that theories of second language acquisition must take into account the social context. She criticizes researchers for only considering psycholinguistic processes and marginalizing social and sociolinguistic factors. Gardner and Wagner (2004) also argue that, unlike conversational data collected in ESL/EFL contexts, the consequences of authentic interaction affect people's real lives. The authenticity and context of the ski shop interaction create an environment in which L1 English-speaking customers to negotiate with Japanese staff, which in turn maximizes the service they can receive for the money they spend.

Successful communication between foreign customers and Japanese staff does not always depend on the staff's proficiency in English. Rather, the real-life contact situations are diverse, and what is more important is the roles and responsibilities the participants are oriented to fulfill throughout the customer service situation, and how much knowledge and skill they have in the matters at the focus of the interaction between the two parties.

Unlike studies on foreigner talk and simplified input in the field of second language learning and teaching, which have mostly focused on static linguistic features, the present study will illustrate how linguistic

modification can typically be demonstrated and observed in the actual activities in which people engage.

3. METHOD

In analyzing the interactions between L1 English-speaking clients and Japanese-speaking staff in this paper, I will examine them from the perspective of interaction analysis, as exemplified by Jordan and Henderson (1995), for example, in the field of workplace studies and Firth and Wagner (1997) and Gardner and Wagner (2004) in the field of second language studies. Interaction analysis has its roots in several research fields, including ethnomethodology and conversation analysis (EMCA). Through detailed analyses of recordings of people's daily interactions, interaction analysis seeks to clarify how people understand each other and organize their activities. Within this research paradigm, the analysis of interaction is characterized by a multimodal approach. This means researchers examine not only participants' talk, but also embodiments such as gaze, gestures, and posture. Additionally, researchers consider how participants use objects within the social and material environments where human actions and activities are situated.

Whether they are native speakers or not, people do not converse with others in order to use language; rather, they converse using language as one resource to carry out various daily practices. From this perspective, what is important is to carefully describe how the parties involved participate in mutual action and organize their specific activities. In this study, I analyze authentic conversational interactions in customer service based on video and textual data. My focus is on the participants' speech and embodiment, including eye gaze and pointing, and the use of artifacts.

Here, we consider how validity is addressed in interaction analysis, specifically how the credibility of data analysis is ensured. According to Peräkylä (2011), in conversation analytical research, the aim is to investigate interaction not as a representation of other social phenomena but "as a phenomenon in its own right," and "(t)his commitment to naturalistic description of interaction gives a distinctive shape to the issues of validation" (p.367). One way of ensuring the credibility of data analysis

is “validation through next turn” (p.368). In conversation analytical research, it is assumed that a hearer’s interpretation of the preceding utterance is displayed in the subsequent actions. Thus, “the next turn will show whether the interactants themselves treat the utterance in ways that align with the analyst’s interpretation” (p.368).

Regarding the reliability of the issues in conversation analytical research, ten Have (2007) claims that “(i)f, as the corpus of CA studies suggests, the way people organize their talk-in-interaction is ‘ordinary,’ that is, based on a set of formal procedures with immense generalizability, then it does not matter very much which particular specimens one collects to study that order” (p.50). It goes without saying that we must accumulate as much data and analysis from similar contexts as possible to enhance the accuracy of investigation.

3.1 Data Collection Site

The conversational data analyzed in the present study was collected at a rental ski shop in a winter resort located in a rural town in Hokkaido, the northernmost island of Japan, a few years before the Covid-19 pandemic. At the time of data collection, the town had a population of only about 15,000 people, but because of the economic boom created by the influx of foreign skiers and English-speaking workers in the service sector in winter, from 2015 to 2020 the town experienced the highest increase in land prices in Japan. Property purchases, particularly by foreign investors and wealthy individuals, are still ongoing, at the time of writing. Additionally, hotel room rates and food prices in the area have been significantly higher than the national average in Japan.

The ski rental shop where the video data was collected was owned by an Australian and managed by a New Zealander. The shop’s customers were almost all individual overseas travelers, and, with the exception of a few Japanese-speaking employees, the majority of employees were also from overseas, including North America, Europe, and Oceania. Most of the employees were in their 20s and 30s and on seasonal contracts. After the ski season in Japan, some of them moved to the Southern Hemisphere where they continued to work in ski-related jobs year after year. English was the *lingua franca* of the shop and the entire resort area. All signs in

the store were in English, and during the five days the researcher videotaped conversations in the store, he did not see a single Japanese customer.

3.2 Data

After obtaining permission from store management and individual customers and employees, the researcher video-recorded service encounters between Japanese staff members and foreign customers at three different sections of the shop over a five-day period: the reception desk, snow board checkout & return counter, and ski repair department. There were two types of Japanese employees working in the shop.

The first group consisted of staff members in their 20s and early 30s who traveled extensively to foreign ski resorts and could speak natural English. Their English was characterized by its frequent use of casual expressions like ‘you guys’, ‘cool’, as well as its natural pronunciation. Although their English ability was limited to informal conversational contexts, the management highly evaluated the Japanese employees in this group. The management of the ski store required its employees to treat customers in the friendly manner that one would expect in an English-speaking country, rather than in a formal manner common in Japan, and the approach and use of English by this type of Japanese staff was consistent with that policy. This group of Japanese staff was stationed at the general reception counter and the check-out counter for rental items.

The second group, the two Japanese-speaking workers assigned to the ski repair department of the shop, is the subject of the analysis of the present study. These two are slightly older than the first group, married with families, and less seasonally mobile. They had some overseas experience and could communicate in English, but not as fluently as the Japanese speakers in the first group. As we will see in the following sections, they often have difficulties in producing English, and sometimes interjected Japanese fillers, words, or phrases in their speech. They worked in the repair department of the shop where proficiency in conversational English was less important than in the other sections.

3.3 Asymmetries between employees and customers

In this paper, we will look at conversations in which the two Japanese repair shop employees from the second group serve customers in English. The key issue here is the asymmetries between the employees and customers in terms of language, participation roles, and ski repair knowledge. As mentioned above, the linguistic asymmetry between L1 and L2 English speakers is not the only issue that needs to be addressed in this context. The history of language modification research in the field of second language acquisition and teaching has focused on this point alone, but in this particular context of authentic language communication, it is also important to consider the participatory roles of customers and employees and their knowledge and skills related to ski repair. These asymmetries play a significant role in organizing their interaction.

First, we will look at the asymmetry in the participation roles of both parties in the conversation. The primary concern of the Japanese-speaking staff working in the repair department is safety. When equipment is brought in for repair, the problem areas are thoroughly inspected, and careful and reliable repair procedures are suggested to the customer to prevent unforeseen accidents while the customer is skiing after the repair. However, the customer does not necessarily share the same orientation or participatory role. Rather, many customers want repairs to be inexpensive and quick. From their perspective, not only are meticulous repairs more expensive, but the longer it takes to fix the equipment means less time to enjoy skiing during their stay. The two examples of conversations examined in the following sections also demonstrate that customers are orienting toward this point. There is an asymmetry between the employees and customers regarding responsibilities and expectations concerning the services provided and received, and the conversations that develop are organized around the value of the service and trust in the other party.

The second asymmetry is that the staff has a great deal of experience, skill, and knowledge in ski equipment repair, while many customers are not as experienced as the staff in these areas. In other words, there is a knowledge and skill asymmetry between experts and amateurs. When skis are brought in for repair, the staff first assesses the condition of the skis. They then provide the customer with an estimate of the repair method,

cost, and expected time frame. The customer then confirms the repair method and time frame or makes any additional requests. As observed in the data in the present study, these procedures and negotiations take place between the two parties, but only the staff has the expertise to make the final decision. The fact that the Japanese staff has more repair knowledge and skills than their English-speaking customers allows them to provide recommendations and expert advice during negotiations, demonstrating the staff's superiority in this context. This contrasts with the third point described below, which is the asymmetry in English proficiency.

The third asymmetry relates to English proficiency between the Japanese staff in the repair department and their customers. As mentioned above, although the Japanese employees are experts in ski repair work, their English communication skills are significantly inferior to those of other Japanese employees working at the front desk and ski rental counters. Most of the customers they serve are native speakers of English from North America, Oceania, and Europe, or well-educated, affluent travelers from countries or regions where English is a second language, such as Singapore, Hong Kong, and Malaysia, with English proficiency comparable to that of L1 English speakers. The asymmetry in English proficiency between staff and customers is striking. This asymmetry is worsened by the nature of the work in a ski repair shop, which requires a fairly complicated and in-depth conversation to diagnose the condition of the skis, explain and persuade the customer about complicated issues such as repair procedures, number of days required, and cost.

As we have seen above, there are asymmetries between the two Japanese speakers working in the repair department and their customers in three areas: participation roles, knowledge and skills, and English proficiency. In this paper, we will examine the nature of conversations between L1 and L2 speakers and how, if at all, linguistic modifications occur in these interactions. In doing so, however, we will not focus solely on the English proficiency of the Japanese staff and clients. As mentioned in the purpose section of this paper, the situations in which people interact in everyday life are diverse, and it is necessary to look at the whole picture of the roles, responsibilities, and knowledge of the people who participate in these situations. The linguistic and interactional modifications that occur between L1 and L2 speakers need to be examined holistically. To

this end, the following analysis section will focus not only on English proficiency but also on the participatory roles, knowledge, and skills of the parties involved.

4. ANALYSIS

In this section, we will examine two excerpts from conversations videotaped between the two Japanese-speaking employees in the repair department and their customers, and examine in detail their speech and embodiment, including eye gaze, pointing, and artifacts. In doing so, using a transcription convention that accurately reflects actual conversations is critical. The present study employs transcription symbols that are widely used in the field of conversation analysis (ten Have 2007). An explanation of each symbol used in the excerpts is given at the end of this report.

4.1 Linguistic modification by L1 speaker

In Excerpt 1 below, C1, a customer from the U.S., brings a pair of skis in need of repair to the repair department's reception counter and consults with the two Japanese employees in the workshop, S1 and S2. The entire over-the-counter exchange between C1, S1, and S2 lasts exactly 6 minutes, and Excerpt 1 below is from the last 25 seconds of the exchange. The staff has already finished diagnosing the skis, and in the conversation preceding this excerpt, a proposal to repair the problem area has been made by the staff. Customer C1's main concern in the following conversation is whether he should repair his skis and continue to use them, or whether he should not repair them and buy new ones. Therefore, C1 is trying to confirm with staff S1 and S2 whether the repair process proposed by them is really worth doing, and whether it is better to repair the skis as proposed than to buy new ones. However, in the first few lines of this excerpt, the staff is unable to understand C1's request for confirmation, and thereafter, an exchange to achieve mutual understanding between the two continues.

L1 Speaker Modification in L1-L2 Conversation

[Excerpt 1] C1=Customer 1, S1=Staff 1, S2=Staff 2
001 C1: Is it worth doing? ((while looking at S1))
002 S1: Uh?
003 C1: Is it <worth> doing? Should I do it?
004 (1.8)
005 S1: W- wh- what's doing?
006 C1: The skis are still pretty good right?
007 (0.8)
008 S1: Ye:ah [I- u- u- yeah
009 C1: [(looks at S2))
010 S2: [Ye:s good uhn
011 C1: [°The ski[s are () good [right? So:°
012 S1: [So: [(So:)
013 S2: [Uhn [uhn
014 : Yeah jus' (0.8) big problem ((points to problem areas,
015 S1: Repair make this- () are you Figure 1)
016 : sure you should repairs keps s-
017 C1: °O'key°
018 S2: Uhn [un
019 S1: [No- not buy new ones
020 C1: No=
021 S1: =[No not buy new ones tha- doesn't need it
022 S2: =[U:n never buy new one
023 C1: Great
024 S1: O'kay so we do

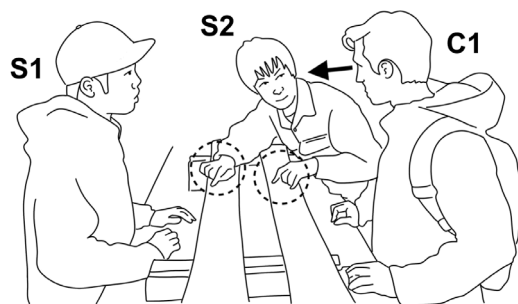


Figure 1: S2 points at problem areas (line 014)

First, in line 001, customer C1 looks at staff member S1, and asks, “Is it worth doing?, ” trying to confirm whether the repair procedure proposed is really necessary. However, in line 002, S1 responds with “Uh?” indicating that he did not understand C1’s request for confirmation in line 001. Here, C1 repeats the initial request by asking “Is it <worth> doing?” in line 003, pronouncing “worth” slowly and prominently. C1 further rephrases the confirmation request using a more common expression, “Should I do it?” The slowing down of speech, rephrasing, and the use of high-frequency expressions are all typical of foreigner talk by L1 speakers. Following a 1.8-second silence in line 004, S1 asks in line 005, “W- wh- what’s doing?” requesting confirmation. S1’s “what’s doing?” appears to be a mishearing of C1’s “worth doing?” in lines 001 and 003. The long silence in line 004 and S1’s request in line 005 make it clear that S1 does not understand C1’s request. In line 006, C1 rephrases the request using a different format: “The skis are still pretty good right?” Notably, the original interrogative sentence is now replaced by a declarative sentence, with “good” pronounced prominently, and “right” added at the end to signal that C1 is seeking confirmation. This shift to a declarative sentence, emphasis on important words, and the use of a tag question are also characteristic of foreigner talk.

Then, after a 0.8-second silence in line 007, S1 responds with “Ye:ah” in line 008, indicating that he understands C1’s request and that the skis can be repaired. C1 then looks at S2 in line 009, the other employee to his right, and S2 says in line 010 “Ye:s good,” repeating the word “good” that S1 emphasized in line 006. S2 confirms that there is no major problem with the skis in response to C1’s request for confirmation. In line 011, C1 repeats the confirmation request with “The skis are () good right?” in a quiet voice, indicating that he has understood the staff’s response and acknowledges that the skis are still in a repairable condition. In line 014, S2 points to the problem areas of the skis with his index fingers and says, “jus’ (0.8) big problem” (Figure 1). S2 informs C1 that there are issues with the skis but suggests that C1 only needs to address the problem areas. It is important to note that when S2 points to the problem areas, both S2 and S1 look at C1 and ensure he understands S2’s explanation.

In lines 015-016, S1 explains to C1 in imperfect English that the skis need repair, and C1 responds with “O’key” in line 017. This sequence repeats in lines 019 and 020, where S1 gives further advice and C1 acknowledges with “No.” Then, in lines 021 and 022, S1 and S2 inform C1 that there is no need to buy new skis, to which C1 responds with an emphatic “Great” in line 023. Finally, in line 024, S1 says “O’kay so we do,” indicating that both parties have agreed to proceed with the repair.

Let us examine Excerpt 1 from the perspective of the asymmetries observed between the employees and customer. First, the English proficiency of S1 and S2 is rudimentary in both speaking and understanding, highlighting a significant gap compared to C1, who is a native speaker. To bridge this gap, C1 made several changes and simplifications in his speech, characteristic of foreigner talk, to make his speech more comprehensible for the staff. However, it should be noted that during the exchange, neither the employees nor the client publicly acknowledged any specific issues with the employee’s English proficiency. This is evidenced by the fact that the staff did not apologize to for their own lack of English, nor did they openly admit to not understanding C1’s English. Similarly, the customer did not express irritation or frustration at the staff’s English skills. Instead, when mutual understanding was not achieved, C1 spontaneously modified and simplified his speech, seeking clear advice on whether the ski needed repair. Despite the language asymmetries, the participants’ mutual understanding and the completion of their immediate tasks were not negatively affected.

The two staff members participated in this exchange as experts, prioritizing accident prevention by providing the correct solution to the customer’s problem, while the customer relied on their expertise to find a quick and cost-effective solution. In this excerpt, since mutual understanding had not been achieved between the two parties, C1 adjusted his speech by rephrasing into simpler expressions, adjusting speech rate, and emphasizing certain points. It is important to note that these adjustments were made to ensure that the customer received accurate information and advice on ski repair, despite the asymmetry between the knowledgeable staff and the customer lacking such expertise.

The above excerpt demonstrates that the Japanese staff in the repair department and the English-speaking customer are more oriented to their roles as participants (customer and employee) and their respective knowledge and experience in ski repair than to any asymmetry in English proficiency. In the next section, we will further explore this point by examining another excerpt.

4.2 Gestures and careful speech by L1 speaker

In Excerpt 2 below, C2, a skier from Australia, requests a snowboard repair, and S1, who also appeared in Excerpt 1 assists him. The entire exchange between C2 and S1 lasts 5 minutes and 5 seconds, with Excerpt 2 focusing on the initial 34-second section of their interaction. In this excerpt, C2's snowboard has a broken metal edge that has popped out from the side. To save money and time, C2 suggests using a patch to cover the damaged area on the board's surface, instead of replacing the metal part of the edge, which is the standard repair method. Although C2 is aware that this simple repair method might not be entirely effective, he seeks S1's opinion whether this approach could be feasible.

[Excerpt 2] C2=Customer 2, S1=Staff 1
001 C2: I just wanna (.) I'd like-
002 : I just wanna cut it near here?
003 S1: Uhuh
004 C2: and not put in the right one an'
005 : [jus' (2.2) put the patch on?
006 : [(covers problem area w/ RH and looks at S1, Figure 2))
007 S1: No- no edges?
008 C2: Yeah no edge the[re?
009 S1: [Just a patch on?
010 C2: Yeah
011 (0.6)
012 C2: Woul- what would happen if I do that
013 S1: Mmmh O:k[:
014 C2: [Would it just be a bi'
015 : sof- sof[ter?
016 : [(moves palms up & down and looks at S1, Figure 3))
017 (0.2)
018 S1: Nn?

L1 Speaker Modification in L1-L2 Conversation

019 C2: Would my board be softer?
 020 S1: If (.) w[e don't put the edges?
 021 C2: [No edge
 022 : Yeah
 023 (1.2)
 024 S1: Weaker
 025 C2: Weaker [ye:s s- ju[s' like (1.0) bendier=
 026 : [((moves palms up and down, Figure 4))
 027 S1: [U:n
 028 : =Un
 029 C2: Yeah an' [might snap easier
 030 : [((gestures board to break, Figure 5))
 031 S1: U:n easier
 032 C2: an' that's it?
 033 (1.2)
 034 S1: That's it=
 035 C2: =Yeah

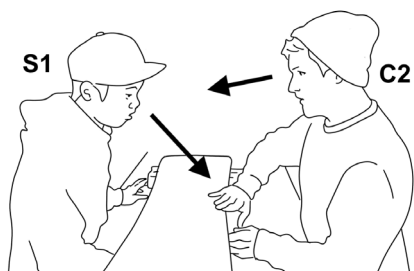


Figure 2: C2 covers problem area (line 006)

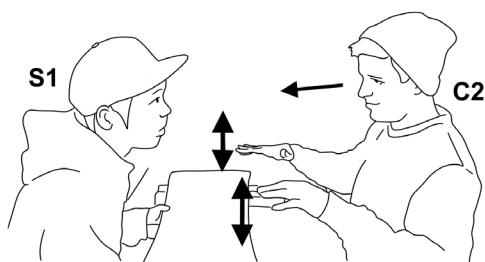


Figure 3: C2 moves palms up and down (line 016)

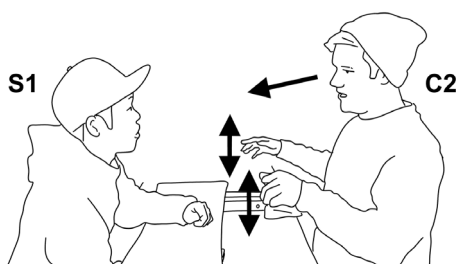


Figure 4: C2 moves palms up and down (line 026)

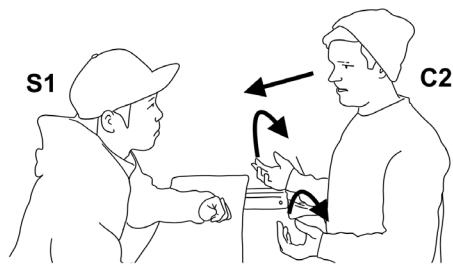


Figure 5: C2 gestures board to break (line 030)

In lines 001-002 and 004-006 of this excerpt, customer C2 informs employee S1 that he prefers a simple repair method. In lines 001 and 002, C2's speech features repetition, a hedge ("just"), a short pause, a cutoff, and rising intonation at the end of the turn, all of which indicate that C2 is cautious about presenting his proposal. Rather than opting for the usual repair method ("right one") mentioned in line 004, C2 suggests using a patch instead in line 005. Once again, there is a significant 2.2-second pause before saying "put the patch on," with the hedging word "just." Furthermore, when C2 says "just" in line 005, he covers the damaged area on the snowboard with the palm of his right hand, simulating the process of putting the patch on the area, and he looks at S1 to see how his suggestion is received (line 006, Figure 2).

In line 007, S1 asks for confirmation with "no edges?" to which C2 responds affirmatively in line 008. However, in line 009, S1 seeks further clarification with "Just a patch on?" and C2 confirms again in line 010. The sequence of confirmation requests from S1, combined with the pause in line 011, indicates that the proposal is difficult for S1 to accept immediately.

In line 012, C2 carefully asks S1 what would happen if such a repair were made, using an utterance that includes a cut-off and paraphrase. When S1's response in line 013 turns out to be not very positive, C2 carefully makes another confirmation request in lines 014-015, asking if the snowboard will soften a little, again with speech containing multiple cutoffs. Midway through the word "softer" in line 015, C2 looks at S1 and gestures by moving his palms down and up to illustrate the softening of the snowboard in line 016 (Figure 3). As in line 006, C2 looks at S1 to see how his suggestion is received by S1.

