The acquisition of definiteness expressions in non-native Japanese by English- and Korean-speaking learners

Keisuke Kume

PhD

University of York

Education

May 2021

Abstract

This thesis investigates learnability problems and the role of L1 in L2 syntax-semantics mappings, through a comparison of English- and Korean-speaking Japanese learners' acquisition of two Japanese definiteness properties: overt definiteness marking by the demonstrative *sono* and covert definiteness distinction through word order change between numeral quantifier constructions. The main objective is to examine distinct predictions based on recent accounts of learnability and L1 influence: the Feature Reassembly Hypothesis (FRH) (Lardiere, 2008, 2009) and the cline of difficulty in feature acquisition (Cho & Slabakova, 2014; Slabakova, 2009). The FRH predicts that both definiteness properties will be acquired faster by Korean speakers due to positive transfer of the L1 corresponding properties. On the other hand, according to the cline of difficulty, the overt property will be acquired before the covert property, irrespective of the L1; and the L1 advantage of Korean is expected with the overt property but not with the covert property. Specifically, the corresponding covert L1 definiteness property (Korean numeral constructions) will not facilitate the acquisition of the same property in the L2, but the L1 functional morphology distinguishing definiteness overtly (English articles) will, hence English speakers will acquire the covert property faster than Korean speakers. These predictions were tested by means of an acceptability judgement task and a self-paced reading task. Although the self-paced reading data offer no evidence of the relevant knowledge even with native Japanese controls, the judgement data indicate that whereas both L2 groups can acquire the overt property equally well, the Korean group tend to exhibit more consistent target-like performance with the covert property than the English group. Based on these results, it is proposed that the necessity of feature reassembly (reconciliation of L1-L2 differences within overt/covert category) plays a bigger role

than overt vs. covert feature realisation, contra the cline of difficulty.

Table of contents

Abstract 2

Table of contents 4

List of tables 11

List of figures 16

Acknowledgements 19

Author's declaration 22

Chapter 1: Introduction 23

1.1 Aims and objectives of the thesis 23

1.2 Learning problems in L2 syntax-semantics mappings 24

1.3 Organisation of the thesis 29

Chapter 2: Theories of the L2 acquisition of syntax-semantics interface properties and learnability problems 31

2.1 Introduction 31

2.2 L2 acquisition at the syntax-semantics interface 32

2.2.1 Syntax-semantics mismatches 32

2.2.2 Poverty of stimulus at the syntax-semantics interface 36

2.2.3 Main findings of L2 research on the syntax-semantics interface 42

2.3 Feature-based contrastive approaches to L2 acquisition 44

2.3.1 Current generative approaches to learnability problems in L2 acquisition 44

2.3.2 The Feature Reassembly Hypothesis 45

2.3.3 Previous studies testing the FRH 48

2.3.4 Cline of difficulty in feature acquisition 49

2.3.5 Previous studies testing the cline of difficulty 56

2.4 Remaining questions about feature contrastive approaches to L2 acquisition 63

Chapter 3: Semantics of definiteness and its expressions in English, Japanese, and Korean 66
3.1 Introduction 66
3.2 What is definiteness? 66
3.3 Two types of definiteness: anaphoricity and uniqueness 69
3.4 Crosslinguistic expressions of definiteness (English, Japanese, and Korean) . 73
3.4.1 Definiteness in English 73
3.4.2 Definiteness in article-less languages, Japanese and Korean 75
3.5 Overt marking of anaphoric definiteness by Japanese *sono* (in comparison to Korean *ku* and English *the* and *that*) 79
3.6 Covert marking of definiteness: Japanese numeral quantifiers (NQs) and semantic constraint in comparison with the Korean and English NQs 85
3.6.1 Numeral classifier constructions and their essential properties 85
3.6.2 Definiteness constraint on NQs 91
3.6.2.1 Prediction A: movement analysis 95
3.6.2.2 Prediction B: size analysis 100
3.6.2.3 Prediction C: floating-NQs-as-adverbial analysis 102
3.6.2.4 English NQ construction 106
3.6.2.5 Which account is the most plausible? 106
3.7 Summary: properties of overt and covert realisations of definiteness in Japanese by the demonstrative *sono* and numeral classifier constructions 108
Chapter 4: Definiteness in second language 112
4.1 Introduction 112
4.2 L2 acquisition of English article semantics 114
4.2.1 Role of semantic universals in L2 article choice 114
4.2.2 Influence of L1 demonstratives on L2 article semantics 124