After a 0.2-second silence in line 017, S1 responds with "Nn?" in line 018, indicating that mutual understanding has not been reached. C2 then repeats in line 019, "Would my board be softer?," paraphrasing the confirmation request of lines 014-015. After the confirmation request and confirmation insertion sequence by S1 and C2 in lines 020-022, S1 responds with "Weaker" in line 024 to C2's confirmation request of line 019, saying that there will be a problem with the strength of the snowboard. C2 then repeats "Weaker" in line 025, rephrasing it with "jus' like (1.0) bendier" and showing how the snowboard becomes bendier through a

gesture of moving both palms up and down (line 026, Figure 4), which is immediately confirmed by S1 in line 028 with the Japanese affirmative token “Un.” C2 further replaces “weaker” and “bendier” with the expression “snap easier” in line 029, and shows how the snowboard is folded with both hands in line 030 (Figure 5) to confirm with S1 if this is correct. Then, in line 031, after the affirmative token “U:n,” S1 repeats the word “easier” used by C2 and confirms C2’s statement. In line 034, C2 says “That’s it?” and tells S1 that even if S1 has such a concern, it is not a big problem for him.

What did we observe in Excerpt 2? First, when C2 was confirming to S1 how the snowboard would be affected by the unusual repair, C2 reformulated a series of expressions related to the strength of the snowboard, along with gestures with both hands, to demonstrate the condition to S1. These reformulations and gestures were used by C2 to ensure that S1 understood his proposed method of repair, and C2 made these linguistic modifications and used embodiment for his own benefit. They were part of the negotiation of the technical aspects of ski repair, and in this regard, the linguistic modifications and embodiment were motivated by the same considerations as C1 in Excerpt 1.

What is noteworthy about this excerpt is that C2’s talk was very carefully crafted as he proposed his own repair method and sought confirmation from S1. This was indicated by numerous cutoffs, paraphrases, repetitions, pauses, and rising tones at the end of sentences in his talk. Furthermore, when making such a request, he looked at S1 to see how his suggestion was received. Conversely, S1 did not immediately accept C2’s alternative suggestion, but carefully examined the condition of the snowboard. Despite communicating in poor English, S1 pointed out to C2 that there were issues with the suggestion. Similar to Excerpt 1, the customer and employee did not indicate that S1’s language ability was a problem; instead, both were oriented toward their participatory roles and expertise.

5. DISCUSSION

After analyzing the above two excerpts, it became clear that the English produced by the staff working in the repair department contained numerous grammatical errors, lacked fluency and naturalness, and they also had experienced difficulties with listening comprehension. Consequently, the native speakers would adjust or simplify their speech when communicating with the Japanese staff. For instance, when Customer C1 realized that the phrase “Is it worth doing?” was not understood by the staff, he would slow down his speech or rephrase it as “Should I do it?” Such linguistic modifications included paraphrasing, slower speech rate, and careful articulation, which are all characteristic of foreigner talk. Through these linguistic adjustments, the English L1 clients aimed to ensure their speech was comprehensible to the employees.

Additionally, as is particularly evident in Excerpt 2, Customer C2, an L1 English speaker, used a series of gestures along with the expressions “softer,” “bendier,” and “snap easier” to describe the condition of his snowboard during negotiations with the staff. This combination of gestures and his simplified speech illustrates another example of foreigner talk employed by L1 speakers when interacting with L2 speakers.

The question to consider here is what the linguistic modifications and embodiment observed in the excerpts were for in this context. For instance, in the case of customer C2 using a series of paraphrases and gestures, these adjustments and embodiment were not primarily aimed at aiding S1, who had limited fluency in English, in understanding the specific vocabulary. Instead, they were employed by C2 to ensure his own precise understanding of the potential impact on the snowboard if standard repairs were not undertaken. Only an employee with repair knowledge and experience could provide insights into this impact, and C2 needed S1 to accurately understand his speech correctly to obtain this information. Therefore, C2’s use of paraphrasing and gestures was directed towards achieving his own practical objectives. The same is true for customer C1 in Excerpt 1.

Comparing the conversation between the English teacher and the learner in the classroom with the conversation in the ski shop yields contrasting insights into the recipient of linguistic modifications by the L1.

The modifications in the classroom are for the L2 learners, whereas in the ski shop they serve the needs of the English L1 customers themselves. The phenomenon of modification occurring for the L2 learner in the classroom can be explained by the asymmetry in linguistic competence between L1 and L2 speakers. However, this asymmetry does not explain the phenomenon of L1 customers making linguistic modifications for themselves in the ski shop. The only way to explain this is to consider asymmetries in participation roles and expertise.

The employees' primary concern was preventing accidents and ensuring customer safety through thorough repairs, and they advised their customers accordingly. Conversely, the customers were primarily focused on the economic rationale for repairs, seeking expert advice from the employees with greater knowledge and experience in repair work. The customers' requests for confirmation were characterized by cutoffs, pauses, paraphrasing, and gazing at the staff. It was observed that the customers communicated with the expert staff in a very cautious manner, based on each other's participatory roles, knowledge, and skills. Despite the staff's rudimentary level of English, which required linguistic correction by the clients, the linguistic asymmetry did not appear to pose a problem for either party.

Historically, the conversational data used to form theories and hypotheses in SLA research have been collected in classrooms and laboratories. In L1-L2 conversations, the teacher or L1 speaker had an advantage not only due to linguistic asymmetry, but also because of participation roles and knowledge. Thus, the issue of asymmetries in participation roles and knowledge was less prominent. However, when we examine everyday conversations that occur outside the classroom, a 'twist' emerges where asymmetries show the linguistically dominant L1 client may lack knowledge and skills. Given that most real L1-L2 conversations take place outside the classroom, there are compelling reasons to reassess theories and practices in SLA research within the social contexts in which speakers are situated.

This study has several implications for research and practice in the teaching and learning of English and other foreign languages. Firstly, it highlights the importance of understanding the realities of English communication, particularly how L1-L2 communication is organized in

authentic contexts, as demonstrated in this paper. By sharing these insights with learners and teachers, they can develop concrete images and practical ideas for English learning and teaching. This approach can help bridge the gap between classroom learning and real-world communication scenarios. In the scenario widely shared by people, L1 speakers are often portrayed as having no problems in any situation, while L2 speakers are seen as always having difficulties and needing assistance from L1 speakers. However, as this study has demonstrated, this perception greatly oversimplifies and distorts the reality of language interactions.

As we have seen in previous sections, linguistic ability alone does not solely dictate the success or failure of conversations and activities in the real world. Conversation participants are not colorless entities; instead, their social roles, knowledge, and skills are projected into the conversation context and play crucial roles in its organization. For example, if the L1 speaker has less knowledge about the topic of the conversation compared to the L2 speaker, it is the L1 speaker who is more likely to encounter difficulties. In such cases, the L1 speaker may be more motivated to accurately understand the L2 speaker's speech and actively engage in the interaction. The participants' roles in interaction, along with their responsibilities and duties, can also serve as resources for organizing the conversation. As Gardner and Wagner (2004) claim, unlike conversational data collected in classroom contexts, the outcomes and consequences of authentic interaction affect individuals' real lives. Whether they are native speakers or not, the success or failure of their conversations directly influences their lives.

Some readers may question whether it is acceptable to leave L2 speakers at a low level of English proficiency if L1-L2 communication can be established in real life. This is not a question with a straightforward answer. For instance, L2 speakers themselves may feel their current ability allows them to perform adequately in their jobs, leading them to believe the status quo is sufficient. On the other hand, others may aspire to achieve a higher level of proficiency. These preferences can vary depending on each individual's social, economic circumstances, and personal perspectives.

While the present study cannot provide a clear answer to the question above, it raises some important issues based on its findings. By describing

and demonstrating how L1-L2 interactions are actually organized, the present study provides an opportunity for researchers, teachers, learners, and the general public to enhance their awareness. The relativization of the linguistic superiority among L1 speakers discussed in this paper may be linked to a cautious approach toward native speakerism, as discussed in the field of English as a Lingua Franca (ELF). Jenkins (2017), for example, characterizes ELF in terms of its scale (number of speakers) and emphasizes that the majority use English as a foreign language, rather than as a second language. Individuals for whom English is a foreign language, such as the repair shop workers in this article, have lower levels of English proficiency compared to native or second language speakers. However, many current English conversations worldwide involve these EFL speakers. For instance, in a scenario a businessperson from an English L1 region wants to sell a product to an Asian company, the L1 speaker must effectively participate in a business meeting, understand the non-standard English of her counterpart, and establish mutual understanding. The insights from this paper should raise people's awareness about the dynamics of global communication.

6. CONCLUSION

Considerable attention has been given to the question of how learners and their teachers should address learners' low levels of proficiency in the target language. However, the English used by non-native speakers is not *deficient* but merely *different* from that of native speakers (Jenkins 2006), and native speakers also bear responsibilities for bridging communication gaps. The prevalence of 'native-speakerism' (Holliday 2005, 2006) may have contributed to the limited acceptance of this idea. As this study has clearly demonstrated, communication is fundamentally a mutual act, requiring the coordination from both parties involved. The assumption or premise that only one side should unilaterally strive for mutual understanding is unnatural, and this argument applies not only to L1-L2 conversations but also to conversations between L1 speakers. Therefore, raising awareness among L1 speakers is crucial. It is hoped that this reconceptualization will be embraced by second language researchers and

Tomoharu Yanagimachi

teachers, as well as the general public. This study investigated conversations between L1 and L2 speakers within a specific social context. In the future, conducting similar studies across diverse social contexts and accumulating relevant knowledge will allow for a reassessment that better aligns with the complexities of L1-L2 conversations.

REFERENCES

- Ferguson, Charles A. 1971. *Language Structure and Language Use: Essays*. Stanford: Stanford University Press.
- Ferguson, Charles A. 1975. Toward a characterization of English foreigner talk. *Anthropological Linguistics* 17(1):1-14.
- Firth, Alan, and Johannes Wagner. 1997. On discourse, communication and (some) fundamental concepts in SLA research. *Modern Language Journal* 81:285-300.
- Gardner, Rod, and Johannes Wagner. (2004). Introduction. In *Second Language Conversations*, eds. Rod Gardner, and Johannes Wagner, pp.1-17. London: Continuum.
- Gass, Susan M., and Larry Selinker. 2008. *Second Language Acquisition: An Introductory Course (3rd ed.)*. New York: Routledge.
- Holliday, Adrian. 2005. *The Struggle to Teach English as an International Language*. Oxford: Oxford University.
- Holliday, Adrian. 2006. Native-speakerism. *ELT Journal* 60:385-387.
- Jenkins, Jennifer. 2006. Current perspectives on teaching World Englishes and English as a Lingua Franca. *TESOL Quarterly* 40:157-181.
- Jenkins, Jennifer. 2017. English as a Lingua Franca in the Expanding Circle. In *The Oxford Handbook of World Englishes*, eds. Markku Filppula, Juhani Klemola, and Devyani Sharma, pp.549-566. New York: Oxford University Press.
- Jordan, Brigitte, and Austin Henderson. 1995. Interaction Analysis: Foundations and practice. *Journal of the Learning Sciences* 4(1):39-103.
- Krashen, Stephen D. 1981. *Second Language Acquisition and Second Language Learning*. Oxford: Pergamon.
- Krashen, Stephen D. 1982. *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon.
- Long, Michael. 1985. Input and second language acquisition theory. In *Input in Second Language Acquisition*, eds. Susan Gass, and Carolyn Madden, pp.377-393. Rowley, M.A.: Newbury House.
- Peräkylä, Anssi. 2011. Validity in research on naturally occurring social interaction. In *Qualitative Research, Third Edition*, ed. David Silverman, pp.365-382. London: Sage.
- Saville-Troike, Muriel. 2012. *Introducing Second Language Acquisition*. Cambridge: Cambridge University Press.
- Tarone, Elaine. 2000. Still wrestling with 'context' in interlanguage theory. *Annual Review of Applied Linguistics* 20:182-198.
- ten Have, Paul. 2007. *Doing Conversation Analysis: A Practical Guide, Second Edition*. London: Sage.

Tomoharu Yanagimachi

*Tomoharu Yanagimachi
School of Letters
Hokusei Gakuen University
Sapporo 004-8631 Japan
tyana@hokusei.ac.jp*

APPENDIX

Transcript conventions (based on ten Have 2007)

1. Sequencing

- [A single left bracket indicates the point of overlap onset.
- = Equal signs, one at the end of one line and one at the beginning of a next, indicate no 'gap' between the two lines.

2. Timed intervals

- (0.0) Numbers in parentheses indicate elapsed time in silence by tenth of seconds,
- (.) A dot in parentheses indicates a tiny 'gap' within or between utterances.

3. Characteristics of speech production

- word Underscoring indicates some form of stress, via pitch and/or amplitude.
- :: Colons indicate prolongation of the immediately prior sound. Multiple colons indicate a more prolonged sound.
- A dash indicates a cut-off.
- ? A question mark indicates a rising intonation.
- ° ° Utterances or utterance parts bracketed by degree signs are relatively quieter than the surrounding talk.
- < > Left/right carets bracketing an utterance or utterance-part indicate slowing down.
- () Empty parentheses indicate the transcriber's inability to hear what was said. The length of the parenthesized space indicates the length of the untranscribed talk.

4. Transcriber's doubts and comments

- (word) Parenthesized words are especially dubious hearings or speaker identifications.
- (()) Double parentheses contain transcriber's descriptions rather than, or in addition to, transcriptions.

L1-L2 對話中的 L1 話者修改為誰而做？：
從互動分析的角度重新概念化外國人講話

柳町智治
北星學園大學

先前關於母語為第一語言者在母語-第二語言對話中語言修改的研究，大多數依賴於在 ESL/EFL 課堂或實驗環境中收集的數據。然而，大多數母語-第二語言的互動發生在日常互動情境中，而非如此受控的環境。本研究採用互動分析的框架，探討母語為英語的人士如何在日本的滑雪用品店中進行語言修改，即所謂的「外國人講話」。分析基於錄影資料，記錄了英語使用者顧客與日本店員交流，詢問如何修理滑雪板。資料中反覆出現的現象是顧客調整他們的言談，以達成彼此理解，並引出店員對於他們滑雪板狀況及必要修理工作的專業評估與建議。員工與顧客都致力於對方的參與角色、知識和技能，而非彼此間的語言不對等。研究發現提出了一些教學上的意涵，並能提高教師、學習者、母語使用者和研究者的意識。

關鍵字：言談修改、外國人講話、真實性、社會情境、互動分析、錄影資料

PROBING THE IMPACT OF TIME-RELATED FACTORS ON KOREAN LEARNERS' PROFICIENCY IN ACQUIRING THE VIETNAMESE LANGUAGE

Trang Phan¹

Ca' Foscari University of Venice

Lan Chu

VNU University of Languages & International Studies

ABSTRACT

In recent years, the interest in learning Vietnamese among foreign learners, particularly Koreans, has surged, driven by the need for effective communication and professional growth. This study examines the factors influencing Vietnamese language acquisition among 63 Korean students, comprising 38 males and 25 females, from Hankuk University of Foreign Studies in Seoul and the Kmove Program at VNU University of Languages & International Studies in Hanoi. The research focuses on three key areas: placement in language classes, self-assessed linguistic capabilities, and a multiple-choice test on the verb đi 'to go' in directed motion constructions. This test highlights key differences between Korean and Vietnamese, serving as a valuable proficiency metric. Additionally, time-related variables such as extended exposure to a native-speaking environment (over 6 months), duration of language study (over 2 years), and early start age (under 12 years) are crucial in language acquisition. These findings aim to provide insights for teaching Vietnamese to Korean learners and other foreigners engaging with Vietnam's evolving cultural landscape.

Keywords: Vietnamese as a second language, Korean, time-related factor

¹ We are thankful to the audience of UNC (ULIS national conference), held by VNU University of Languages & International Studies, Vietnam National University Hanoi on April 24, 2023, and COTSEAL (Council of Teachers of Southeast Asian languages) 39th Conference held by University of Wisconsin-Madison on July 21-22, 2023. For academic purposes, Trang Phan is responsible for Sections 2, 3.2, 4.1, 4.3, 5. Lan Chu is responsible for Sections 1, 3.1, 4.2, 4.4.

1. INTRODUCTION

Recently, the demand for learning Vietnamese from foreign, particularly Korean, learners has been on the rise. The majority of these learners aim to achieve fluency in the language for work and communication purposes. This study focuses on a sample of 63 Korean students, who were enrolled at Hankuk University of Foreign Studies in Seoul and The Kmove Program based at VNU University of Languages & International Studies in Hanoi. The objective of this research is to examine the time-related factors influencing Vietnamese language acquisition among Korean learners. In addition to differences in linguistic typology, three key time-related variables were identified as impacting the acquisition process. These variables include the duration and intensity of exposure to a Vietnamese-speaking environment, the overall length of time committed to Vietnamese language study, and the age at which the study of Vietnamese begins. These preliminary findings aim to offer practical insights for teaching Vietnamese to Korean learners, as well as to foreigners in general.

2. WHY KOREAN LEARNERS OF VIETNAMESE?

The Korean and Vietnamese languages were chosen for this study due to their typological differences.

In particular, Vietnamese and Korean exhibit typological differences, particularly in their word order patterns. Vietnamese is classified as a head-initial language, where the head (e.g., the verb) is positioned before the modifier (e.g., the object) and the location is preceded by the preposition (Cao 2003; Duffield 2013, Phan and Duffield 2022), as seen below:

- (1) a. Tôi đọc sách
 1SG read book
 'I read a book.'
 b. Tôi đi đến công viên
 1SG go to park
 'I go to the park.'

In contrast, Korean is considered a head-final language, with the modifier (e.g., the object) preceding the head (e.g., the verb) and the location being followed by the postposition (Cho and Whitman 2020), as illustrated in (2).

- (2) a. 나는 책을 읽어요.
Na-neun cha-keul ik-eo-yo
1SG-TOP book-ACC read-PRES
'I read a book.'
- b. 나는 공원에 가요.
Na-neun gong-won-e ga-yo
1SG-TOP park-DAT go-PRES.
'I go to the park.'

Korean and Vietnamese are also different in terms of the expression of directed motion. Talmy (1985, 2000) categorizes languages worldwide into two contrasting typological frames: verb-framed languages such as Spanish, Japanese, and Turkish where the direction element is marked directly on the main verb and satellite-framed languages such as English where the direction element is marked within the complement of the verb.

Korean is typically classified as a verb-framed language (Choi and Bowerman 1991; Choi 2006, 2009) due to the directional element marked in the verb, whereas Vietnamese, according to the works of Brown (1999); Nguyen (2001); Beecher (2004); Pace (2009) and Nguyen (2019), can be classified as both a satellite-framed language and a third type known as equipollently-framed language, as proposed by Slobin (1996, 2004, 2006), in which direction is encoded in serial verbs. The following two sentences illustrate the differences in the typology of Vietnamese and Korean based on the criterion of directionality in motion constructions:

- (3) a. 병이 동굴 속으로 떴다
Byeong-i dong-gul sokeuro tteotda
bottle-SUB cave inside-ADV float-PAST
'The bottle floated into the cave.'
- b. Cái chai trôi vào trong hang
CLF bottle float enter in cave
'The bottle floated into the cave.'

The examples in (3) illustrate three difficulties that Korean learners of Vietnamese may encounter. First, the verb 'float' is head-final in Korean, whereas it is head-initial in Vietnamese. Second, in Korean, the location 'cave' comes before the postposition, whereas in Vietnamese, it follows the preposition. And finally, the direction is encoded within the verb 'float' in Korean but is expressed separately from the verb in Vietnamese. Due to space constraints, detailed discussion is omitted here. For a theoretical examination of the typological and language-specific differences between Vietnamese and Korean motion constructions, please refer to Phan, Basciano and Chu (2024). These typological differences make the acquisition of Vietnamese by Korean learners an interesting subject for examination.

3. PARTICIPANTS AND PROFICIENCY MEASUREMENT

3.1. Participants

This study was conducted on 63 Korean students with a ratio of 38 males and 25 females. Following ethical guidelines for research involving human participants, we recruited participants from two distinct sources: undergraduate students enrolled in the Vietnamese Language program at Hankuk University of Foreign Studies in Seoul, and participants from the Kmove Program, operated at VNU University of Languages & International Studies in Hanoi at the time of testing.

Regarding the students of Hankuk University of Foreign Studies, these are full-time students specializing in Vietnamese within the Vietnamese Language Program. Over the span of four years, students will undergo a comprehensive study of Vietnamese language, ranging

from fundamental to specialized aspects. The ultimate objective is for students to attain advanced proficiency in professional and academic Vietnamese and develop a solid understanding of Vietnamese culture and society. Upon obtaining an BA degree, students will have various career prospects, including Vietnamese language interpreters, Korean-language teachers for Vietnamese learners, and researchers in the field of Vietnamese studies, etc.

The Kmove Program, on the other hand, is part of a project initiated by the Human Resources Development Service of Korea (HRD Korea), an agency under the Ministry of Labor of Korea. HRD Korea is the managing and funding entity responsible for training Korean learners who will work in Korean companies in Vietnam as key managerial personnel. According to the plan, the learners will be taught Vietnamese language for a duration of two months in Korea, followed by a six to seven-month training period in Vietnam. The training program in Vietnam includes instruction in Vietnamese language, accounting, and the current economic laws applicable in Vietnam. Upon graduation, the learners will be provided with free referrals to Korean companies operating in Vietnam. During their stay in Vietnam, the learners will study approximately 20 hours per week for a period of six to seven months, with the aim of achieving an intermediate level or higher proficiency. The objective of the course is to enable the learners to effectively use the Vietnamese language in workplace communication and daily life situations.

3.2. Proficiency measurement

In order to assess the proficiency levels of Korean students, we considered three distinct factors: (1) their placement in appropriate language classes, (2) their self-reported evaluation of their linguistic capabilities, and (3) a multiple-choice assessment that centers on the placement of the verb *đi* 'to go' in directed motion constructions.

Firstly, we took into account their placement in appropriate language classes. For Seoul-based students, at the time of testing, they were second-year students at Hankuk University of Foreign Studies. On the other hand, Hanoi-based students had been part of the Kmove program

in Hanoi for 6 months. Hence, all of them belonged to the intermediate to advanced level.

Secondly, we evaluated their linguistic capabilities through their self-reported assessments. We requested them to rank their proficiency approximately according to the ACTFL scale. For an introduction to the ACTFL scale and its guidelines for Vietnamese oral tests, see Phan and Chu (2024). Specifically, they had four options: Intermediate Low, Intermediate Mid, Intermediate High, and Advanced Low. Notably, none of them self-identified as Intermediate Low.

Finally, a multiple-choice survey of 33 sentences containing the verb *đi* ('to go') was conducted. The 33 sentences were divided into directed and non-directed motion constructions. The directed motion constructions was further categorized into Goal-oriented, Source-oriented, and Route-oriented. Examples of each category is given below.

(4) Goal-oriented motion constructions:

- a. Hoa đi chợ
Hoa go market
'Hoa goes to the market.'
- b. Hoa 씨는 시장에 간다.
Hoa-ssi-neun si-jang-e gan-da.
Hoa-POL-TOP market-DAT go-PRES
'Hoa goes to the market.'

(5) Source-oriented motion constructions:

- a. Hoa đi từ chợ
Hoa go from market
'Hoa goes from the market.'
- b. Hoa 씨는 시장에서 간다.
Hoa-ssi-neun si-jang-eseo gan-da.
Hoa-POL-TOP market-LOC go-PRES
'Hoa goes from the market.'

(6) Route-oriented motion constructions:

- a. Hoa đi qua chợ
Hoa go across market
'Hoa goes across the market.'
- b. Hoa 씨는 시장을 지나간다.
Hoa-ssi-neun si-jang-eul ji-na-gan-da.
Hoa-POL-TOP market-ACC pass-PRES
'Hoa goes across the market.'

(7) Non-directed motion constructions:

- a. Hoa đi mua thịt ở chợ
Hoa go buy meat at market
'Hoa goes buy meat at the market.'
- b. Hoa 씨는 시장에서 고기를 산다.
Hoa-ssi-neun si-jang-eseo go-gi-reul san-da.
Hoa-POL-TOP market-LOC meat-ACC buy-PRES
'Hoa goes buy meat at the market.'

Korean learners of Vietnamese may encounter a number of challenges, at least three of which deserve particular attention. Firstly, there is a notable difference in the verb's position between the two languages, with Korean exhibiting a head-final pattern while Vietnamese follows a head-initial pattern. Secondly, the ordering of location and postposition varies between Korean and Vietnamese, as Korean places the location before the postposition whereas Vietnamese follows a preposition followed by the location. Lastly, the encoding of direction is internal to the main verb in Korean, whereas in Vietnamese, it is expressed as a separate element from the verb. These typological distinctions present intriguing areas for investigation when examining the acquisition of Vietnamese by Korean learners. The mastery of directed motion constructions in Vietnamese by Korean students signifies their ability to surpass the influence of their first language (L1) and acquire second language (L2) patterns that deviate from their L1.

Phan, Basciano, and Chu (2024) offer a comprehensive discussion on the typological and language-specific variations between Vietnamese and Korean motion constructions, which form the theoretical basis of our

study. Our survey aims to test the linguistic construct of motion constructions, focusing on the verb *đi* 'to go', to assess the proficiency of Korean learners in Vietnamese. This construct was chosen because it highlights significant syntactic and semantic differences between the two languages.

The data was collected through a survey conducted on Google Forms and through direct surveys in class. The survey consisted of 33 multiple choice questions, each with 3 choices, of which 1 was the correct option regarding the order of the verb *đi* 'to go' in relation to other surrounding elements in Vietnamese motion constructions.

For participants who self-identified as Intermediate Mid and achieved approximately 50% correct answers out of the 33 stimulus sentences, they were categorized as Intermediate Mid. Self-identifying as Intermediate High and achieving approximately 70% correct answers results in classification as Intermediate High, while self-identifying as Advanced Low and achieving approximately 90% correct answers leads to being labeled as Advanced Low. Furthermore, we also took their placement in appropriate language classes into account.

We acknowledge that our approach is not ideal, however, it serves our specific purpose. It provides a more precise measure of learners' linguistic competence in areas where Korean and Vietnamese diverge significantly. By analyzing responses to a specific linguistic construct, we aim to gain deeper insights into the learners' proficiency and the impact of their exposure to the L2 speaking environment.

Based on the findings, out of the 63 participants, 28 were classified as Intermediate Mid learners, 25 as Intermediate High learners, and 10 as Advanced Low learners.

Table 1. Participants' proficiency level

Proficiency level	Number of participants
Intermediate Mid	28
Intermediate High	25
Advanced Low	10

4. RESULTS AND DISCUSSION

The survey findings demonstrate the influence of extralinguistic time-related factors on the Vietnamese language learning process among Korean students. These factors encompass the length and intensity of exposure to a native Vietnamese-speaking environment, the overall duration dedicated to Vietnamese language study, and the age of initiation in learning Vietnamese. We present the results for each variable and discuss their implication sequentially.

4.1. Length and intensity of exposure to a native Vietnamese-speaking environment

The initial time-related factor under consideration is the duration and intensity of exposure to a native Vietnamese-speaking environment. The impact of time spent in an immersive environment where the target language is spoken has garnered significant attention and has become a focal point of investigation in the existing literature on second language acquisition (see Gu 2018 for a more recent review). There is a rapid growth of research comparing language gains attained from time spent abroad with those achieved in other learning contexts, including foreign language settings and domestic immersion programs. Researchers have dedicated their efforts to examining the variations in the acquisition of different aspects of language ability across these learning contexts. These aspects include listening skills, reading proficiency, grammar and vocabulary knowledge, oral proficiency, writing skills, and pragmatic competence (Freed 1995; Llanes 2011; among others). The findings so far support the hypothesis that a longer duration of stay in a foreign country positively correlates with increased linguistic competence in the respective language.

By integrating these insights, we aim to shed light on the specific influence of exposure duration and intensity to a native Vietnamese-speaking environment on the acquisition process of grammar knowledge (i.e. knowledge of directed motion) among Korean learners. In particular, this study categorizes this time-related factor into two groups: those who have never been to Vietnam, those who stayed in Vietnam for

at least 6 months. This classification aims to investigate the impact of living in Vietnam, as opposed to not residing there, on the acquisition of the Vietnamese language. To explore this, we examined two groups of Korean learners, comprising students from Seoul and Hanoi. The study examines both the duration and intensity of L2 exposure on language acquisition, a point emphasized by a reviewer, highlighting the significance of immersive experiences.

The analysis involved 33 stimulus sentences administered to two groups of participants: 36 individuals who had not visited Vietnam and 27 who had spent at least six months in the country. The former group provided a total of 1188 responses, while the latter contributed 891 responses. This comparison aims to evaluate whether prolonged and intense exposure to authentic Vietnamese in a native environment yields a better language acquisition outcomes.

Table 2. Participants' length and intensity of exposure to a native Vietnamese-speaking environment

	Number of students	Number of answers	Number of correct answers
Never been to Vietnam	36	1188	945
Been to Vietnam	27	891	757

Among those who hadn't been to Vietnam, 945 responses were correct, equating to an 80% accuracy rate. Conversely, participants with over six months of exposure to Vietnam achieved an 85% accuracy rate, with 757 correct responses out of 891. A one-tailed p-value of 0.0008 indicates statistical significance.

Time-Related Factors in Korean Learners' Vietnamese Acquisition

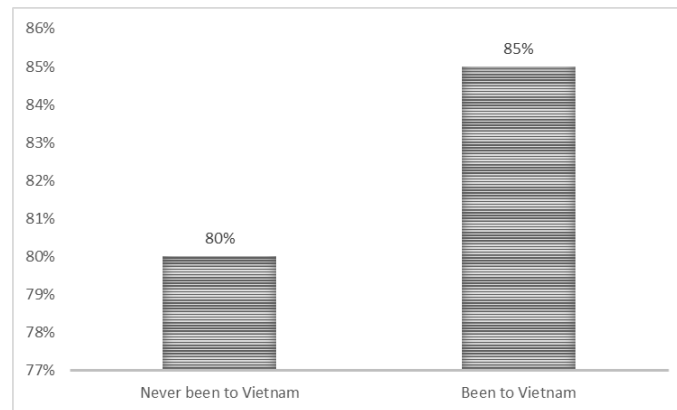


Figure 1: Length and intensity of exposure to a native Vietnamese speaking environment and Accuracy rate

The findings underscore the significant influence of immersion in a language-speaking environment on language acquisition.² Exposure to an immersive environment where the target language is spoken, even for a relatively short period such as six months, can positively affect the acquisition of motion constructions. Specifically, these results emphasize the importance of experiencing a native Vietnamese-speaking environment for Korean learners to effectively acquire and use the motion constructions. They highlight the beneficial impact of immersion in Vietnam for language development, indicating that authentic exposure provide greater opportunities for learners to internalize and utilize the linguistic patterns associated with motion constructions. These findings have implications for language education and curriculum planning. They suggest that incorporating opportunities for immersive experiences and

² As one of the anonymous reviewers suggested, it is also important to consider the medium of instruction used for teaching Vietnamese to Korean speakers in Korea. This raises the question of whether bilingual instruction, which may involve using Korean alongside Vietnamese, significantly impacts language acquisition compared to exclusive immersion in an L2-speaking environment. The findings suggest that while extended study in Korea can facilitate learning, meaningful progress in acquiring nuanced language skills, particularly in using motion constructions, might require the intensity and authenticity of immersion in a Vietnamese-speaking environment. This is a compelling area for future research, and we leave it for further exploration.

creating study abroad programs or language exchange programs can greatly enhance the acquisition of motion constructions and overall proficiency in Vietnamese among Korean learners.

4.2. Overall length of time of learning Vietnamese

It can be observed that learning a foreign language requires a considerable investment of time and effort. Determining the exact amount of time required to reach a desired level of proficiency in a foreign language can be influenced by various factors, including individual abilities, motivation, learning strategies, and the complexity of the language itself. Although estimating the timeline for language proficiency is challenging, educators and researchers have made efforts to provide approximate guidelines. For instance, Hakuta et al. (2000) presents research findings on the length of time language minority students require special services, such as ESL courses or bilingual education, to develop oral and academic English proficiency. The study examines four school districts in California and Canada and concludes that it typically takes 3 to 5 years to develop oral proficiency and 4 to 7 years to achieve academic English proficiency, even in the most successful districts. Their research report challenges policies that assume rapid language acquisition within a short period and emphasizes the need for realistic expectations and adequate language instruction. Realistic expectations should acknowledge the gradual nature of language acquisition, avoiding the assumption that learners can achieve proficiency quickly. Adequate language instruction must include sustained, high-quality teaching that supports both oral and academic language development. This involves not only immersive exposure to the target language but also structured learning environments that cater to the individual needs of learners, providing ample opportunities for practice, feedback, and cultural integration.

It is widely acknowledged that the longer one studies a foreign language, the higher the likelihood of reaching proficiency. Normally, simply investing a short amount of time does not guarantee adequate language skills. The current study conducted among Korean students examined the relationship between the duration of studying Vietnamese

and the level of language output achieved. The results show distinct average scores for two different categories: less than 2 years of study, and more than 2 years of study.

In the analysis, 33 stimulus sentences were presented to two participant groups: 38 individuals who had studied Vietnamese for less than 2 years and 25 who had studied for at least 2 years. The former group generated a total of 1254 responses, while the latter contributed 825 responses.

Table 3. Participants' length of exposure to a native Vietnamese-speaking environment

	Number of students	Number of answers	Number of correct answers
Studied Vietnamese for less than 2 years	38	1254	1000
Studies Vietnamese for at least 2 years	25	825	702

Among participants with less than two years of Vietnamese study, 1000 responses were correct, resulting in an 80% accuracy rate. Conversely, those with over 2 years of Vietnamese study achieved an 85% accuracy rate, with 702 correct responses out of 825. A one-tailed p-value of 0.0001 was obtained, indicating statistical significance in the relationship between study duration and language proficiency level.

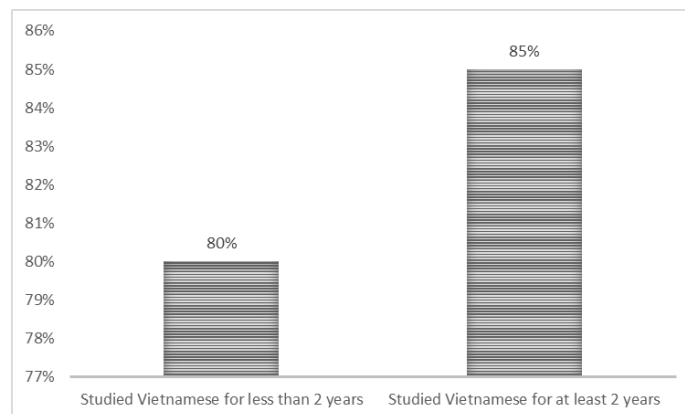


Figure 2. Overall length of time of learning Vietnamese and Accuracy rate

These results highlight a clear trend: as the duration of study increases, the average level of language output also increases. This suggests that longer periods of study provide students with more opportunities to practice and develop their language skills, resulting in improved proficiency and output.

These findings emphasize the importance of investing sufficient time and effort in language learning. For Korean students interested in achieving higher levels of proficiency in Vietnamese, the results suggest that studying for more than 2 years yields the best outcomes. It implies that language learners should be patient and persistent, understanding that language acquisition is a gradual process that requires dedicated time and practice. Furthermore, educators and language learning institutions can take these results into consideration when designing language programs. They can provide students with realistic expectations regarding the time and effort required to achieve specific language proficiency goals. Additionally, the findings suggest that longer-duration language courses or extended study opportunities may be more effective in facilitating language output and proficiency. Overall, the study's results underscore the positive relationship between the duration of language study and language output, highlighting the significance of prolonged and consistent effort in language learning to achieve higher levels of proficiency.

4.3. The relationship between the length and intensity of exposure to a Vietnamese-speaking environment and the duration of study

The current study not only examined the relationship between the duration of studying Vietnamese and the level of language output but also established a link between the duration of study and the time spent living in a native environment, specifically in Vietnam. This additional factor further influences the accuracy rate, leading to variations among different groups of learners.

Table 4. Participants' length and intensity of exposure to a native Vietnamese speaking environment & the duration of study

		Number of students	Number of answers	Number of correct answers
Never have been to Vietnam	Studied Vietnamese for less than 2 years	23	759	602
	Studies Vietnamese for at least 2 years	13	429	343
Been to Vietnam for at least 6 months	Studied Vietnamese for less than 2 years	15	495	398
	Studies Vietnamese for at least 2 years	12	396	359

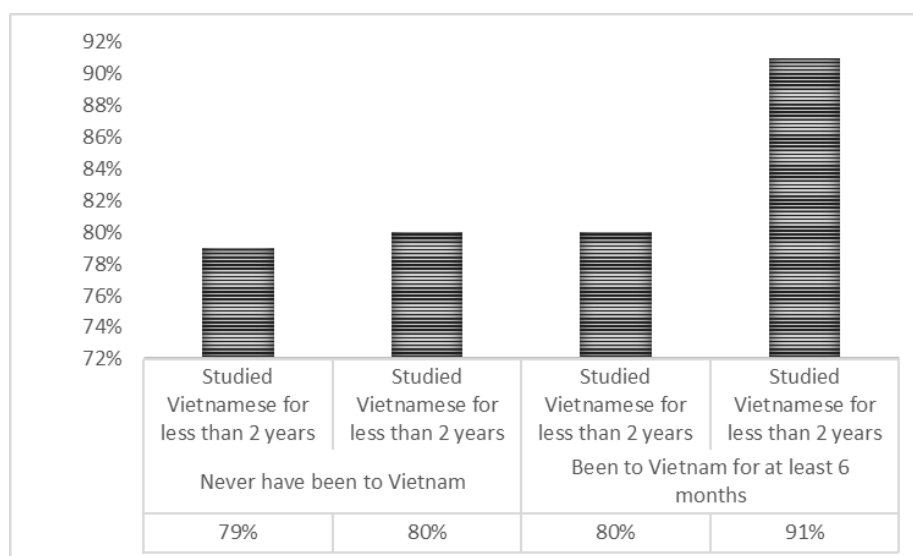


Figure 3. The length and intensity of exposure to a native Vietnamese speaking environment & the duration of study and Accuracy

The results, as presented in the table, show two categories based on whether the learners have been to Vietnam or not. Within each category,

the accuracy percentages are further differentiated based on the duration of study: less than 2 years, and at least 2 years.

Comparing the scores for each category, we can observe interesting patterns. Among those who have never been to Vietnam, 23 individuals studied Vietnamese for less than 2 years, yielding a total of 759 answers, with 602 being correct, resulting in an accuracy rate of 79%. Additionally, 13 individuals studied Vietnamese for at least 2 years, obtaining a total of 429 answers, with 343 being correct, yielding an accuracy rate of 80%. These findings suggest that among those who have never been to Vietnam, there is no significant difference between those who study Vietnamese for more than 2 years compared to those who study for less than 2 years. It seems challenging to make significant progress solely by studying Vietnamese in South Korea without experiencing immersion in Vietnam.

Among those who have been to Vietnam for at least 6 months, 15 individuals studied Vietnamese for less than 2 years, garnering a total of 495 answers, of which 398 were correct, resulting in an accuracy rate of 80%. In contrast, 12 individuals studied Vietnamese for at least 2 years, accumulating a total of 396 answers, with 359 being correct, yielding an impressive accuracy rate of 91%. The analysis reveals a significant disparity (one-tailed p -value < 0.001) between individuals who have been to Vietnam for 6 months or more and have studied Vietnamese for over 2 years versus those who have been to Vietnam for the same duration but have studied Vietnamese for less than 2 years. This underscores the substantial impact of prolonged study duration on language proficiency, particularly when coupled with real-world immersion experiences.

When comparing individuals with the same duration of study (less than 2 years), there is no significant difference between those who have never been to Vietnam and those who have spent at least 6 months in the country. However, it is only when the study duration extends to 2 years that a significant difference is observed between individuals who have never visited Vietnam and those who have spent 6 months there (one-tailed p -value < 0.0001). This suggests that to make substantial progress, students should immerse themselves in a native environment and at the same time engage in long-term study.

The implications of these findings highlight the significance of both study duration and exposure to the native environment in language acquisition. Simply dedicating a certain amount of time to studying the language is beneficial, but the results suggest that learners who have the opportunity to live in Vietnam, even for a short duration, tend to achieve higher levels of language output and proficiency. This suggests that immersion in the native environment provides valuable opportunities for practicing the language, experiencing cultural nuances, and enhancing overall language skills. Educational institutions and language programs can consider incorporating immersive experiences or study abroad opportunities to complement formal language instruction. This can provide learners with the chance to apply their language skills in real-life contexts, further accelerating their language acquisition process. Overall, the results highlight the importance of both study duration and exposure to the native environment, shedding light on the benefits of combining formal study with immersion experiences to maximize language proficiency.

4.4. Age of onset

The impact of age of onset on language learning has been extensively discussed in the literature, with numerous studies highlighting differences in foreign language acquisition across various age groups (Krashen 1973; Krashen et al. 1979; Harley 1986; Harley and Hart 1997; Singleton and Ryan 2004; Abrahamsson and Hyltenstam 2009; Birdsong 2018). Patkowski (1980) examined the correlation between age and phonetic perception in language learning, specifically comparing teenagers and adults. The results indicated that teenagers achieved a higher level of proficiency in the target language compared to adults, supporting the theory that age influences the ability to master phonetics in a second language. Wu (2022) proposed the perspective of "the younger the better" to explore the relationship between age and foreign language learning, acknowledging that age does affect the ability to acquire a second language, but emphasizing that being young does not guarantee an advantage in learning a foreign language at any given time.

Our research on the acquisition of Vietnamese by Korean students with different age groups has identified three variables: under 12 years old, 12-18 years old, and over 18 years old. The results are shown as follows:

Table 5. Participants' age of onset

	Number of students	Number of answers	Number of correct answers
< 12 years old	8	264	227
12-18 years old	9	297	244
>18 years old	46	1518	1231

Among the students examined, 8 individuals commenced studying Vietnamese before the age of 12, amassing a total of 264 answers, of which 227 were correct, yielding an accuracy rate of 86%. Additionally, 9 individuals began studying Vietnamese between the ages of 12 and 18, gathering a total of 297 answers, with 244 being correct, resulting in an accuracy rate of 82%. Furthermore, 46 students initiated their Vietnamese studies after the age of 18, accumulating a total of 1518 answers, with 1231 being correct, yielding an accuracy rate of 81%.

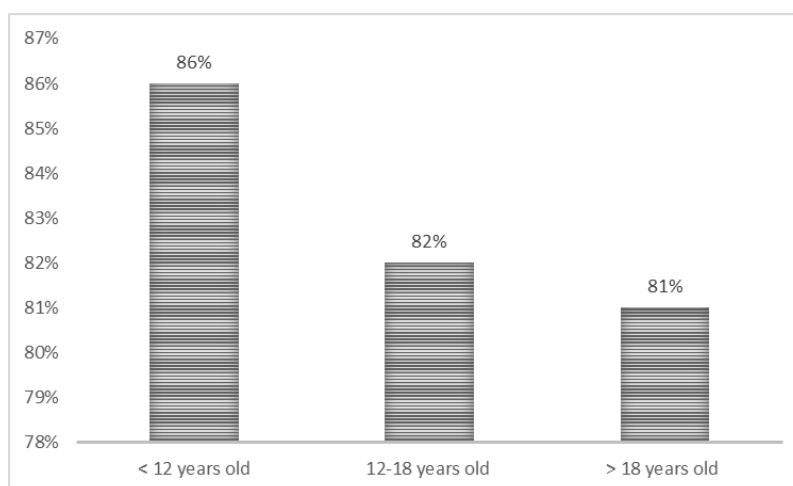


Figure 4: Age of onset and Accuracy

The analysis uncovers a significant discrepancy (one-tailed p -value < 0.05) between individuals who commenced their Vietnamese studies before the age of 12 and those who began after the age of 12.