4.2.3 L2 acquisition of definiteness expressions within the Feature Reassembly Hypothesis 128
4.3 Definiteness in article-less L2s: influence of L1 article semantics on L2 demonstratives 139
4.4 Discussion 141
4.4.1 L2 learnability of bridging reference and L1 transfer 141
4.4.2 Semantics of different definiteness 144
4.5 Implications for the present study 147
Chapter 5: The present study: acquisition tasks and predictions 150
5.1 Introduction 150
5.2 Overt definiteness marking with the demonstrative *sono* 152
5.2.1 Acquisition tasks 152
5.2.2 Relative ease/difficulty and acquisition order 156
5.3 Covert definiteness marking with numeral quantifier constructions 159
5.3.1 Acquisition tasks 159
5.3.2 Relative ease/difficulty and acquisition order 161
5.4 Potential confounding factors 163
5.5 Conclusion: overall predictions 164
Chapter 6: Experiment design and preliminary studies with native Japanese speakers 167
6.1 Introduction 167
6.2 Rationales for combining online and offline methods 169
6.3 Pilot AJT ver. 1 171
6.3.1 Method 172
6.3.1.1 Participants 172
6.3.1.2 Test instrument 172
6.3.1.3 Procedure 175

6.3.2 Results 176
6.3.3 Discussion 178
6.4 Pilot AJT ver. 2 181
6.4.1 Method 181
6.4.1.1 Participants 181
6.4.1.2 Test instrument 182
6.4.1.3 Procedure 185
6.4.2 Results 185
6.4.2.1 *Sono* 185
6.4.2.2 NQ constructions 187
6.4.3 Discussion 189
6.5 Pilot SPRTs ver. 1 & ver. 2 190
6.5.1 Key assumptions in self-paced reading 191
6.5.2 Japanese orthography 192
6.5.3 Processing of *sono* 193
6.5.4 Processing of NQ constructions 193
6.5.5 Method 195
6.5.5.1 Participants 195
6.5.5.2 Test instruments 196
6.5.5.3 Procedure 204
6.5.6 Results 205
6.5.6.1 SPRT data preparation 205
6.5.6.2 Summary of results 207
6.6 General discussion 208
6.6.1 Key findings of the pilot studies 208
6.6.2 Further revisions for main studies 210
6.6.3 Implications for acquisition tasks and learnability predictions 211

Chapter 7: Main study 1 (non-web based SPRT and AJT with L2 Japanese learners) 213
7.1 Introduction....................213
7.2 Method....................213
7.2.1 Overall design....................213
7.2.2 Participants....................214
7.2.3 Test instruments....................215
7.2.3.1 Proficiency test....................215
7.2.3.2 SPRT....................217
7.2.3.3 AJT....................220
7.2.4 Procedure....................222
7.3 Results....................222
7.3.1 SPRT results....................222
7.3.1.1 SPRT data preparation....................222
7.3.1.2 *Sono*: SPRT results....................224
7.3.1.3 NQ constructions: SPRT results....................238
7.3.2 AJT results....................246
7.3.2.1 *Sono*: AJT results....................246
7.3.2.2 NQ constructions: AJT results....................254
7.4 Discussion....................261
7.4.1 Summary of the results....................261
7.4.2 Comparisons with the pilot studies....................262
7.4.3 Potential causes for the difference between the pilot and main AJTs...265
7.4.4 Interpretation of L2 data....................268
7.5 Conclusion....................270
Chapter 8: Main study 2 (web-based AJT with L2 Japanese learners)....................271
8.1 Introduction....................271

8.2 Method 272
8.2.1 Participants 272
8.2.2 Test instruments 273
8.2.2.1 Proficiency test 273
8.2.2.2 AJT 274
8.2.3 Procedure 277
8.3 Results 277
8.3.1 *Sono* 277
8.3.2 NQ constructions 285
8.4 Discussion 291
8.4.1 Comparisons with Main study 1 and Pilot ver. 2 291
8.4.2 Evaluation of the predictions for L2 acquisition of the target properties 295
8.5 Conclusion 296
Chapter 9: General discussion 298
9.1 Introduction 298
9.2 Summary of the main findings 298
9.3 Implications for the cline of difficulty 303
9.4 Acquisition mechanisms 305
9.4.1 Developmental stages of the L2 acquisition of *sono* by L1-English learners 306
9.4.2 How L1-English learners could overcome the poverty of stimulus regarding the definiteness constraint on NQs 308
9.5 Implications for research on L2 acquisition of definiteness expressions 311
9.6 Summary of the chapter 314
Chapter 10: Conclusion 315
10.1 Contributions 315

10.2 Limitations & directions for future research 316

10.3 A final remark 319

Appendices 320

Appendix 1: Consent form & information sheet 320

Appendix 1A: Sample consent form (English ver.): 320

Appendix 1B: Sample information sheet (English ver.): 321

Appendix 2: Language background questionnaire 324

Appendix 3: Proficiency task 327

Appendix 3A: Proficiency test & choices with correct answers 327

Appendix 3B: English translation of the proficiency test passage: 330

Appendix 4: Pilot data 331

Appendix 4A: Pilot AJT ver. 1 331

Appendix 4B: Pilot SPRT ver. 1 331

Appendix 4C: Pilot SPRT ver. 2 333

Appendix 5: Test items for the main AJT and SPRT 336

Abbreviations 380

References 385

List of tables

Chapter 2

Table 2.1 Syntax-semantics mappings in English and Spanish plural NPs in subject position ... 33