These findings suggest that individuals under the age of 12 exhibit a greater ability to absorb and assimilate a new language, as evidenced by their higher average scores compared to those over 12. This information holds valuable implications for teaching Vietnamese to native Korean speakers and other foreign language learners. Recognizing the critical age for acquiring new language grammar becomes a strategic consideration for educators (Johnson and Newport 1989, Herschensohn 2007).

The findings underscores the benefit of initiating language acquisition during the pre-adolescent stage. The superior average scores evident in the group under the age of 12 signify their heightened language proficiency. This observation resonates with the concept of a sensitive period or critical phase for language learning, during which individuals in their pre-teen years exhibit heightened cognitive capacities and aptitude for language acquisition. Leveraging this critical phase can maximize the efficacy of language learning endeavors. The implications of these findings extend to language pedagogy, especially in programs for young Vietnamese learners in different countries around the world. Educators should pay particular attention to the age factor when teaching Vietnamese to native Korean speakers and other foreign language learners. Designing language programs that cater to the specific needs and abilities of different age groups can enhance the effectiveness of language instruction. Moreover, acknowledging the critical age range for language acquisition can inform curriculum development, instructional strategies, and assessment methods.

It is essential to note that individual differences, motivation, and exposure to the language also contribute to language learning outcomes, and age alone should not be the sole determinant. However, the present study underscores the significance of early language learning and the potential advantages associated with language acquisition during the pre-adolescent period.

Further research could delve into the underlying cognitive and neurological factors that contribute to the observed age-related differences in language proficiency. Additionally, exploring the impact of age on different aspects of language learning, such as pronunciation, vocabulary acquisition, and pragmatic skills, could provide a more comprehensive understanding of the relationship between age and language acquisition.

In summary, the study's analysis highlights the superior language proficiency observed in individuals under the age of 12 compared to those over 12. This information has practical implications for teaching Vietnamese to native Korean speakers and other foreign language learners. Acknowledging the critical age range for language acquisition and tailoring instructional approaches accordingly can optimize language learning outcomes.

5. CONCLUSION

This study offers practical insights into how Korean learners acquire proficiency in the Vietnamese language. We collected a unique dataset including 63 Korean students (38 males and 25 females), enrolled at Hankuk University of Foreign Studies in Seoul and the Kmove Program at VNU University of Languages & International Studies in Hanoi. To assess the proficiency levels of Korean students, three distinct factors were considered: (1) placement in appropriate language classes, (2) self-reported evaluation of linguistic capabilities, and (3) a multiple-choice assessment focused on the placement of the verb *đi* 'to go' in directed motion construction. This assessment holds particular significance as it reveals notable differences between Korean and Vietnamese, making it a valuable metric for measuring language proficiency in Korean learners.

In addition to the linguistic typological disparities between the two languages, several time-related variables were examined for their influence on Vietnamese language acquisition. These variables include the duration and intensity of exposure to a Vietnamese-speaking environment (preferably exceeding 6 months), the overall length of time dedicated to studying Vietnamese (preferably exceeding 2 years), and

the age at which learners commence studying Vietnamese (preferably younger than 12 years old).

The implications of the above results are significant for the teaching and learning of Vietnamese among Korean learners, as well as for foreigners in general who are interested in engaging with the Vietnamese language and culture.

Firstly, the findings underscore the importance of considering time-related factors in Vietnamese language acquisition. The study highlights the influence of variables such as the length of exposure to a Vietnamese-speaking environment, the overall duration of Vietnamese language study, and the age at which learners commence their Vietnamese language journey. These factors should be taken into account when designing language programs and curricula, as well as when providing guidance to Korean learners and foreign individuals who aspire to achieve proficiency in Vietnamese. Teachers and educational institutions can use this information to create effective learning pathways and tailor instruction to optimize language acquisition outcomes.

Secondly, the recognition of the significant differences between Korean and Vietnamese, particularly evident in the assessment focused on the placement of the verb *đi* 'to go' in directed motion construction, highlights the areas where Korean learners may face specific challenges. Educators can use this knowledge to design targeted instructional strategies that address these specific linguistic differences and promote accurate language usage. By addressing these areas of difficulty, teachers can facilitate the development of greater language proficiency among Korean learners.

Additionally, the findings emphasize the importance of creating immersive language learning environments and opportunities for Korean learners to engage with native Vietnamese speakers. Extended exposure to a Vietnamese-speaking environment for at least six months was identified as a beneficial factor for language acquisition. This suggests that study abroad programs, cultural exchanges, and language immersion experiences can play a crucial role in enhancing Vietnamese language proficiency.

Furthermore, the age at which learners begin their Vietnamese language study also emerges as a significant factor. The ideal age range of younger than 12 years old is associated with greater language

acquisition potential. These findings can inform decisions regarding when to introduce Vietnamese language learning opportunities within the educational system or language programs, maximizing the chances of successful language acquisition.

Given the limited studies on Vietnamese as a second language, this paper offers valuable insights for L2 teaching professionals and policymakers, especially for the development of Vietnamese programs and exchange initiatives globally. This research highlights the necessity of immersive experiences and suggests that program designers should incorporate substantial in-country exposure to enhance language acquisition. Our data offer a unique contribution by specifically addressing the acquisition of Vietnamese by Korean learners, an underexplored area. This paper serves as a foundational reference for future research and development in Vietnamese language education, providing actionable insights for improving L2 Vietnamese teaching methodologies and program effectiveness.

REFERENCES

- Abrahamsson, Niclas, and Kenneth Hyttenstam. 2009. Age of onset and native likeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning* 59(2):249-306.
- Beecher, Henry. 2004. *Three varieties of serial verb constructions in Vietnamese*. Unpublished Manuscript. University of California San Diego.
- Birdsong, David. 2018. Plasticity, variability and age in second language acquisition and bilingualism. *Frontiers in psychology* 9:81.
- Brown, Linh-Chan. 1999. *The grammaticalization of directed motion in Vietnamese*. Boulder: University of Colorado dissertation.
- Cao, Xuan Hao. 2003. Tiếng Việt. Mấy Vấn đề Ngữ Âm, Ngữ Pháp, Ngữ Nghĩa [Vietnamese: some phonetic, syntactic, and semantic issues]. Đà Nẵng: Education Publishing House.
- Cho, Sungdai, and Jonh Whitman. 2020. *Korean: A Linguistic Introduction*. Cambridge: Cambridge University Press.
- Choi, Soonja, and Melissa Bowerman. 1991. Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition* 41(1-3):83-121.
- Choi, Soonja. 2006. Influence of language-specific input on spatial cognition: Categories of containment. *First Language* 26(2):207-232.
- Choi, Soonja. 2009. Typological differences in syntactic expressions of Path and Causation. In *Routes to Language: Studies in Honor of Melissa Bowerman*, ed. Virginia C. Mueller Gathercole:169-194. Mahwah, NJ: Lawrence Erlbaum.
- Duffield, Nigel. 2013. Head-first: on the head-initiality of Vietnamese clauses. In *Linguistics of Vietnamese: An International Survey*, eds. Daniel Hole, and Elisabeth Löbel, pp.127-155. Berlin: De Gruyter Mouton. DOI: 10.1515/9783110289411.127.
- Freed, Barbara (Ed.). 1995. *Second language acquisition in a study abroad context*. Amsterdam, Netherlands: John Benjamins.
- Gu, Lin. 2018. Length of Time Abroad and Language Ability Comparing Means Using a Latent Approach. In *The Routledge Handbook of Study Abroad Research and Practice*, eds. Cristina Sanz, and Alfonso Morales-Front, pp.277-292. New York: Routledge.
- Hakuta, Kenji, Yuko Goto Bulter, and Daria Witt. 2000. *How Long Does It Take English Learners to Attain Proficiency?* (The University of California Linguistic Minority Research Institute Policy Report 2000-1). California University Santa Barbara.
- Harley, Birgit. 1986. *Age in Second Language Acquisition*. Clevedon: Multilingual Matters.
- Harley, Birgit, and Doug Hart. 1997. Language aptitude and second language proficiency in classroom learners of different starting ages. *Studies in Second Language Acquisition* 19(3):379-400.
- Herschensohn, Julia. 2007. *Language Development and Age*. Cambridge: Cambridge University Press.

- Johnson, Jacqueline, and Elissa Newport. 1989. Critical period effects in second language learning: the influence of maturational state on the acquisition of ESL. *Cognitive Psychology* 21(1):60–99.
- Krashen, Stephen. 1973. Lateralization, language learning and the critical period: Some new evidence. *Language Learning* 23:63–74.
- Krashen, Stephen, Michael Long, and Robin Scarcella. 1979. Age, rate and eventual attainment in second language acquisition. *TESOL Quarterly* 13:573–82.
- Nguyen, Lai. 2001. *Nhóm từ chỉ hướng vận động trong tiếng Việt* [Words of motion direction in Vietnamese]. Hà Nội: Khoa học Xã hội Publisher.
- Nguyen, Phuong Hoang. 2019. *The preservation of path verb meanings in prepositional extensions in Vietnamese*. Brisbane: University of Queensland dissertation, Australia.
- Llanes, Àngels. 2011. The many faces of study abroad: An update on the research on L2 gains emerged during a study abroad experience. *International Journal of Multilingualism* 8:189–215.
- Pace, Cassandra. 2009. The typology of motion verbs in Northern Vietnamese. *Rice Working Papers in Linguistics* 1:52–64.
- Patkowski, Mark. 1980. The sensitive period for the acquisition of syntax in a second language. *Journal of Research in Language Studies* 30(2): 449–468.
- Phan, Trang, and Lan Chu. 2024. *A handbook of the Vietnamese Oral Proficiency Interview (OPI) Assessment: According to the ACTFL Framework by the American Council on the Teaching of Foreign Languages* [written bilingually Vietnamese-Korean]. Hanoi: Vietnam National University Publishing House.
- Phan, Trang, and Nigel Duffield. 2022. “A road map to Vietnamese phrase structure.” In *The Routledge Handbook of Asian Linguistics*, eds. Chris Shei, and Saihong Li, pp. 165–185. London & New York: Routledge.
- Phan, Trang, Bianca Basciano, and Lan Chu. 2024. Asymmetry in the acquisition of directed motion constructions in L2 Vietnamese: A comparative study of Chinese and Korean learners. *Annali di Ca' Foscari. Serie Orientale* 60:1–44.
- Singleton, David, and Lisa Ryan. 2004. *Language Acquisition: the age factor*. UK: Multilingual Matters.
- Slobin, Dan. 1996. Two ways to travel: Verbs of motion in English and French, In *Grammatical Constructions: Their Form and Meaning*, eds. Masayoshi Shibatani, and Sandra A. Thompson, pp.195–220. Oxford: Clarendon Press.
- Slobin, Dan. 2004. The many ways to search for a frog: Linguistic typology and the expression of motion events. In *Relating Events in Narrative: Typological and Contextual Perspectives*, eds. Sven Stromqvist, and Ludo Verhoeven, pp.219–257. New York: Psychology Press.
- Slobin, Dan. 2006. What makes manner of motion salient? Explorations in linguistic typology, discourse, and cognition. In *Space in Languages: Linguistic Systems and Cognitive Categories*, eds. Maya Hickmann, and Stéphane Robert, pp.59–81. Amsterdam: John Benjamins.

Time-Related Factors in Korean Learners' Vietnamese Acquisition

- Talmy, Leonakd. 1985. Lexicalization patterns: Semantics structure in lexical forms. *In* Language Typology and Syntactic Description: Grammatical Categories and the Lexicon, ed. T. Shopen, pp.36-149. Cambridge: Cambridge University Press.
- Talmy, Leonakd. 2000. *Towards a Cognitive Semantics: Conceptual Structuring Systems*. Cambridge: MIT Press.
- Wu, Elise Yi. 2022. A brife review of effects of age on second language acquisition: Is younger better. *Journal of language teaching* 2(1):1-5.

Trang Phan

*Department of Asian and North African Studies,
Ca' Foscari University of Venice
Dorsoduro 3246, 30123 Venice
thihuyentrang.phan@unive.it*

Lan Chu

*Department of Vietnamese Language and Culture,
VNU University of Languages & International Studies,
Vietnam National University Hanoi
2 Pham Van Dong, Cau Giay district, Hanoi
lanctp@vnu.edu.vn*

Trang Phan, Lan Chu

在探討時間相關因素對韓國學習者習得越南語能力的影響

Trang Phan

威尼斯卡·弗斯卡里大學

Lan Chu

越南河內國家大學

近年來，外國學習者，尤其是韓國學習者，對學習越南語的興趣不斷增加，這主要是為了有效的溝通和職業發展。本研究調查了 63 名韓國學生學習越南語的影響因素，這些學生包括來自首爾韓國外國語大學的 38 名男性和 25 名女性，以及河內越南國家大學語言與國際研究學院的 Kmove 計劃的學生。研究重點是三個關鍵領域：語言課程的安排、自我評估的語言能力以及關於動詞「去」在定向動作結構中位置的多選測試。這項測試強調了韓語和越南語之間的重要差異，是一個有價值的語言能力指標。此外，時間相關的變數如長期暴露於母語環境（超過 6 個月）、學習語言的時間長度（超過 2 年）和學習越南語的開始年齡（12 歲以下）對語言習得過程至關重要。這些研究結果旨在為韓國學習者和其他越來越多接觸越南文化的人提供有關學習越南語的寶貴見解。

關鍵字：越南語作為第二語言、韓國語、時間相關因素

INTRODUCING AGIBIS (AUTOMATIC GRAMMAR INSTRUCTION BASED ON INPUT SENTENCES): A PILOT STUDY OF A CHATBOT FOR AI-ASSISTED L2 CHINESE LEARNING*

Jiun-Shiung Wu, Hsin-Hui Chen, Hui-Ru Yang
National Chung Cheng University

ABSTRACT

This paper presents an automatic grammar instruction system based on input sentences (AGIBIS). AGIBIS takes a Chinese sentence as an input and outputs a list of grammar points present in the sentence after analyzing the structure of the sentence. A user is allowed to choose one or more grammar points (s)he is interested in and the system generates explanations of the grammar points chosen. AGIBIS is built on Linguistics-Oriented Keyword Interface (LOKI), a natural language understanding engine relying on small data. We present the three features of LOKI, which uses only small data to accomplish natural language understanding. We also present the architecture of AGIBIS, where each intent takes care of analyzing a sentence for a particular grammar point. We, then, demonstrate how AGIBIS works and identify its niche in comparison to studies devoted to applying natural language processing models to sentence structure processing, that is, ungrammatical Chinese sentence detection and correction. Finally, we discuss the applications and implications of AGIBIS on Teaching Chinese as a Second language.

Keywords: Automatic Grammar Instruction, AI-assisted Chinese Learning, AI-assisted Chinese Teaching

* The division of labor of the three authors is as follows. The first author comes up with the topic and takes charge of system development and programming. The second author provides the list of grammar points covered here and is responsible for the part related to Chinese teaching. The third author collects and analyzes appropriate training sentences.

1. INTRODUCTION

It is inevitable for Chinese as a Second Language (hereafter CSL) learners to consult a grammar book to learn about sentence structures. But, scanning a grammar book from the beginning takes a lot of time and cannot provide answers targeting at CSL learner's specific questions. The index section of a grammar book, if any, relies on grammar terms to locate the elaborations of specific grammar points. But, if a learner encounters a sentence whose grammatical structure (s)he is interested to learn more, an index of this type is not helpful unless the learner is capable of analyzing the sentence grammatically (syntactically) and of coming up with the grammar points present in the sentence.

In this paper, we propose an automatic grammar instruction system based on input sentences (for short, AGIBIS), which retrieves grammar points from an input sentence and allows a user to choose from the set of grammar points retrieved to learn more about the chosen ones.

One of the invaluable advantages of this system is that CLS learners can use sentences whose grammatical structure(s) they are curious about to extract illustration of these structure(s). To put it another way, this system can target at learners' questions, specific to which elaborations can be provided.

AGIBIS is built on a natural language understanding engine known as Linguistics-oriented Keyword Interface (LOKI, Wang et al. 2019). LOKI relies on linguistics to achieve the goal of natural language understanding through small data, contra the most popular big data trend, as required in Deep Learning Neural Network, e.g. Chollet (2017), and Transformers, e.g. Vaswani et al. (2017).

This paper is organized as follow. In Section Two, we state that LOKI is a commercial product developed by Droidtown Linguistic Technology and provide a brief introduction to the essences of LOKI and explain why small data works. In Section Three, the architecture of AGIBIS is presented. In Section Four, demonstration of AGIBIS is given. Moreover, the niche of AGIBIS LOKI model is presented, in comparison to the studies on applying NLP models to processing sentence structure, i.e. ungrammatical Chinese sentence detection and correction. In Section Five, we discuss the application and implication of AGIBIS on Teaching

Chinese as a Second/Foreign Language. Finally, Section Six concludes this paper.

2. ESSENCES OF LOKI

LOKI is a commercial product developed by Droidtown Linguistic Technology Ltd. It achieves natural language understanding based on small data because of its three features: (i) syntax, including syntactic categories (a.k.a. parts of speech, word classes, etc.) and phrase structure rules (PS rules), (ii) directed graphs, which formalize PS rules, and (iii) pattern matching. Let's start with the first one.

LOKI, as its name suggests, is linguistics-oriented. It relies on the ideas of syntactic categories and PS rules. How is it determined that a group of words belong to the same syntactic category? As stated in Aarts (2008:26-48), Radford (1988:60-64), and so on, distributional criteria are relied on to determine whether a group of words are of the same syntactic category. That is, words of the same syntactic category appear in the same position in a sentence.

PS rules are context-free rules which can generate well-formed (i.e. grammatical) sentences in a language. Let's look at one example:

- (1) a. NP $\rightarrow$ (Det) (CL) (A)* N
 b. *gǒu* 'dog', *xiōngměng-de gǒu* 'fierce dog',¹ *xiōngměng měilì de gǒu* 'beautiful, fierce dog', *nà zhī xiōngměng měilì de gǒu* 'that classifier beautiful, fierce dog', etc.

(1a) basically says that a nominal phrase (NP) consists an optional determiner (Det), an optional classifier (CL), any number of adjectives (A)

¹ There is a disagreement concerning whether *adjective* is a valid syntactic category in Chinese, e.g. Chen and Huang (1993), who propose that words which are usually referred to as adjectives in Chinese and which can function as predicates are actually stative intransitive verbs, vs. Liu (2007, 2010), who agrees that *adjective* is a valid syntactic category in Chinese. While it does not really affect how LOKI works, we take Liu's (2007, 2010) position for now. Please note that Droidtown's inventory of syntactic categories does not include *adjective*, but *modifier*. See <https://api.droidtown.co/document/#Pos>.

and an obligatory head noun. In the notation of PS rules, parentheses indicate optionality and Kleene star stands for repetition of any number greater or equal to zero. This phrase structure alone can generate an infinite number of nominal phrases, partially exemplified by the NPs in (1b).

That is, once a string is segmented and paired with syntactic categories, the (phrasal) structure of the string can be abstracted over to get a PS rule, which, in turns, can generate strings of the same phrase structure by means of substituting the words in the string with words of the same syntactic category.

In order to get a feeling how LOKI really uses the idea of syntactic category and PS rules, let's use LOKI's syntactic categories below for the purpose of illustration:

- (2) a. Zhāngsān kàn-le yì-běn shū.
Zhangsan read-PFV² one-CL book
'Zhangsan read a book.'
- b. <ENTITY_person>張三</ENTITY_person>
<VerbP>看了</VerbP>
<ENTITY_classifier>一本
</ENTITY_classifier>
<ENTITY_noun>書</ENTITY_noun>
- c. S → ENTITY_person VerbP (ENTITY_classifier)
ENTITY_noun

The sentence in (2a) is segmented and paired with syntactic categories in LOKI's convention, which is shown in (2b). The syntactic category for a word is marked with XML conventions. <ENTITY_person> and </ENTITY_person> surround a word which is an entity and a person. *Kàn-le* 'read-PFV' is treated as a verbal phrase (VerbP).³ A numeral and a classifier together are treated as ENTITY_classifier. And finally, *shū* 'book' is classified as ENTITY_noun.

² The abbreviations used in this paper include: CL for a classifier, PFV for a perfective aspect marker.

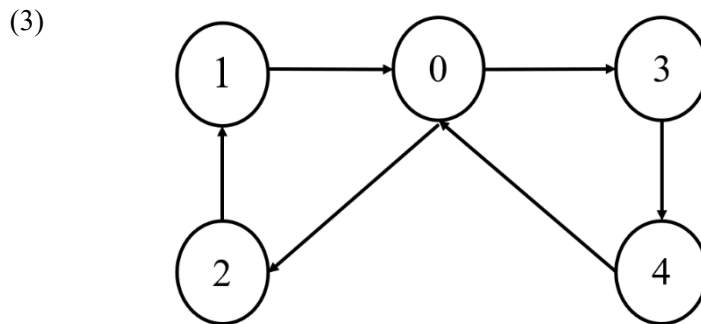
³ This VerbP is different from VP used in syntax. VerbP in LOKI basically is a constituent composed of a verb plus an aspect marker.

Moreover, the segmentation and tagging of (2a) is abstracted over to form a PS rule as (2c). Essentially, (2c) says that a sentence is composed of a person, a verbal phrase, an optional classifier and a noun. Please note that, due to implementation limitation, the PS rules used in LOKI do not necessarily conform to linguistic theories. However, LOKI mostly captures the spirit of PS rules.

The PS rule in (2c) can generate an infinite number of sentences by replacing the words in this sentence with words of the same syntactic category. For example, *Zhāngsān* can be replaced by any names such as *Lǐsì*, *Xiǎohuá*, etc. The verbal phrase *kàn-le* ‘read-PFV’ can be replaced by any (transitive) verbal phrase. The classifier *yì-běn* ‘one-CL’ is optional and therefore can be omitted or substituted by any classifier. And, the noun *shū* ‘book’ can be replaced by any noun. This way, an infinite number of sentences can be generated and recognized in the task of natural language understanding.

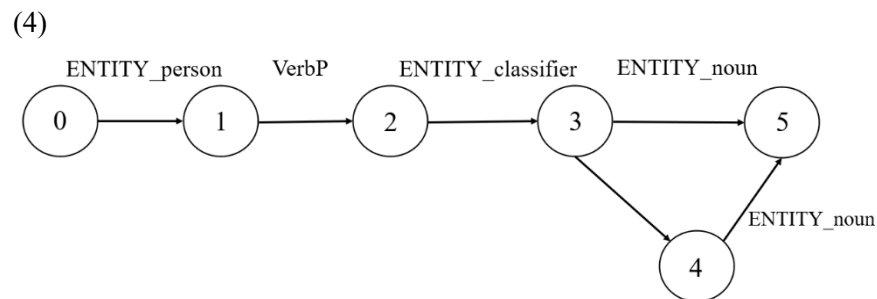
With the distributional definition of syntactic categories and the idea of PS rules, which can generate an infinite number of sentences, we need a way to formalize the idea. Directed Graphs (DG) have been an option to formalize syntactic structure. For example, Matthews (1998, Chapters 8-10) demonstrates how to use finite state automata (a type of DG) to formalize sentence structure and how to implement these finite state automata.

A DG is a graph composed of states and directed paths among states. For example,



(3) is a directed graph, with five states numbered from 0 to 4 and directed paths between these states. Imagine that (3) is a map and you are on this map. Starting from state 1, you can only walk to state 0, but not to states 2, 3, or 4. You cannot walk to state 2 because the path between 1 and 2 points toward 1, not the other way around. You cannot walk to 3 or 4 because there is no path connecting 1 to either 3 or 4.

DG can be used to model syntactic structure. Let's take (3c) as an example. (3c) can be modeled by the DG (4):



(4) works in the same way as (3). From state 0 to state 1, a word of ENTITY_person must be read. From state 1 to state 2, a string of VerbP, composed of a verb and an aspectual marker, must be read. State 2 has two paths from it: one toward state 3, reading ENTITY_classifier and the other to state 4, reading nothing. The two paths from state 2 model optionality of ENTITY_classifier: if one goes from state 2 to state 3, then ENTITY_classifier must be read. If one walks from state 2 to state 4, nothing is read. Since there are two paths from state 2 and one of the paths does not read anything, ENTITY_classifier does not always have to appear in the sentence and therefore is optional. Both states 3 and 4 go to state 5, reading ENTITY_noun.

The reason why DG is used to model syntactic structure is because DG can be easily implemented computationally and therefore it is a very popular model used to represent sentence structures.⁴

⁴ Please note that DG might not have the same expressive power as natural language syntax. To put it another way, whether DG can completely model natural language syntax is an issue under debate. Interested readers are referred to Pullum and Gazdar (1982) for an enlightening discussion. While the answer to this question is commonly accepted as

Finally, LOKI relies on pattern matching to achieve the purpose of recognizing an infinite number of sentences based on small set of training data. Let's take (2a) as an example again. How can we recognize *nǐ zhīdào Zhāngsān kàn-le yì-běn shū* 'you knew that Zhangsan read a book', *Xiǎomíng tīngshuō nǐ zhīdào Zhāngsān kàn-le yì-běn shū* 'Xiaoming heard that you knew that Zhangsan read a book', *wǒ yǐwéi Xiǎomíng tīngshuō nǐ zhīdào Zhāngsān kàn-le yì-běn shū* 'I thought Xiaoming heard that you knew that Zhangsan read a book', etc. by only one training sentence *Zhāngsān kàn-le yì-běn shū* 'Zhangsan read a book'?

The answer lies in pattern matching. Let's examine the segmentation and tagging of *wǒ yǐwéi Xiǎomíng tīngshuō nǐ zhīdào Zhāngsān kàn-le yì-běn shū* 'I thought Xiaoming heard that you knew that Zhangsan read a book' and compare it with (2b), repeated below (5a) for the sake of explanation.

- (5) a. <ENTITY_person>張三</ENTITY_person>
 <VerbP>看了</VerbP>
 <ENTITY_classifier>一本
 </ENTITY_classifier>
 <ENTITY_noun>書</ENTITY_noun>
 b. <ENTITY_pronoun>我</ENTITY_pronoun>
 <ACTION_verb>以為</ACTION_verb>
 <ENTITY_nouny>小明</ENTITY_nouny>
 <ACTION_verb>聽說</ACTION_verb>
 <ENTITY_pronoun>你</ENTITY_pronoun>
 <ACTION_verb>知道</ACTION_verb>
 <ENTITY_person>張三</ENTITY_person>
 <VerbP>看了</VerbP>
 <ENTITY_classifier>一本
 </ENTITY_classifier>
 <ENTITY_noun>書</ENTITY_noun>

In (5), the part overlapping (5a) is marked with different fonts. As we can see, the segmentation and tagging of the longer sentence under

negative, in terms of natural language processing, DG is still frequently used to model language structure, because its easy implementability and wide coverage.

discussion contains that of (5a). Therefore, the way to recognize a longer sentence with a short training one is to examine whether the segmentation and tagging of the training sentence can be found in the longer one.

One might ask why we need to recognize a longer sentence based a shorter training one. Suppose that the system outputs a list of grammar points for a user to choose from: *wánchéng mào* ‘perfective aspect’, *fēnlèi cí* ‘classifier’ and *zhíshùjù* ‘declarative sentence’. The user can responses *wánchéng mào* ‘perfective aspect’, *wǒ xiǎng zhīdào wánchéng mào* ‘I want to know (learn) about perfective aspect’, *wǒ xiǎng duō zhīdào yìxiē wánchéng mào* ‘I want to know (learn) more about perfective aspect’, etc.

Given that the user can respond with a variety of ways, it is desirable to recognize as many potential responses as possible without using the same number of training sentences. The way of pattern matching adopted by LOKI as introduced above can easily achieve the purpose: all LOKI needs is one training string: *wánchéng mào* ‘perfective aspect’. The other longer potential responses can be captured as long as the pattern of *wánchéng mào* ‘perfective aspect’ can be found included in them.

While LOKI is capable of capturing many sentences with a small set of training sentences, dependent on linguistic research, it has its own weaknesses. Here, we mention one. Given that LOKI is syntactically based, it has difficulties dealing with semantics. For example, LOKI cannot take care of selectional restrictions. It takes extra efforts for LOKI to deal with semantics. But, we will not go into details here.

An interim summary. AGIBIS is built on LOKI, which is a natural language understanding engine based on linguistics. To achieve the purpose of natural language understanding with small data, LOKI relies on three mechanisms: (i) the distributional definition of syntactic categories and PS rules, (ii) directed graph, and (iii) pattern matching.

3. ARCHITECTURE OF AGIBIS

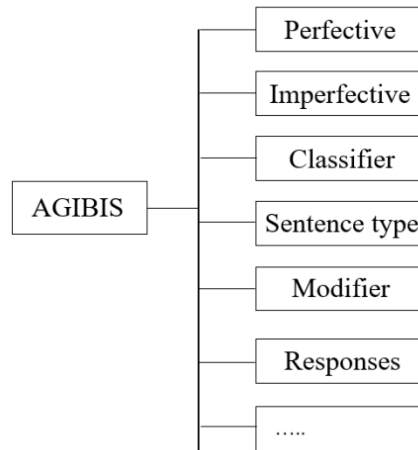
A LOKI project is composed of different intents, each of which performs a unified function and serves a consistent purpose. AGIBIS consists of multiple intents, each of which scans an input sentence and retrieves the corresponding grammar point from the input sentence, if the

grammar point *is* present in the sentence, and nothing is retrieved if the grammar point is not present in the sentence.

The current version reported in this paper contains six intents, five of which deal with grammar points and the other of which takes care of responses and does not analyze sentence structure. The five intents taking care of grammar points in this version of AGIBIS are: Perfective, Imperfective, Classifier, Sentence Type, and Modifier.

The architecture is as follows:

(6) Intents of AGIBIS LOKI Model



As stated above, each intent takes care of a particular grammar point. Let's elaborate on the intents specified in the architecture (6). Perfective scans through a sentence and identifies whether the sentence is either perfective or experiential. Imperfective identifies an input sentence either as durative or progressive. Classifier locates the classifier used in an input sentence, if any. Sentence type determines the sentence type of an input sentence, i.e. declarative or different sub-types of interrogatives, e.g. WhatQ, HowQ, WhereQ, etc. To make it clear, this paper presents a pilot study and therefore this version of AGIBIS deals with declarative and *what* questions only. Modifier takes care of nominal modifiers, such as possessive/part-whole relationship, e.g. *Zhāngsān-de shū* 'Zhangsan's

book’ or *gōngchē-de yǐnqíng* ‘the engine of a bus’, relative clauses, such as *zhàn zài ménkǒu de nǚhái* ‘the girl who is/was standing at the door’, *wú jiàoshòu xiě de lùnwén* ‘the paper which Professor Wu write’, or locative/temporal modifier, such as *zhuōzi shàng de bǐ* ‘the pen on the table’ or *sāntiān qián de shìguò* ‘the accident three days ago’.

Responses is the only intent which does not deal with retrieving grammar points from an input sentence. This intent is required for a technical reason. AGIBIS is designed to scan through a sentence and to output a list of grammar points present in the sentence. Then, a user can choose from the list and AGIBIS generates explanations of the grammar point(s) chosen. The intent *Responses* scans through user’s input and determines whether the input is a response revealing what grammar points the user is interested in, and, if yes, what grammar point(s) should be elaborated on.

Please note that, once a sentence is fed into AGIBIS, the sentence is scanned by every intent. An intent outputs the corresponding grammar point, if the grammar point is detected in the sentence. On the other hand, an intent outputs nothing if the corresponding grammar point is not detected in the input sentence.

Let’s use *Xiǎolǐ yǎng-le sān-zhī lùguī* ‘Xiaoli keeps three tortoises (as pets)’ as an example to elaborate on the architecture (6). As stated before, when an input sentence is fed into the AGIBIS model, it goes through every intent so that all of the significant grammar points present in the sentence can be retrieved. When the sentence goes into the intent *Perfective*, it is identified as a perfective sentence with *le*, after being analyzed by the intent. Intents *Imperfective* and *Modifiers* do not detect any grammar point of these two types in the input sentence and therefore do not output anything. When the sentence goes into the intent *Classifier*, the intent locates *zhī* ‘CL’ as the classifier in this sentence and marks this sentence as containing a classifier, which is *zhī*. The intent *Sentence type* identifies this sentence as declarative. When the sentence is scanned by all of the intents, AGIBIS outputs a list of grammar points retrieved from the input sentence: perfective *le*, classifier *zhī* and declarative.

After outputting the list of grammar points, AGIBIS allows a user to choose one or more grammar points present in the sentence in which (s)he

is interested. Then, AGIBIS generates explanations of the grammar points chosen.

4. DEMONSTRATION OF AGIBIS AND ITS NICHE

In this section, we, first, present a demonstration of AGIBIS. AGIBIS is a chatbot, which is currently deployed on Discord, a free instant messaging and digital publication social platform. Then, we point out the niche of AGIBIS, compared to the studies on applying NLP models to sentence structure processing, specifically, ungrammatical Chinese sentence detection and correction.

4.1 Demonstration

Once it is activated, one can talk to AGIBIS and ask it to conduct automatic grammar instruction based on the sentences a user inputs. Let's see some examples where AGIBIS is put to work. The first input sentence we wish AGIBIS to process is: *zhuō shàng de xiǎoshuō shōuqǐ lái le ma* 'has the novel on the desk been put away?'

(7)

The screenshot shows a Discord chat window with the following messages and annotations:

- HunterWu** (2023/10/03 07:59): @GrammarInstructor hi
- @GrammarInstructor** (2023/10/03 07:59): hi
- GrammarInstructor** (2023/10/03 07:59): Hi! 輸入一個句子，我們來學習這個句子的語法吧！
- HunterWu** (2023/10/03 07:59): @GrammarInstructor 桌上的小說收起來了嗎？
- @GrammarInstructor** (2023/10/03 07:59): 桌上的小說收起來了嗎？
- GrammarInstructor** (2023/10/03 07:59): 你的輸入的句子有修飾語，疑問句，這幾個語法點，你想知道哪個語法點的說明？
- HunterWu** (2023/10/03 08:00): @GrammarInstructor 我想知道修飾語，疑問句
- @GrammarInstructor** (2023/10/03 08:00): Your sentence has a modifier that expresses a time or location. You usually will find a word expressing relative position after the time/location noun. This is a yes-no question. In Chinese such a question usually ends with 'ma'. You can answer a yes-no question with 'shi, bushi, dui, budui', etc.

Annotations on the right side of the chat window:

- Greetings**: Points to the initial 'hi' messages.
- Input sentence**: Points to the user's question '桌上的小說收起來了嗎？'.
- Retrieving list of grammar points**: Points to the bot's response listing grammar points.
- User's responses**: Points to the user's follow-up question '我想知道修飾語，疑問句'.
- Output explanations**: Points to the bot's detailed explanation of the grammar points.

As shown in (7), we greet with the chatbot GrammarInstructor first. Then, we can input a sentence. The input sentence is fed into AGIBIS LOKI model. AGIBIS analyzes the sentence and outputs a list of grammar points for the user to choose. The user inputs the grammar point(s) he is interested in. Please note that here the user input a complete sentence *wǒ xiǎng zhīdào xiūshìyǔ, yíwènjù* ‘I want to know about modifier and interrogative’, as marked by the first red underline. Then, AGIBIS analyzes the input again. The chatbot decides that this is a response from the user and retrieves what the user is interested in. Finally, AGIBIS outputs explanations in English. Please pay attention to the second and third red underlined parts. AGIBIS identifies the modifier present in this sentence is one that expresses locative/temporal modification, and the sentence type as a subtype of interrogative, i.e. a yes-no question.

Let’s look at another example. The sentence to be processed is: *zài kān shū de tóngxué biǎoxiàn hén hǎo* ‘the classmate who is reading a book performs well’.

(8)



The workflow of GrammarInstructor is the same as the first demonstration. What deserves the attention in (8) is underlined in red. First, in (8), the user's response is a short one, without using a complete sentence. As explained in Section 2, LOKI uses only one training string to capture both a long response as a complete sentence as the one in (7) or a short one like this here, with pattern matching. Second, AGIBIS correctly identifies the subtype of imperfective of this input sentence: progressive *zài/zhèngzài*. Third, it, again, accurately identifies that the modifier in this sentence is a relative clause.

Let's see the last demonstration. The sentence to process is: *wǒ kàn-guò sān-zhī xióng* 'I have seen three bears (before)'.

(9)



(9) is just like (7) and (8). Please pay attention to the underlined parts in red. As we can see, AGIBIS can retrieve the actual classifier used in a sentence, *zhī* in this example, and offer explanation specific to this classifier. Furthermore, it can pinpoint the subtype of perfective, i.e. experiential *guò* in this example.

4.2 Niche of AGIBIS

Ungrammatical sentence detection and/or correction is one of the popular issues in NLP, for example Yeh et al. (2016), Lee et al. (2017), Ren et al. (2018), Xiang (2018), Zhang and Wang (2019), Li et al. (2019), Zhao and Wang (2020), Wang et al. (2020), Cheng and Duan (2020), Lee et al. (2021), Chen and Zhang (2022), Kuang et al. (2022), etc.

These works either treat ungrammatical sentence detection as a classification task or think of ungrammatical sentence correction as a translation task. The former studies use Deep Learning or Transformers Models to predict whether a sentence is ungrammatical or not, while the latter ones depend on these models to “translate” an ungrammatical sentence into a grammatical one.

While these studies rely on different models, they share one problem: CSL learners do not know why a certain sentence is ungrammatical or should be corrected in the specified way. It would be of great help if the learners can know more about the grammatical structure of a corrected sentence.

To achieve the purpose stated above, AGIBIS can be used, following these models. Once a sentence is classified as ungrammatical and is translated into a grammatical one, the grammatical sentence can be fed into AGIBIS to get a list of grammar points from the sentence for a Chinese learner to learn more about the structure of the sentence. In this way, the learner can get a better idea of how a grammatical sentence should be and acquire more about the key grammar points of the corrected sentence.

On the other hand, Chung (2023) proposes a model which can conduct ungrammatical sentence detection, correction and provide revision suggestions. While Chung’s model provides explanations, these explanations are fixed and attached to the error types, instead of automatically retrieving from input sentences by analyzing the input sentences. That is, Chung offers explanations specific to errors. Yet, her model cannot provide a complete analysis of the grammatical structure of a sentence and does not allow a learner to know more about the grammar points. AGIBIS can be of help in this aspect.

In sum, AGIBIS can be complementary to Deep Learning/Transformers-based ungrammatical Chinese sentence detection/correction or Chung's LOKI-based model to provide explanations on grammar points retrieved from corrected sentences. The most significant niche of AGIBIS is that the model provides explanation of grammar points based on input sentences. When encountering a sentence whose grammar points a user is interested to know more, AGIBIS can be of tremendous help, which grammar books or the studies on ungrammatical Chinese sentence detection and/or correction can hardly provide.

5. APPLICATION, IMPLICATIONS AND FUTURE DEVELOPMENT

In this section, we discuss the application and implications of AGIBIS to Chinese teaching. We present a few cases where AGIBIS will come in handy for Chinese learners and hope that this presentation can inspire more applications of AGIBIS. We also discuss the implications of AGIBIS to Chinese teaching. And, finally, we point out some possible future developments for AGIBIS to make it a more full-fledged tool for individualized, specialized, targeted and autonomous L2 learning.

5.1 Application of AGIBIS to Chinese Teaching

The role of linguistic knowledge in successful L2 learning has undergone dramatic change in the past decades, e.g. DeKeyser (2003), DeKeyser and Juffs (2005), Ellis (2006), etc. Some scholars suggest that explicit structural knowledge of a target language is of great help to language learners' production, once the knowledge is internalized and automatized, e.g. DeKeyser (1998), Johnson (1996), etc., whereas others do not believe in the contribution of explicit linguistic knowledge in the success of L2 learning, such as Ellis (2004), Krashen (1987), Paradis (1994), etc.

For example, Krashen's Monitor Theory stipulates that explicit linguistic knowledge hinders fluent production of a target language

because an L2 learner will monitor his/her production to see if his/her production is “grammatical”, that is, whether his/her utterance conforms to the rules.

In a more recent study, Hu (2011:74) concludes that “[...] metalinguistic awareness has an important role to play in L2 learning and use”, and also points out that “[...] the need for further research that explores how metalinguistic and metalingual knowledge can be fruitfully integrated into meaning-focused, communication-oriented L2 instruction.”

The purpose of grammar instruction is *not* to train L2 learners to do grammatical analysis. Instead, grammar provides rules, which tell L2 learners what kind of word orders and/or structure is “acceptable” or commonly used by native speakers. Let’s take an extreme example. If a beginning learner of Chinese does not know that the Chinese word order is normally Subject-Verb-Object, then (s)he may have trouble producing Chinese sentences whose word order is commonly accepted by native speakers. But, once they are introduced this word order of Chinese, they can easily grasp this word order and can generate an infinite number of “grammatical” sentences. That is, grammar instruction is to help L2 learners to generate grammatical ($\approx$ acceptable) sentences, not to equip L2 learners with the ability to do structural analysis. However, we would like to point out that, grammatical awareness, i.e. conscious knowledge of structure, plays different, but significant, roles for L2 learners of different levels of proficiency. In order to achieve an advanced, near-native level, it is crucial to develop an awareness of structure so that L2 learners can express themselves correctly and naturally.

That said, AGIBIS can be of great value to Chinese learners to facilitate their knowledge of grammar point with the aid of LOKI because the system can provide specialized, individualized and targeted instruction of grammar points. It is specialized because it produces grammar points based on sentences a learner is interested in, unlike grammar books, which present according to a pre-arranged order of grammar points. It is individualized because each learner can input sentences (s)he is interested in and every learner likely take interest in different sets of sentences. It is targeted because it targets at learners’ needs.

AGIBIS can be invaluable to activities which require students to know about the grammar points in them. We present four tasks that take

excellent advantage of AGIBIS: picture description task, sentence grammaticality distinction task, sentence unscrambling task, and sentence creation by analogy task.

A picture description task is a natural way to examine whether students can produce grammatical (acceptable) sentences. For example, after students learn a set of grammar points, e.g. declarative vs. question, perfective *le*, progressive *zài* and a few classifiers, the teacher can show (motion or static) pictures, which require to use some of the set of grammar points to describe, and ask students to write down, for each picture, a sentence, which can best depict the picture. After the students write down their sentences for the pictures, the teacher provides an answer key. The students check their sentences against the answer key. For those sentences they make mistakes with, students can input the sentences in the answer key into AGIBIS, which generates a set of grammar point based on the input sentences. Then students can choose the grammar point(s) they want to learn more for AGIBIS to yield explanations for the grammar point(s) chosen.

AGIBIS is of great help in this task for the following reasons. First, students will less likely make mistakes with the same sentences. If (s)he is to discuss students' mistakes, the teacher will need to talk about more mistakes than students need individually. AGIBIS is specialized and individualized, so each student only needs to input the sentence(s) puzzling him/her. Second, it is very difficult, if possible at all, for a learner to look up, in a grammar book, the grammar points of a sentence with which (s)he makes mistakes, but this is exactly what AGIBIS is intended to achieve. Third, AGIBIS is targeted and hence learners will get explanations of the grammar points they wish to explore and do not need to spend time on other students' mistakes which they do not make.

A sentence grammaticality distinction task is to help students spot an ungrammatical (unacceptable) sentence. If an L2 learner can do this (at least partially), it will be very useful because the learner can always look back at the sentences (s)he writes and performs a preliminary check.

For this task, the teacher prepares a few sets of sentences, some of which are grammatical but the others of which are not. Students form several groups and each group gets a set of sentences. The students of each group look at the sentences and discuss which is grammatical and which

is not. When the given period of time is up, the teacher gives each group the answer key, which specifies which sentence is grammatical and which is not. For ungrammatical sentences, the teacher provides correct ones.

Again, just like the picture description task, AGIBIS can contribute a great deal in this sentence grammaticality distinction task. As stated above, AGIBIS is specialized, individualized and targeted. Therefore, students can use AGIBIS to explore the grammar points of the sentences in which they make mistakes and find out what they may be missing so that they make the mistakes.

A sentence unscrambling task is to ask students to restore a set of strings of randomized order back to grammatical sentences. It goes as follows. The teacher gives students set of randomized strings of randomized order. Students perform group discussion and decide how to re-order the strings assigned to each group so that grammatical sentences can be derived. After students finish re-ordering the strings given to them, the teacher provides an answer key for students to check their results against. Then, students can get individualized and specialized explanations of grammar points of the sentences with which they make mistakes by inputting these sentences into AGIBIS to retrieve the grammar points of each sentence, from which students can choose to get AGIBIS to generate explanations targeted at their needs.

The last task to present here is the sentence creation by analogy task. Gentner (2003:203) points out that highlighting and schema abstraction is one of the methods in which the attainment of knowledge can be facilitated. Under the scenario of L2 learning, creating new sentences by analogy is one of the important ways to achieve schema (= grammar here) abstraction. The foundation of the analogy task relies on the principles of Structure-Mapping Theory. The Structure-Mapping Theory posits that the comparison process involves aligning and mapping structured conceptual representations (Gentner 2003; Gentner and Namy 2006; Gentner and Markman 1997). The fundamental concept in the structure-mapping theory is that an analogy involves transferring knowledge from one domain (referred to as the base) to another (known as the target). A sentence creation by analogy task is to give students a list of sentences and to ask students to create new sentences by analogy. Students look at each sentence given and create a new sentence by means of analogy to each

given sentence. Then, students are asked to analyze the grammar points of the sentences they create. By means of creating new sentences in analogy to a given sentence, L2 learners can come up with the abstract structure, which governs the generation of these sentences. This task helps to reinforce their understanding of Chinese grammar and to apply their grammatical knowledge to practical language expression.

Obviously, AGIBIS is helpful in this task because students can use AGIBIS to check whether their analysis is on the right track and, if not, AGIBIS can provide the explanation of the grammar point they have difficulties with, so that they can learn autonomously and get targeted responses. Analogy task can help students improve their Chinese language skills, grammar understanding, and creative thinking. This task also encourages students to actively participate and transform their language learning into practical language use.

5.2 Implications of AGIBIS to Chinese Teaching

In this section, we present three important implications of AGIBIS on Chinese Teaching. First, we discuss the advantage of the linguistics-based explanations AGIBIS provides. Second, AGIBIS can help to deal with global errors (Selinker 1969). Third, AGIBIS complements the inadequacies of existing digital language learning tools, such as Lingvist, Duolingo, etc.

AGIBIS is a linguistics-based AI-assisted Chinese teaching system. It is linguistics-based, not only because it is built on the natural language understanding engine LOKI, but also because the grammar points encoded in AGIBIS is based on Chinese linguistics. For example, Chinese is a language without tense morphology, e.g. Li and Thompson (1981), and hence it is not appropriate to talk about present tense and past tense when teaching Chinese. Moreover, perfective *le* does not indicate past tense. AGIBIS generates explanations of grammar points based on Chinese linguistics. Therefore, AGIBIS will not provide tense as a grammar point. Rather, it presents Perfective and Imperfective as grammar points. This way, students can get an accurate characterization of Chinese grammatical structure and can avoid misunderstanding and misinterpretation once Chinese and the native languages of students have subtle differences.

Second, Selinker (1969) highlights that the errors serve as valuable indicators of a learner's language learning progress and play a crucial role in testing L2 learners' hypothesis of target language. Additionally, Burt and Kiparsky (1978) distinguishes global error from local error and suggests that error plays a significant role in L2 acquisition because error provides insights into language acquisition and helps students and teachers evaluate the progress of L2 acquisition. Global error refers to structural error, while local error to lexical error. AGIBIS can be used to assist students to learn, autonomously, about potential global errors, in ways similar to the four tasks discussed in Section 5.1.

Third, existing digital language learning tools such as Lingvist, Dulingo, etc. all focus on vocabulary learning. Typically, these tools encourage students to learn vocabulary, phrases, and simple sentences by using flashcards, visuals, listening exercises, and writing activities. Moreover, ChatGPT, which has been used increasingly as a language learning tool, provides a very shallow, and sometimes incorrect and/or insufficient, syntactic analysis. Let's take a look at an example:

(10) ChatGPT's response while asked to provide *wǒ kàn-guò sān-zhī xiōng* 'I have seen three bears' (retrieved October 30th, 2023)

a.

找出以下句子的語法點：我看過三隻熊

這個句子的語法點如下：

1. 主語: "我" - 這是句子的主要動作執行者，也就是主詞。

2. 動詞: "看過" - 這是動詞短語，表示動作。"看過"組成了動詞"看"和助動詞"過"，表示過去的動作。

3. 受詞: "三隻熊" - 這是動詞的受詞，描述了動詞的操作對象。在這個句子中，"三隻熊"是你看過的東西。

所以，這個句子的基本結構是"主語 + 動詞 + 受詞"，描述了一個過去的動作，即你看過三隻熊。

b. English translation:

Find the grammar points of the following sentence:

wǒ kàn-guò sān-zhī xióng ‘I see-Exp three-cl bear’

The grammar points of this sentence are as follows:

1. Subject: wǒ ‘I’ – the doer of the main action of this sentence, that is, the subject
2. Verb: kàn-guò ‘see-Exp’: this is a verbal phrase, describing an action. It is composed of a verb kàn ‘see’ and an auxiliary verb guò ‘Exp’, indicating a past event.
3. Object: sān-zhī xióng ‘three-cl bear’— this is the object of the verb, describing the entity acted upon. In this sentence, this phrase is what you saw.

So, the basic structure of this sentence is Subject + Verb + Object. It describes a past event.

As we can see from ChatGPT’s response, essentially, only the subject-verb-object of the given sentence is analyzed, unlike what we demonstrate previously where AGIBIS provides a list of important grammar points: sentence type, experiential aspect and classifiers, which are more helpful to Chinese learners than what ChatGPT provides. Moreover, ChatGPT fails to identify guò as an experiential marker and to mistakenly associate the experiential marker with past tense, which is a significant disadvantage since it is widely accepted that perfective markers, i.e. perfective le and experiential guò, do not express past. And, it does not report the sentence type of the given sentence.

This is probably because it is relatively easy to conduct a shallow structural analysis. It requires much more knowledge to perform a complete and detailed structural analysis and takes a lot more efforts to train an AI model to conduct structural analyses, for example, to retrieve grammar points based on input sentences, as what AGIBIS is capable of. Moreover, it goes without saying that students will need to go beyond lexical level and learn structure-level knowledge of a target language. AGIBIS can be handy because it provides structural knowledge.

Furthermore, it has been recognized that the use of appropriate teaching strategies and error correction feedback is crucial, e.g. Chang (1992), Zhang (2020), etc. While AGIBIS does not correct ungrammatical sentences, it can provide information about the correct sentence once an ungrammatical sentence is corrected. Students will not feel threatened by AGIBIS, which is an AI model and is not judgmental. That is, there is a niche in Computer-assisted Language Learning (CALL) for AGIBIS, which is capable of structural analysis and of providing explanations targeted at students' individual needs.

To sum up, we discuss three implications of AGIBIS on Chinese teaching. It provides explanations based on Chinese linguistics and avoid confusing, misleading statements. It helps to take care of global errors. And, finally, there is a niche for AGIBIS because few, if any, digital language learning tools can deal with structure.

5.3 Future Development

Kim and Kim (2020) point out that AI tutoring is a recent trend in educational technology because personalized educational service can be provided. As pointed out above, AGIBIS, as it is the current way, is specialized, individualized and targeted, and, therefore, fits in the AI tutoring trend in educational technology. In order to make AGIBIS fulfill its expectation in assisting structural teaching and learning of Chinese, there are a few points where AGIBIS needs more work.

First, more grammar points should be taken into consideration so that AGIBIS can perform much better than it can now. It is ideal to consult pedagogical grammar books in Chinese to expand the grammar points AGIBIS can recognize to generate targeted explanations. The wider the coverage of AGIBIS is, the more useful it can be.

Second, if students allow it, AGIBIS can keep a record of each student who uses this tool. A record of this type is valuable to students themselves and to teachers. This record is useful to students themselves because they can always look back and review what they may have problems with. Students can use these learning records to track their progress and effectively manage their learning priorities. Teachers can also provide tailored advice and guidance based on students' records,

allowing them to create customized learning plans for maximum teaching effectiveness.

Such a record provides useful information for teachers as well. Teachers can analyze such record to categorize students' questions. If associated with students' information, e.g. native languages, length of learning, etc., teachers get to understand the interaction between students' traits and potential questions. With this knowledge, teacher can provide more targeted instruction.

In addition, such a record can also reveal the frequency of each grammar points being inquired about. The information of such frequency can be potentially helpful for teachers to arrange his/her teaching materials or arrange chapters, compiling a textbook, etc. The frequency information can serve as a basis for determining which grammar points should receive more detailed explanations and should be focused on for extensive practice.

6. CONCLUSION

In this paper, we present a system, which performs automatic grammar instruction based on input sentences (known as AGIBIS). This system accepts a Chinese sentence as an input and generates a list of grammar points of this sentence for a learner to choose. After the learner chooses the grammar point(s), in which (s)he is interested, the system outputs explanations of these chosen grammar points.

AGIBIS is built on LOKI (Linguistics-Oriented Keyword Indexing), which performs natural language understanding (NLU) based on a small amount of data. Three features enable LOKI to achieve NLP with small data: (i) linguistics, specifically distributional definitions of parts of speech and phrase structure rules, (ii) directed graphs, and (iii) pattern matching.

The current version of AGIBIS is composed of six intents, e.g. Perfective, Imperfective, Modifier, Sentence Type, Classifier and Responses, five of which are responsible for identifying one grammar point by analyzing an input sentence. The intent Responses reads a user's response to decipher what intent(s) the user is interested in. It functions in

the following steps: a CSL learner feeds, into AGIBIS, a sentence (s)he is interested in, and AGIBIS outputs a list of grammar points. The learner picks the grammar point(s) which (s)he wishes to learn more and AGIBIS generates the explanations of the chosen grammar point(s).

We propose that AGIBIS is autonomous, specialized, individualized, and targeted. It is autonomous because a CSL learner can learn autonomously whenever (s)he feels like learning, without having to wait for teachers' help. It is specialized with respect to input sentences. One can hardly find the grammar point(s) of a sentence by looking the sentence up in a grammar book. It is individualized because each learner can learn individually without interfering with each other. Finally, it is targeted because it provides explanations of grammar points targeted at a specific sentence.

Then, the applications and implications of AGIBIS to Chinese teaching are presented. The future development is also covered. A few cases of exercises or class activities are presented, where AGIBIS can be helpful. Three implications are given. First, AGIBIS provides sound explanations of grammar points, based on Chinese linguistics and avoid confusing, misleading statements. Second, it helps to take care of global errors. Third, there is a niche for AGIBIS because few, if any, digital language learning tool can deal with structure. Concerning future developments, two types are identified. First, AGIBIS should be expanded so that it can take care of more grammar points and becomes a full-fledged grammar instruction tool. Second, if students allow it, AGIBIS is suggested to keep a record of each student's inquiries. Such a record can be valuable to students and to teachers. A record of this type is helpful to students because the record can track their progress and manage their learning priorities. Such a record is helpful to teachers because teachers can analyze students' questions and provide instruction targeting at students' needs.

REFERENCE

- Aarts, Bas. 2008. *English Syntax and Argumentation (3rd ed.)*. New York: Palgrave MacMillan.
- Burt, Marina, and Carol Kiparsky 1978. Global and local mistakes. In *New frontiers in second language learning*, eds. John Schumann, and Nancy J. Stenson. Massachusetts: Newbury House.
- Chang, Hsing-Wu. 1992. The Acquisition of Chinese Syntax. *Advances in psychology*, 90:277-311.
- Chen, Keh-Jiaan, and Churen Huang 1993. *Zhongwen cilei fenxi* [Analysis of Parts of Speech in Chinese]. (CKIP Technical Report No. 93-05). Chinese Knowledge and Information Processing (CKIP), Institute of Information Science, Academic Sinica.
- Chen, Binbin, and Jingyu Zhang. 2022. Pre-Training-Based Grammatical Error Correction Model for the Written Language of Chinese Hearing-Impaired Students. *IEEE Access* 10:35061-35072.
- Cheng, Yong, and Mofan Duan. 2020. Chinese Grammatical Error Detection Based on BERT Model. In *Proceedings of the 6th Workshop on Natural Language Processing Techniques for Educational Applications*, Suzhou, pp.108-113.
- Chung, Chi-Hsuan. 2023. *AI-assisted Chinese Language Teaching: A Study of Ungrammatical Sentence Detection, Correction and Revision Suggestions*. Chiayi: National Chung Cheng University MA thesis, Taiwan.
- Chollet, François. 2017. *Deep Learning with Python*. Shelter Island, New York: Manning.
- DeKeyser, Robert. 1998. Beyond Focus on Form: Cognitive Perspectives on Learning and Practicing Second Language Grammar. In *Focus on Form in Second Language Acquisition*, eds. Catherine J. Doughty, and Jessica Williams, pp.42-63. Cambridge: Cambridge University.
- DeKeyser, Robert. 2003. Implicit and explicit learning. In *Handbook of Second Language Acquisition*, eds. Catherine J. Doughty, and Michael H. Long, pp.313-348. Oxford: Blackwell.
- DeKeyser, Robert, and Alan Juffs. 2005. Cognitive Conservation in L2 Learning. In *Handbook of Research in Second Language Teaching and Learning*, ed. Eli Hinkel, pp.437-454. Mahwah: Erlbaum.
- Ellis, Rod. 2004. The Definition and Measurement of L2 Explicit Knowledge. *Language Learning* 54:227-275.
- Ellis, Rod. 2006. Modeling Learning Difficulty and Second Language Proficiency: The Differential Contributions of Implicit and Explicit Knowledge. *Applied Linguistics* 27:431-463.
- Gentner, Dedre. 2003. Why we're so smart. In *Language in mind: Advances in the study of language and thought*, eds. Dedre Gentner, and Susan Goldin-Meadow, pp.195-235. Cambridge: MIT Press.
- Gentner, Dedre, and Arthur B. Markman. 1997. Structure mapping in analogy and similarity. *American Psychologist* 52:45-56.
- Gentner, Dedre, and Laura L. Namy. 2006. Analogical Processes in Language Learning.

- Current Directions in Psychological Science* 15(6):297–301.
- Hu, Guangwei. 2011. Metalinguistic Knowledge, Metalanguage and Their Relationship in L2 Learners. *System* 39:63-77.
- Johnson, Keith. 1996. *Language Teaching and Skill Learning*. Oxford: Blackwell.
- Krashen, Steven. 1987. *Principles and Practice in Second Language Acquisition*. Englewood Cliffs: Prentice Hall.
- Kuang, Hailan, Kewen Wu, Xiaolin Ma, and Xinhua Liu. 2022. A Chinese Grammatical Error Correction Method Based on Iterative Training and Sequence Tagging. *Applied Sciences* 12(9):4364.
- Kim, Woo-Hyun, and Jong-Hwan Kim. 2020. Individualized AI Tutor Based on Developmental Learning Networks. *IEEE Access* 8:27927-27937.
- Lee, Lung-Hao, Bo-Lin Lin, Liang-Chih Yu, and Yuen-Hsien Tseng. 2017. Chinese Grammatical Error Detection Using a CNN-LSTM Model. In *Proceedings of the 25th International Conference on Computers in Education (ICCE 2017)*, New Zealand, pp.919-921.
- Lee, Lung-Hao, Man-Chen Hung, Chao-Yi Chen, Rou-An Chen, and Yuen-Hsien Tseng. 2021. Chinese Grammatical Error Detection Using Adversarial ELECTRA Transformers. In *Proceedings of the 29th International Conference on Computers in Education (ICCE 2021)*, Thailand, pp.111-113.
- Li, Charles N., and Sandra Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California.
- Li, Si, Jianbo Zhao, Guirong Shi, Yuanpeng Tan, Huifang Xu, Guang Chen, Haibo Lan, and Zhiqing Lin. 2019. Chinese Grammatical Error Correction Based on Convolutional Sequence to Sequence Model. *IEEE Access* 7:72905-72913.
- Liu, Chen-sheng. 2007. The Week Comparative Morpheme in Mandarin Chinese. *Concentric: Studies in Linguistics* 33(2):53-89.
- Liu, Chen-sheng. 2010. The Positive Morpheme in Chinese and the Adjectival Structure. *Lingua* 120:1010-1056.
- Matthews, Clive. 1998. *An Introduction to Natural Language Processing through Prolog*. London and New York: Longman.
- Paradis, Michel. 1994. Neurolinguistic Aspects of Implicit and Explicit Memory: Implications for Bilingualism and SLA. In *Implicit and Explicit Learning of Language*, ed. Nicholas Carlton Ellis, pp.393-419. San Diego: Academic.
- Pullum, Geoffrey K., and Gerald Gazdar. 1982. Natural Language and Context-free Languages. *Linguistics and Philosophy* 4(4):471-504.
- Radford, Andrew. 1988. *Transformational Grammar: A First Course*. Cambridge: Cambridge University Press.
- Ren, Hongkai, Liner Yang, and Endong Xun. 2018. A Sequence to Sequence Learning for Chinese Grammatical Error Correction. In *Proceedings (part II) of the 7th CCF International Conference on Natural Language Processing and Chinese Computing (NLPCC 2018)*, Hohhot, pp.401-410.
- Selinker, Larry. 1969. Interlanguage. *IRAL* 3:114-132.
- Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N.

- Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in Neural Information Processing Systems 30*, eds. Isabelle Guyon, Ulrike von Luxburg, Samy Bengio, Hans Wallach, Rob Fergus, S. Vishwanathan, and R. Garnett. Red Hook, New York: Curran Associates.
- Wang, Wen-jet, Chia-jung Chen, Chia-ming Lee, Chien-yu Lai, and Hsin-hung Lin. 2019. *Linguistics-Oriented Keyword Interface NLU System* [Computer program]. (<https://api.droidtown.co>)
- Wang, Hongfei, Kurosawa Michiki, Katsumata Satoru, and Komachi Mamoru. 2020. Chinese Grammatical Correction Using BERT-based Pre-trained Model. In *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing (ACL-IJCNLP 2020)*, Suzhou, pp.163-168.
- Xiang, Yang. 2018. *Grammatical Error Identification for Learners of Chinese as a Foreign Language*. Uppsala: Uppsala University M.A. thesis, Sweden
- Yeh, Jui-Feng, Tsung-Wei Hsu, and Chan-Kun Yeh. 2016. Grammatical Error Detection Based on Machine Learning for Mandarin as Second Language Learning. In *Proceedings of the 3rd Workshop on Natural Language Processing Techniques for Educational Applications (NLPTEA2016)*, Osaka, pp.140-147.
- Zhang, Jinbin and Heng Wang. 2019. Multi-task Learning for Chinese Word Usage Errors Detection. In *Proceedings of the 3rd IEEE International Conference on Computational Intelligence and Applications (ICCIA 2018)*, Hong Kong, pp.93-96.
- Zhao, Zewei and Houfeng Wang. 2020. MaskGEC: Improving Neural Grammatical Error Correction via Dynamic Masking. In *Proceedings of the AAAI Conference on Artificial Intelligence 34(1) (AAAI 2020)*, New York, pp.1226-1233.
- Zhang, Junling (2020). A Review of Grammar Corrective Feedback: The Learning Experience in China's English Classes. *Studies in English Language and Teaching*, 8:127-137.

Jiun-Shiung Wu, Hsin-Hui Chen, Hui-Ru Yang

*Jiun-Shiung Wu
Graduate Institute of Linguistics
National Chung Cheng University
Chiayi County, Taiwan 621, ROC
Lngwujs@ccu.edu.tw*

*Hsin-Hui Chen
Chinese Language Center
National Chung Cheng University
Chiayi County, Taiwan 621, ROC
clchenhh@ccu.edu.tw*

*Hui-Ru Yang
Institute of Linguistics
National Chung Cheng University
Chiayi County, Taiwan 621, ROC
d12130002@ccu.edu.tw*

介紹 AGIBIS (基於輸入句的自動語法教學)
一個人工智慧輔助華語第二語言學習聊天機器人的前導研究

吳俊雄、陳欣徽、楊惠如
國立中正大學

本文介紹一個基於輸入句的華語語法自動教學系統，簡稱：AGIBIS。本系統接受使用者輸入一個華語句子，分析其語法點，回傳給使用者選擇。使用者可以選擇有興趣的語法點，請系統回傳其說明。AGIBIS 建立於 LOKI，一個基於語言學研究、可以使用小數據達到自然語言理解的模型。本文介紹讓 LOKI 可以達到小數據分析的三個特徵。此外，我們也介紹 AGIBIS 的架構：此版本的 LOKI 由六個意圖組成，其中五個意圖用來分析句子的語法點，剩下的意圖則觸解讀讀者的回覆。我們示範 AGIBIS 如何運作，也藉由與其他自然語言處理運用於華語結構教學的研究，來討論 AGIBIS 的利基。最後，我們也討論 AGIBIS 在華語教學上的用應用與啟示。

關鍵字：自動語法教學、人工智慧輔助華語學習、人工智慧輔助華語教學

TEACHING AND LEARNING IN AN ENGLISH-MEDIUM INSTRUCTION COURSE IN UNIVERSITY: A TWO-YEAR EXPERIENCE OF THE VEGETABLE DISEASES COURSE FOR GENERAL EDUCATION

Ying Ying Wong, Yuan-Min Shen, Cheng-Chun Huang
National Taiwan University

ABSTRACT

A newly designed English-medium instruction (EMI) course titled “Vegetable Diseases”, originally created for graduate students in the Master Program for Plant Medicine, has been transformed into a general education course for university students who have English as their second language across a wide variety of backgrounds. The teaching and learning experiences of the course have been documented during the 2022-Spring and 2023-Spring semesters. While a few teaching plans were influenced by the COVID-19 pandemic, the students provided positive feedbacks on the EMI learning process. Experiential learning activities, such as vegetable growing, farm visiting, face-to-face interaction with farmers and experts, were included, contributing to the achievement of the learning objectives. Based on the survey results regarding the difficulty level in this EMI learning environment, students from different colleges exhibited similar attitudes. However, most of the foreign students did not encounter difficulties in the EMI subjects, whereas some Taiwanese students faced challenges in the EMI learning process. The supporting teaching practices concerning the experiential learning activities in this study could serve as examples for future EMI curriculum designs.

Keywords: English-medium instruction, second language, agriculture, experiential learning, scholarship of teaching and learning

1. INTRODUCTION

1.1 Introduction to English-medium Instruction (EMI)

English-medium instruction (EMI) involves teaching academic subjects (excluding English) in countries where the primary language spoken by most people isn't English, using the English language as the medium of instruction (Dearden 2014). In contemporary academia, the implementation of EMI in universities has become a prominent concern, particularly in Asian countries such as Malaysia, Indonesia, Vietnam, Hong Kong, Taiwan, China, Japan, and Korea (Kaur 2020; Pun and Jin 2021; Nguyen et al. 2017; Wu 2006; Jiang et al. 2019; Galloway et al. 2017; Kym and Kym 2014). This heightened focus on English can be attributed to two main factors: the prevalence of English as a lingua franca (ELF) and the drive towards higher educational internationalism (Kuteeva 2020; Rose and McKinley 2018). The use of EMI in Asia originated in former British colonies like Malaysia, Hong Kong, and Singapore, which implemented bilingual educational programs. In response to globalization and the desire for higher education quality, non-colonized countries like Japan and Korea began incorporating EMI courses into their educational systems (Tsou and Kao 2017).

In 2016, Taiwan's president launched the "New Southbound Policy" as a strategic initiative to build lasting, mutually beneficial relationships with countries in ASEAN and South Asia (Andoko et al. 2022). Through the New Southbound Policy, Taiwan actively seeks to enhancing cooperation, understanding, and collaboration with these regions. The development of agriculture in Taiwan has garnered significant interest, attracting a diverse group of students, many of whom are international and have a second language other than Chinese. Utilizing English as a medium of instruction can be an effective way to attract students to take these courses. The Ministry of Education (MOE) in Taiwan has launched a strategic initiative to make Taiwan a bilingual country by 2030 (Ferrer and Lin 2024). In 2021, MOE announced the Program on Bilingual Education for Students in College (BEST), which is set to run from 2021 to 2030. National Taiwan University is one of the four focus universities chosen to participate in this initiative. To support this vision of bilingualism, the

MOE has adopted a policy promoting the integration of EMI in higher education. This strategic aim is to create an environment where students can engage with academic content in English. This emphasis on EMI will foster a more inclusive and internationally-oriented educational experience, ultimately enhancing their global competitiveness. Two sides of the same coin: While EMI offers many benefits, we must also be concerned that students may struggle to address the potential challenges of understanding different accent, acquiring subject-specific vocabulary, and keeping pace with spoken English, all while developing listening, speaking, and writing skills. These challenges can affect their self-assurance and involvement in class, as students may not feel comfortable using English consistently in EMI classes (Siddiqui et al. 2021).

1.2 The Theory of Experiential Learning

Experiential learning, as defined by Gentry (1990), is the process of acquiring knowledge through hands-on experiences. This approach finds its foundation in the theories of John Dewey, Kurt Lewin, and Jean Piaget (Miettinen 2000). David Kolb, a significant contributor to this field, enriched the discourse with his learning styles model introduced in 1984. Kolb's experiential learning cycle delineates four distinctive stages, creating a comprehensive framework for understanding the learning process (Kolb 1984).

The experiential learning cycle, as articulated by Kolb, unfolds in four distinct stages. Initially, learners immerse themselves in concrete experiences, engaging directly with real-world situations. This hands-on involvement serves as the foundation for the entire learning process. Subsequently, learners transition to reflective observation, taking a step back to contemplate and analyze their experiences. This phase encourages introspection, deepening their understanding.

Building on reflections, learners progress to abstract conceptualization, where they draw connections between their experiences and existing theoretical concepts. This stage facilitates the development of new ideas and a more profound comprehension of the subject matter. Finally, the cycle culminates in active experimentation, urging learners to apply the insights gained through reflection and conceptualization in real-world

scenarios. This last stage completes the experiential learning journey, emphasizing the practical application of acquired knowledge.

Kolb's experiential learning cycle represents a dynamic, iterative process where individuals engage in real-life experiences, reflect on them, conceptualize abstract ideas, and apply their understanding through active experimentation. Learning, in essence, emerges from the harmonious interplay between individuals and their environment, integrating new experiences into existing concepts and adapting current concepts to align with fresh experiences (Kolb and Kolb 2006).

2. DESIGN OF THE VEGETABLE DISEASES COURSE AND THE STUDENT'S BACKGROUND

The Vegetable Diseases course integrated the concept of EMI and experiential learning in both the spring semesters of 2022 and 2023. The course is designed to cultivate an educational environment centered around the study of vegetable diseases. This approach enables students to acquire fundamental and current knowledge about vegetable diseases, achieved through readings, hands-on practical experiences, and engaging discussions. To assess the effectiveness of this approach, we utilized Google Forms to collect and analyze feedback and comments from students over these two academic years. This study aims to investigate the learning outcomes of students enrolled in a general education course focused on vegetable diseases within an EMI environment. It seeks to assess how students comprehend the course content under EMI in this specialized field.

The composition of students enrolled in the Vegetable Diseases course during the academic years 2022 and 2023. Over this period, the class comprised 28 male students and 19 female students, contributing to a total of 47 students. Notably, the student body consisted of 13 international students and 34 Taiwanese students. In the academic year 2022-2023, the course attracted a diverse group of students. The distribution by academic year is as follows: 38.30% were freshmen, indicating a significant presence of first-year students eager to explore the Vegetable Diseases course. Additionally, 2.13% were sophomores, 28.26% were juniors,

21.28% were seniors, and 10.64% were graduate students. This diverse mix of students from different academic levels contributed to a rich and dynamic learning environment.

Regarding the academic backgrounds of the students, the diversity of disciplines among the participants further enriched the learning experience. Specifically, around 29.79% (14 students) came from the College of Liberal Arts, Social Sciences, or Management, representing a broad spectrum of non-science majors. Additionally, 17.02% (8 students) were affiliated with the College of Science, Engineering, Electrical Engineering, or Computer Science. The largest contingent, comprising approximately 53.19% (25 students), emanated from the College of Bio-resources, Agriculture, or Medicine, bringing a strong foundation in agricultural and life sciences. This varied academic composition fostered interdisciplinary discussions and perspectives throughout the course.

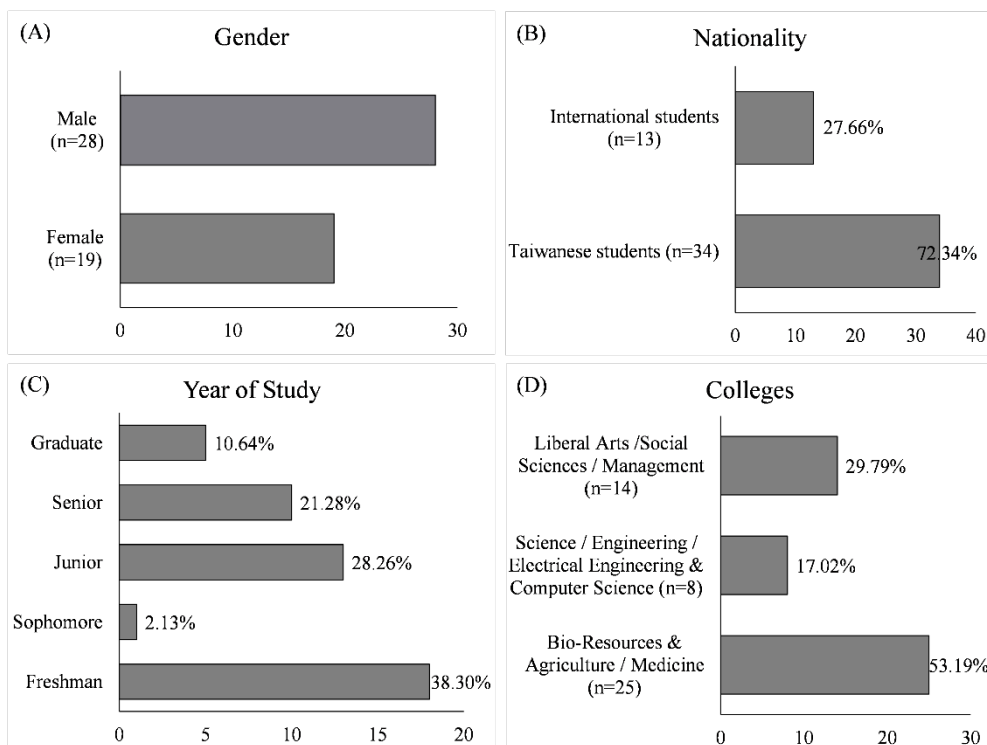


Figure 1. Students' background in the 2022-Spring and 2023-Spring enrollments.

3. MOTIVATION OF STUDENTS ENROLLING IN THE VEGETABLE DISEASES COURSE

During the spring semester of 2023, we conducted a preliminary survey to gain insights into the motivations and expectations of students enrolled in the Vegetable Diseases course. The survey drew participation from 19 students, and the results provided valuable information about their varied interests and goals within the course. Among the key findings, it was observed that nearly 15 students exhibited a keen interest in acquiring skills related to the diagnosis of plant diseases. This enthusiasm indicated a strong desire among participants to delve into the practical aspects of identifying and understanding various plant disease. Furthermore, 13 students expressed a specific interest in engaging in hands-on experiences related to cultivating their own vegetables. This aspect of the course appealed to their desire for experiential learning, allowing them to directly apply theoretical knowledge to real-world scenarios.

Interestingly, the survey unveiled that 9 out of the 19 students recognized the significance of the course's English-medium environment in enhancing their language proficiency. This acknowledgment highlighted the dual benefits of the course, not only focusing on the subject matter but also contributing to the improvement of English language skills. Lastly, 3 students articulated a desire to develop their ability to read and comprehend scientific papers, showcasing a more academic-oriented objective. While not explicitly shown in the provided data, these aspirations underline the diverse set of motivations among students in the Vegetable Diseases course. In summary, the preliminary survey provided a comprehensive understanding of the students' motivations, including their interest in practical skills, language proficiency, and academic pursuits. These insights lay the groundwork for tailoring the course to meet the varied expectations and goals of the participants.

4. CRAFTING EXPERIENTIAL LEARNING FOR VEGETABLE DISEASES COURSE

In the context of EMI, the Vegetable Diseases course underwent a notable transformation, transitioning from a specialized study to a comprehensive liberal education experience (Shen et al. 2023). This evolution was characterized by the deliberate integration of experiential learning activities, reshaping the course to incorporate interactive sessions with farmers, hands-on cultivation, knowledge-sharing, farm visits, and practical experiments. The encouragement for students to cultivate their own vegetables fostered a profound connection with the subject matter. Through midterm and final assessments, students were provided with opportunities for reflection, enabling them to share their insights. The course reached its pinnacle with a visit to the NTU farm, effectively bridging the gap between theoretical and practical learning, thereby enriching the overall educational experience.

4.1 Hands-on Experience in Vegetable Cultivation

During the spring semesters of 2022 and 2023, the course lecturer provided students with a variety of seedlings and seeds for their cultivation practices. A diverse selection of seedlings, including chili, eggplant, lettuce, cucumber, scallion, and garlic chive seedlings, were made available to students. Besides, seeds for tomatoes, green peppers, pak choi, and water spinach were provided, enabling students to explore a range of vegetables in their cultivation journey. Understanding the importance of practical guidance, the course also equipped students with a comprehensive guide covering the fundamentals of vegetable gardening. This guide served as a valuable resource, offering insights into various aspects of cultivation, from planting techniques to crop care.

Approximately two weeks after students received the seeds and seedlings, they were provided with basal and organic fertilizers, ensuring the necessary nitrogen, phosphorus, and potassium nutrients for optimal vegetable growth. This aimed to ensure that the growing plants received the essential nutrients, fostering healthy and robust vegetable growth. This comprehensive approach not only enriched students' practical experiences

Ying Ying Wong, Yuan-Min Shen, Cheng-Chun Huang

but also emphasized the importance of holistic knowledge in successful vegetable cultivation.



Figure 2. Preparation of various vegetable seedlings in the classroom

4.2 Campus Farm Visits

During the spring semester of 2022, the NTU Farm trip followed a planned route, began at Gongtong Lecture Building. Our trip began with a walk along Zhoushan road. Our observations started at the greenhouse, where we examined vegetable diseases and insect pest damage. Subsequently, we proceeded to NTU farm exploring various crops, including corn, cucumber, paddy, broccoli, water spinach and ganges amaranth. Our next destination was the NTU Plant Teaching Hospital (NTUPTH), where we interacted with plant doctors who provided valuable insights into their participation in diagnostic activities related to plant diseases, insect damage, and crop management. The trip culminated in a delightful experience at the NTU Farm Visitor Center, where we savored the delicious NTU sandwich ice cream.

In spring semester of 2023, we implemented a few modifications to the NTU Farm trip route. Our visit to the greenhouse provided an opportunity to observe and diagnose various vegetable diseases. We observed fungal diseases like powdery mildew and downy mildew on cucumbers, as well as insect pest like *Lepidoptera* larvae on cornstalks and thrips on scallions. Our educational trip continued with a visit to the Iso House NTU, a historical place known as the birthplace of Taiwan Penglai Rice. The journey concluded with a delightful experience at the NTU Farm Visitor Center, where we had the pleasure of trying the delicious NTU sandwich ice cream.



Figure 3. Insect damage and vegetable disease observations from NTU farm visiting



Figure 4. Field visiting at NTU farm in spring 2023.

4.3 Face-to-face Interaction with Farmers and Experts

The Vegetable Diseases course actively involved agricultural experts and farmers, allowing students to directly interact with them. This not only enriches the learning experience but also ensures that students stay updated on the latest developments and conditions in agriculture.

In the 2022-Spring semester, we had the privilege of hosting Frank, the owner of De An Gu organic farm, who shared insights into organic vegetable cultivation and the dynamics of the vegetable market. However, due to the challenges posed by the COVID-19 pandemic, the Vegetable Diseases course shifted to an online format. During this period, Willie Chen Wu-Yang from the World Vegetable Center (AVRDC) joined us to discuss the management of solanaceous vegetable diseases.

In the 2023-Spring semester, the course maintained its tradition of inviting experts to enrich our learning experience. Dr. Chen Wei-Ling from the Taichung District Agricultural Research and Extension Station, enlightened us on the current state and future development of the vegetable industry in Taiwan. In addition, we were honored to have a young farmer, Mr. Chen, who shared valuable insights into organic leafy vegetable cultivation and farm management. Notably in this case, Taiwanese students were invited to translate the farmer's supplemental information from Chinese into English, improving the understanding for all international students. The expertise of these diverse professionals significantly enhanced the depth and perspective of our exploration into the realm of vegetable diseases and their management.



Figure 5. Two expert lecture sessions during the 2023-spring semester

4.4 Collecting Field Samples of Vegetable Diseases and Experiments in The Classroom

This course incorporates actual plant samples of vegetable diseases directly into the classroom, with the goal of offering students hands-on learning experiences through practical activities rather than relying solely on PowerPoint slides. The objective is to amplify practical learning opportunities and elevate the overall educational experience.

During the class, the teaching assistant prepared samples relevant to each week's topic and collaborated with plant doctors to obtain field plant samples. In the topic of Solanaceae crops, for instance, we collected plant samples encompassing Tomato Mosaic Virus (ToMV), early blight, late blight, and *Stemphylium* sp. Additionally, when focusing on Cucurbitaceae crops, we collected samples such as powdery mildew and downy mildew on bottle gourd.

In Spring 2022, our class had the privilege of hosting a master's student from Master Program for Plant Medicine, who shared the expertise in diagnosing tomato bacterial wilt caused by *Ralstonia solani*. The practical session involved essential materials such as infected tomato plant samples, florist scissors or knife, a transparent container, and water. The diagnostic procedure commenced with observing the tomato plant's condition, identifying typical symptoms such as wilting of the tree and browning of stems and leaves. To test for bacterial infection, participants use a knife or scissor to cut the tomato stem with browning and immersed it in a transparent container with water. The presence of white to light grey bacterial ooze seeping from the infected tissue, typically observed in the stem section, indicated bacterial infection. This hands-on activity provided students with valuable experience in diagnosing *Ralstonia solani* infection in tomato plants, enhancing their practical skills in plant pathology.

In the year of 2023-spring, we designed a simple experiment to test the efficacy of antibiotics in inhibiting bacterial growth. The requisite materials for this experiment encompassed kasugamycin with copper, kanamycin antibiotics, *Pectobacterium carotovorum* subsp. *carotovorum* (*Pcc*), LB agar plate, sterilized paper disc, forceps, cotton swab, parafilm, disposable rubber glove, marker pen, tissue paper, and alcohol. The

procedural steps for this experiment were as follows: Initially, the cotton swab was immersed in the *Pcc* bacterial culture solution. Then, the swab was gently used to spread the bacterial solution uniformly on the plate. Next, a paper disc pre-soaked with an antibiotic solution was placed on the plate. The plate underwent incubation for 16 to 20 hours within a temperature range of 26 to 30°C. Finally, the measurement and recording of the diameter of the inhibition zone surrounding the antibiotics disc were carried out. While collecting the students' culture media, we found that the majority of their experiments were successful. Given the diverse backgrounds of the students, our aim was to provide each of them with a unique hands-on experience through this simple experiment. This practical experience allowed students to actively engage and understand how to test the microbicidal effects of chemicals on pathogens, contributing to a more dynamic and enjoyable course.

The demonstration using field samples and the planning of in-class experiments across the two semesters engaged students in highly immersive activities. To deepen student understanding of plant diseases in spring 2022, the course featured a session on diagnosing tomato bacterial wilt led by a guest expert. Later in the subsequent 2023-spring semester, we introduced a slight modification. Instead of the previous diagnostic approach, we engaged students in a hands-on experiment centered around the inhibiting effects of antibiotics on bacterial pathogen growth. These two activities served as invaluable learning experiences, not only deepening students' understanding of plant diseases but also refining their practical skills.



Figure 6. Sample preparation for tomato phytophthora disease, cucumber downy mildew, and cucumber powdery mildew



Figure 7. Experiment utilizing antibiotics to inhibit bacterial growth

4.5 Knowledge Sharing

The Vegetable Diseases course is conducted in English and is designed to impart specialized knowledge through a blend of theoretical and practical sessions. One integral aspect is the inclusion of oral presentations within the course framework. During the spring semester of 2022, due to the COVID-19 pandemic, the end-of-semester group presentations were done via video recordings. This approach ensured students' safety and prevented classroom overcrowding. On May 17 and May 24, 2023, group presentations were conducted in classroom, sharing their findings from scientific journal reviews in class. Furthermore, a master's student provided valuable insights by delivering a presentation on soil-borne damping-off disease, offering students an enriched understanding of the topic. In the final course session, students shared their cultivation experiences, revealing the joy they derived from planting. In Taiwan, where English is not commonly used in students' daily lives, the experience of these group presentations could be invaluable. This offered students with unique opportunities to practice and enhance their English speaking and listening skills.



Figure 8. Students' knowledge sharing on scientific paper review and cultivation experiences related to the topics

5. SURVEY ON LEARNING IN THE VEGETABLE DISEASES COURSE

Throughout the academic year 2022-2023, this study collected data during two distinct semesters: 2022-Spring and 2023-Spring. The primary objective was to gain a comprehensive understanding of students' experiences and outcomes over these two semesters, enabling an analysis of the course's impact on their learning journey. Across both semesters, a total of 47 students were enrolled in the Vegetable Diseases course. The survey garnered 16 responses from the 2022-Spring semester and 20 responses from the 2023-Spring semester, totaling 36 collected responses. Additionally, the pre-test administered in the 2023-Spring semester elicited 19 responses.

In the process of data collection, a Likert scale ranging from 1 to 5 was utilized, allowing participants to express their agreement levels, spanning from strongly disagree to strongly agree. The survey for the 2022-Spring semester took place at the end of the semester. For the 2023-Spring semester, a survey was administered at the end of the semester, primarily focusing on students' self-assessment of the planned course activities, such as agricultural specialists' lectures, sharing of farming experiences by

farmers, students' personal vegetable cultivation experiences, farm visits, and other arrangements designed to assess their impact on learning outcomes.

Implement Google Forms to construct a focused questionnaire aimed at systematically exploring and evaluating the impact of an EMI environment on learning outcomes. The findings were graphically illustrating both the average values and standard error calculations. Additionally, students were encouraged to offer concise comments and feedback on their experiences with the Vegetable Diseases course.

Here are the findings for six questions regarding to the EMI learning environment and the impact of course planning on achieving the learning objectives, as presented in Table 1. For the statement "Understanding the content of vegetable diseases in an EMI environment is too difficult for me", the average Likert scale value of 2.69 slightly exceeds the midpoint, indicating a moderate level of difficulty reported by the students. The survey results shed light on participants' perceptions of comprehending vegetable diseases in an EMI environment. In the "agree" category, 4 out of 5 students were Taiwanese. Similarly, in the "strongly agree" category, 4 out of 5 students were also Taiwanese. This indicates a substantial majority of Taiwanese students expressing "agree" or "strongly agree" with the statement, suggesting that while there might be some challenges in an EMI environment. The findings of this study correspond with the conclusions drawn in our previous study (Shen et al. 2023).

Furthermore, the survey delved into participants' responses regarding the statement "Experience sharing of organic vegetable cultivation by the expert contributes to the learning objectives of the Vegetable Disease course", participants expressed a notably positive sentiment with an average Likert scale value of 4.58. This suggests a strong endorsement of the experiential and expert-driven elements in enhancing their learning experience. Furthermore, participants expressed positive sentiments towards additional components of the course. For the statement "Experience sharing of vegetable disease management studies by the expert contributes to the learning objectives of the Vegetable Diseases course", the average Likert scale value was 4.53. This signifies a strong acknowledgment from students that the insights shared by the expert significantly contributed to meeting the learning objectives of the course,

highlighting the value of expert perspectives in enriching the educational experience.

Similarly, for the statement "The field trip to the farm helps me establish knowledge connections with the course content of vegetable diseases", participants demonstrated high agreement with an average Likert scale value of 4.58. This indicates that the practical exposure gained from the farm visit significantly contributed to consolidating their understanding of the course content. These findings collectively underscore the overall positive impact of experiential elements in the course and highlight the effectiveness of these strategies in facilitating comprehension and knowledge integration for the students. The fifth statement, "The activity of growing vegetables by myself helps me establish knowledge connections with the course content of vegetable diseases", garnered the highest average Likert scale value of 4.64. This suggests a particularly strong agreement among participants regarding the positive impact of hands-on vegetable cultivation on reinforcing their understanding of the course content. The experiential aspect of growing vegetables emerged as a highly effective strategy in enhancing their learning experience.

Lastly, for the statement "I am interested in taking care of plants, I will try to grow some plants in the future", participants expressed a positive inclination with an average Likert scale value of 4.42. The positive course feedback suggests the experiential components effectively enhanced their interest in plant care and future application of the knowledge.

Table 1. Survey on EMI and learning outcomes

Questions	Likert Scale Rating					Average	Standard Error
	1	2	3	4	5		
1. It is too difficult for me to understand the content of vegetable diseases in an English-Medium Instruction environment.	8	10	8	5	5	2.69	0.22

2. Experience sharing of organic vegetable cultivation by the expert contributes to the learning objectives of Vegetable Disease course.	0	0	3	9	24	4.58	0.11
3. Experience sharing of vegetable disease management studies by the expert contributes to the learning objectives of Vegetable Diseases course.	0	0	2	13	21	4.53	0.10
4. The field trip to the farm helps me to establish knowledge connections with the course content of vegetable diseases.	0	0	2	12	22	4.58	0.10
5. The activity of growing vegetables by myself helps me to establish knowledge connections with the course content of vegetable diseases.	0	0	3	7	26	4.64	0.11
6. I am interested in taking care of plants, I will try to grow some plants in the future.	0	0	5	11	20	4.42	0.12

6. STUDENTS FEEDBACK AND COMMENT ABOUT THE COURSE

Analyzing the statistical outcomes of the course questionnaire for both the 2022-spring and 2023-spring semesters, a total of 18 participants engaged in the evaluation using the Likert scale. Evaluating the statement "The instructor was devoted to teaching", 6 students (33.33%) agreed, and 12 students (66.67%) strongly agreed. Moving on to the statement "In general, this is an effective course", 2 students (11.11%) were neutral, 7 students (38.89%) agreed, and 9 students (50%) strongly agreed. Regarding the statement "I would recommend this instructor to other students", 9 students (50%) agreed, and 9 students (50%) strongly agreed. Finally, in response to the statement "This course helps me achieve course objectives", 1 student (5.56%) chose neutral, 8 students (44.44%) agreed, and 9 students (50%) strongly disagreed. In summary, based on this teaching evaluation, the course is perceived as effective. Student course evaluations are essential for reflecting the quality of the course and providing instructors with feedback to improve students' learning experiences. In the spring 2023 end-of-semester survey, participant provided concise comment and feedback on their experiences in the Vegetable Diseases course.

- (a) "The class is good, perhaps we could have more practice class outside to gain more experience. – Form Indonesia, Department of Horticulture and Landscape Architecture. "
- (b) "I really liked the course, especially the second part about the management of plant diseases. The first part about the identification of diseases was a bit more boring but I understand that it was necessary. – From Japan, Department of Social Work. "
- (c) “能自己試試看種蔬菜真的很有趣，雖然中間經歷了好多困難，有蟲、風太大、換盆，但還是很有趣的經驗！– Form Taiwan, Department of Plant Pathology and Microbiology.”

- (d) "Thank you very much for this class. I like how the professor have us to gather around and that prevent me to be sleepy. I also enjoy the trip to NTU farm very much. – From Indonesia, Department of Nursing."
- (e) "Both the TA and professor are really invested, which makes it easier for us to be invested as well! – Form French, Department of Foreign Languages and Literature."
- (f) “很喜歡教授和助教上課的方式，輕鬆愉快又能學習到許多知識，且互動性高的課程可以幫助學生更加投入。– Form Taiwan, Department of Agricultural Economic.”
- (g) "Thank you for teaching this course! I think the more practical parts of the course were the most interesting. – Form Dutch, School of Forestry and Resource Conversation. "
- (h) “會選擇這堂課的原因是當初在課程網看到可以種植自己的植物，就覺得這應該是一堂非常有趣的課，經過一整個學期後，真的非常有趣！教授與助教人非常好，雖然自己所種的植物失敗了，但還是從這堂課中學習到許多知識，如：植物病害、抗病蟲的方法等，往若有機會，希望可以再學習有關植物或是微生物的知識，謝謝老師！– Form Taiwan, Department of Bioenvironmental System Engineering.”

The feedback from students provided valuable insights into their diverse perspectives on the Vegetable Diseases course. Students from Indonesia, Japan, Taiwan, France, and other backgrounds expressed their gratitude for the course structure, content, and hands-on experiences. The comments also underscored the importance of striking a balance between theoretical knowledge and its practical application, recognizing the necessity of both aspects in understanding plant diseases. Overall, the feedback reflected the course's success in providing a well-rounded and engaging learning experience for students with diverse academic backgrounds and interests.

7. FINDINGS DISCUSSION

In the academic years spanning from 2022 to 2023, a predominant number of students enrolled in courses were non-native English speakers. The success of effective EMI hinges on the proficiency of both the EMI teachers and their diverse student body. If either party lacks robust language skills, it can significantly impede the effectiveness of the learning experience. The transformation of the Vegetable Diseases course into a general education option for university students has granted valuable insights into the diverse backgrounds of the participants. This approach was expanded, transforming it into a general education course that was accessible to all university students.

The course garnered interest from students with varied academic backgrounds, encompassing fields such as art and electronics, thereby extending its appeal beyond the conventional boundaries of the Bioresources department. To enrich the learning experience, we introduced a antibiotics test activity. This hands-on activity provided students with practical insights into various aspect of vegetable diseases, contributing to a diverse range of learning experiences. Customizing the course to accommodate students from different colleges not only broadened their perspectives but also seamlessly integrated liberal education principles within the EMI framework.

The integration of experiential learning activities, such as vegetable cultivation, farm visits, and direct interactions with farmers and experts, has demonstrated its effectiveness. With meticulous planning, these activities not only boost interest in learning but also demand minimal time and effort. However, the outcomes in this course highlight the potential of EMI, but verification through research is essential to determine its suitability for wider implementation across universities and various courses.

Place-based education (PBE) is an educational method that utilizes the local community and environment, emphasizing hands-on, real-world learning experiences (Sobel 2004). Our teaching strategies may closely link to PBE and similar educational frameworks, highlighting the importance of real-world context and hands-on experiences in learning (Deringer 2017; Powers 2004). Within the Vegetables Diseases course,

this could involve activities such as vegetable growing, farm visits, and interactions with farmers and experts. PBE connects students with their surroundings and community, providing practical insights into agricultural practices and disease management strategies. The positive impact of experiential learning activities on teaching effectiveness suggests potential benefits for broader educational approaches. When thoughtfully designed, these activities not only kindle students' interest but also demand minimal time and effort investment. Future research should explore how these insights can be used in various academic settings, like different courses and institutions. This will help us better understand the benefits of hands-on learning in EMI environments.

Despite their varied origins, students from different colleges showcased similar perspectives on the challenges posed by EMI. Foreign students, generally more accustomed to EMI subjects, demonstrated a smoother adaptation to the language demands of the course. Their familiarity with English as a medium of instruction facilitated a more seamless integration into the academic content. On the other hand, some Taiwanese students faced challenges in the learning process. This exploration of the participants' experiences highlights the importance of tailoring EMI courses to accommodate the diverse linguistic backgrounds of the student body. It serves as a foundation for ongoing efforts to enhance the effectiveness of EMI, ensuring that it remains accessible and beneficial for students from all linguistic backgrounds.

To improve the experience of Taiwanese students encountering difficulties in the EMI learning process, Tsui (2017) proposes a framework for developing EMI teachers comprising three key components: English language training, pedagogical training, and licensing certification. EMI teachers are encouraged to obtain an EMI teaching certificate to validate their qualifications for teaching at the university level. However, determining appropriate English language standards for evaluating lecturers' performance has been challenging (Dimova 2020). Additionally, to encourage more class interaction, teachers may allow students to speak in English or Chinese while responding only in English. Creating an authentic English learning environment is crucial for EMI learners to enhance their language proficiency and build confidence. To support this, teachers should provide

clear preparation guidance, conduct lectures in a fluent and logical manner, and incorporate strategy-based instruction to help students effectively overcome challenges (Xiao and Zou 2020).

Although the data on students exempt from advanced English courses doesn't directly show the English skills of those in the Vegetable Diseases course, it does highlight a pattern. Every year, a significant number of students at the university are exempted, indicating many have strong English abilities. This suggests that university students generally have good English skills or can adapt well to an EMI environment, showing positive results in language proficiency. In future research, the goal is to involve a larger group of students in surveys and interviews. This will help ensure the data collected is both accurate and representative of the diverse student body. Building on the improvements made in this study, a key enhancement involves actively engaging a more extensive pool of students in interviews, a process conducted in either Chinese or English (Teo et al. 2019). During these interviews, participants are encouraged to openly share their observations on the study's findings, offer insights based on their experiences, and respond to specific questions. This approach is expected to provide a deeper understanding of the complex dynamics of the learning process as seen by the students. The valuable feedback gathered through these interviews will serve as a foundation for future adjustments in our teaching methodologies.

8. CONCLUSION

This article, serving as a noteworthy case study, offers valuable insights and guidance for instructors in higher education institutions. Notably, in English major courses, international students exhibit a commendable level of adaptation compared to their English-first-language counterparts, while some Taiwanese students encounter challenges in an all-English learning environment. The findings from English major courses indicate an overall positive adjustment to specialized content across various colleges. To establish the generalizability of these insights, further research on teaching practices in diverse institutions is necessary. The positive outcomes observed in this article underline the effectiveness

of experiential learning, demonstrating that it doesn't necessitate an excessive investment of time and effort. This study showcases how experiential learning can transform professional courses into practical experiences within the realm of English general education.

The future prospects for EMI learning, especially in courses like Vegetable Diseases that incorporate experiential learning components, are promising and multifaceted. As the global demand for interdisciplinary and hands-on education continues to rise, the integration of EMI with experiential learning is expected to play a pivotal role. In the context of courses like Vegetable Diseases, EMI coupled with experiential learning not only enhances language proficiency but also provides students with practical insights into the subject matter. Looking ahead, the future of EMI learning with experiential components in courses like Vegetable Diseases holds great promise. It is poised to contribute to the development of well-rounded, globally aware individuals equipped with both language proficiency and practical skills.

REFERENCES

- Andoko, Effendi, Iqbal Rafani, and Wan-Yu Liu. 2022. Review of implementation of Taiwan's new southbound policy for agriculture. FFTC agricultural policy platform. Food and Fertilizer Technology Center (FFTC). Retrieved February 15, 2024, from <https://ap.ffc.org.tw/article/3033>.
- Dearden, Julie. 2014. *English as a Medium of Instruction-a Growing Global Phenomenon*. London: British Council.
- Deringer, S. Anthony. 2017. Mindful place-based education: Mapping the literature. *Journal of Experiential Education* 40(4):333-348.
- Dimova, Slobodanka. 2020. Language assessment of EMI content teachers: What norms. In *Language Perceptions and Practices in Multilingual Universities*, eds. Maria Kuteeva, Kathrin Kaufhold, and Niina Hynninen, pp.351-378. London: Palgrave Macmillan.
- Ferrer, Alessandra, and Tzu-Bin Lin. 2024. Official bilingualism in a multilingual nation: a study of the 2030 bilingual nation policy in Taiwan. *Journal of Multilingual and Multicultural Development*, 45(2):551-563.
- Galloway, Nicola, Jaroslaw Kriukow, and Takuya Numajiri. 2017. *Internationalization, Higher Education and the Growing Demand for English: An Investigation into the English Medium of Instruction (EMI) Movement in China and Japan*. London: British Council.
- Gentry, James W. 1990. What is experiential learning. In *Guide to business gaming and experiential learning*, ed. J. W. Gentry, pp.9-20. London, England: Nichols/GP Publishing.
- Jiang, Li, Lawrence Jun Zhang, and Stephen May. 2019. Implementing English-medium instruction (EMI) in China: Teachers' practices and perceptions, and students' learning motivation and needs. *International Journal of Bilingual Education and Bilingualism* 22(2):107-119.
- Kaur, Jagdish. 2020. Using English for interaction in the EMI classroom: Experiences and challenges in a Malaysian public university. In *English-medium instruction and the internationalization of universities*, eds. Hugo Bowles, and Amanda C. Murphy, pp.129-154. London: Palgrave Macmillan.
- Kolb, Alice Y, and David A Kolb. 2006. Learning styles and learning spaces: A review of the multidisciplinary application of experiential learning in higher education. In *Learning Styles and Learning: A Key to Meeting the Accountability Demands in Education*, eds. Ronalds R. Sims, and Serbrenia J. Sims, pp.45-91. New York: Nova Publishers.
- Kolb, David A. 1984. *Experiential Learning: Experience as the Source of Learning and Development*. (Vol. 1). Englewood Cliffs, NJ: Prentice Hall.
- Kuteeva, Maria. 2020. Revisiting the 'E' in EMI: Students' perceptions of standard English, lingua franca and translingual practices. *International Journal of Bilingual Education and Bilingualism* 23(3):287-300.

- Kym, Inyoung, and Moon Hyun Kym. 2014. Students' perceptions of EMI in higher education in Korea. *The Journal of Asia TEFL* 11(2):35-61.
- Miettinen, Reijo. 2000. The concept of experiential learning and John Dewey's theory of reflective thought and action. *International Journal of Lifelong Education* 19(1):54-72.
- Nguyen, Huong Thu, Ian Walkinshaw, and Hiep Hoa Pham. 2017. EMI programs in a Vietnamese university: Language, pedagogy and policy issues. In *English Medium Instruction in Higher Education in Asia-Pacific: From Policy to Pedagogy*, eds. Ben Fenton-Smith, Pamela Humphreys and Ian Walkinshaw, pp.37-52. Cham, Switzerland: Springer.
- Pun, Jack, and Xina Jin. 2021. Student challenges and learning strategies at Hong Kong EMI universities. *Plos One* 16(5):e0251564.
- Powers, Amy L. 2004. An evaluation of four place-based education programs. *The Journal of Environmental Education* 35(4):17-32.
- Rose, Heath, and McKinley Jim. 2018. Japan's English-medium instruction initiatives and the globalization of higher education. *Higher Education* 75:111-129.
- Shen, Yuan-Min, Ying Ying Wong, and Cheng-Chun Huang. 2023. Conversion of a specialized course into liberal education in an English-medium instruction environment through experiential learning: A case study in the course of vegetable diseases. *Journal of Scholarship of Teaching and Learning* 3(3):1-28. DOI: 10.7007/JSoTL.202309_3(3).0001
- Sobel, David. 2004. *Place-based Education: Connecting Classrooms and Communities*. Great Barrington, MA: The Orion Society.
- Siddiqui, Kamran Akhtar, Hassan Syed, and Zafarullah Sahito. 2021. Challenges and strategies of university students in English-medium instruction (EMI) classes: A case study. *Foro Educacional* 36:9-25.
- Teo, Timothy, Cathy Ka Weng Hoi, Xuesong Gao, and Leining LV. 2019. What motivates Chinese university students to learn Japanese? Understanding their motivation in terms of 'posture'. *The Modern Language Journal* 103(1):327-342.
- Tsou, Wenli, and Shin-Mei Kao. 2017. Overview of EMI development. In *English as a Medium of Instruction in Higher Education: Implementations and Classroom Practices in Taiwan*, eds. Wenli Tsou, and Shin-Mei Kao, pp.3-18. Singapore: Springer.
- Tsui, Cynthia. 2017. EMI teacher development programs in Taiwan. In *English as a Medium of Instruction in Higher Education: Implementations and Classroom Practices in Taiwan*, eds. Wenli Tsou, and Shin-Mei Kao, pp.165-182. Singapore: Springer.
- Wu, Wen-Shuenn. 2006. Students' attitude toward EMI: Using Chung Hua University as an example. *Journal of English and Foreign Language and Literature* 4(1):67-84.
- Xiao, JianFang, and XuanQi Zou. 2020. Strategy use to overcome difficulties in EMI courses in higher education. *IJIREs* 7(6):503-519.

Ying Ying Wong, Yuan-Min Shen, Cheng-Chun Huang

*Ying Ying Wong
Master Program for Plant Medicine
National Taiwan University
Taipei, Taiwan 106, ROC
r09645004@ntu.edu.tw*

*Yuan-Min Shen
Master Program for Plant Medicine
National Taiwan University
Taipei, Taiwan 106, ROC
shenym@ntu.edu.tw*

*Cheng-Chun Huang
Master Program for Plant Medicine
National Taiwan University
Taipei, Taiwan 106, ROC
r10645004@ntu.edu.tw*

以全英語教學進行大學授課：
通識教育蔬菜病害課程兩年來之經驗

黃楹瑩、沈原民、黃承濬
國立臺灣大學

「蔬菜病害」課程最初為植物醫學碩士課程研究生所設計的英語媒介教學（EMI）課程，現已轉變為適用於大學學生的通識課程，面向具有英語作為第二語言的多樣背景學生。在 2022 年春季和 2023 年春季學期，我們紀錄了該課程的教學和學習經驗。儘管受到 COVID-19 疫情的影響，但學生們對 EMI 學習過程的反饋仍然積極。課程包含了蔬菜種植、農場參訪以及與農民和專家進行面對面交流等實地學習活動，這有助於實現學習目標。根據對本課程 EMI 學習環境困難程度的調查結果，來自不同學院的學生呈現出相似的態度；然而，大多數外國學生在學習 EMI 主題方面不致於遇到困難，而一些臺灣學生則在 EMI 學習過程中面臨挑戰。本研究中有關體驗式學習活動的支持教學實踐可作為未來 EMI 課程設計的參考。

關鍵字：全英語教學、第二語言、農業、體驗式學習、教學實踐研究

Current Issues in L2 Teaching and Research

作者：Stano Kong (江丕賢)、Yu-Hsin Huang (黃郁欣)、Miao-Chin Chiu (邱妙津)
Kai-Lin Wu (吳凱琳)、One-Soon Her (何萬順)

發行人：張富恭

出版者：文鶴出版有限公司

地址：106 台北市金山南路二段 200 號 8 樓

電話：(02) 2393-4497

傳真：(02) 2394-6822

總公司：106 台北市金山南路二段 200 號 8 樓

TEL: (02) 2393-4497 FAX: (02) 2394-6822

各區業務及門市

北區：106 台北市金山南路二段 200 號 8 樓

TEL: (02) 2393-4497 FAX: (02) 2394-3810

南區：802 高雄市同慶路 88 號 2 樓

TEL: (07) 2270-888 FAX: (07) 2270-801

網路書店：<https://www.crane.com.tw>

網路書店客服信箱：cranebookshop@gmail.com

法律顧問：蘇家宏律師 TEL: (02) 2322-3588

劃撥帳號：01079261

戶名：文鶴出版有限公司

出版日期：2024 年 12 月初版一刷

I S B N : 978-626-7239-67-4

※ 著作權所有，翻印必究，非經本公司同意不得轉載、仿製或局部使用於雜誌、網路或其他書籍、媒體上。

※ 若有缺頁，裝訂錯誤，煩請寄回更換。