Table 2.2 Syntax-semantics mappings in a POS situation ... 36

Table 2.3 Lexical features of English and Mandarin plural markers ... 47

Table 2.4 Crosslinguistic variation in kind-referring count NPs in the four languages tested by Köylü (2019) ... 60

Chapter 3

Table 3.1 Comparison of demonstratives in Japanese, Korean, and English ... 80

Table 3.2 Crosslinguistic differences between English determiners and Japanese and Korean demonstratives ... 84

Table 3.3 Three different predictions for interpretation of post-nominal and floating NQs ... 93

Chapter 4

Table 4.1 Predictions for L2 English article choice (based on Ionin et al., 2004, pp. 18–19) ... 117

Chapter 5

Table 5.1 Crosslinguistic differences between English determiners and Japanese and Korean demonstratives ... 152

Table 5.2 Summary of acquisition scenarios for each L1 group ... 157

Table 5.3 Summary of predictions for each research question (RQ) ... 166

Chapter 6

Table 6.1 Acceptability ratings for NQ construction items: Pilot AJT ver. 1 (SD in

parentheses)..176
Table 6.2 Fixed effect estimates of critical linear mixed-effects models and likelihood-ratio testing results of critical vs. null model comparisons for NQ items: Pilot AJT ver. 1 178
Table 6.3 Mean acceptability ratings for *sono* items: Pilot AJT ver. 2 (SD in parentheses)..186
Table 6.4 Fixed effect estimates of critical linear mixed-effects models and likelihood-ratio testing results of critical vs. null model comparisons for *sono* items: Pilot AJT ver. 2 187
Table 6.5 Mean acceptability ratings for NQ construction items: Pilot AJT ver. 2 (SD in parentheses)..188
Table 6.6 Fixed effect estimates of critical linear mixed-effects models and likelihood-ratio testing results of critical vs. null model comparisons for NQ construction items: Pilot AJT ver. 2 189
<u>Chapter 7</u>
Table 7.1 Participant information: Main study 1 215
Table 7.2 Scores on cloze test (0–42 points): Main study 1....................................216
Table 7.3 Amount of data affected 223
Table 7.4 *Sono*: Mean log-transformed RRTs: native controls (SD in parentheses) 225
Table 7.5 *Sono*: mean log-transformed RRTs: Korean speakers (SD in parentheses) 226
Table 7.6 *Sono*: mean log-transformed RRTs: English speakers (SD in parentheses) 227
Table 7.7 Fixed effects estimates for linear mixed effects model of log-transformed RRTs for *sono* items in the critical region...231

Table 7.8 Fixed effects estimates for linear mixed effects model of log-transformed RRTs for *sono* items in the spillover 1 region..234
Table 7.9 Fixed effects estimates for nested linear mixed effects model of log-transformed RRTs for *sono* items in the spillover 1 region: effects of NP type within each context...235
Table 7.10 Fixed effects estimates for linear mixed effects model of log-transformed RRTs for *sono* items in the spillover 2 region......................................236
Table 7.11 Results of nested models for RRTs for each context for *sono* items in the spillover 2 region...237
Table 7.12 NQ constructions: mean log-transformed RRTs: native controls (SD in parentheses)..238
Table 7.13 NQ constructions: mean log-transformed RRTs : Korean speakers (SD in parentheses)..239
Table 7.14 NQ constructions: mean log-transformed RRTs: English speakers (SD in parentheses)..239
Table 7.15 Fixed effects estimates for linear mixed effects model of log-transformed RRTs for NQ construction items in the critical region..........................243
Table 7.16 Fixed effects estimates for linear mixed effects model of log-transformed RRTs for NQ construction items in the spillover 1 region: NQ constructions 244
Table 7.17 Fixed effects estimates for linear mixed effects model of log-transformed RRTs for NQ construction items in the spillover 2 region....................245
Table 7.18 Mean acceptability ratings by context and by group: Main study 1 *sono* items (SD in parentheses)..246
Table 7.19 Results of the omnibus ordinal model for acceptability ratings for *sono* items 250

Table 7.20 Results of the nested ordinal model for acceptability ratings : the effect of NP type by context 252

Table 7.21 Results of the nested ordinal model for acceptability ratings within each group for the anaphoric context 252

Table 7.22 Distribution of response patterns within each group, in the anaphoric context 253

Table 7.23 Mean acceptability ratings by context and by group: Main study 1 NQ construction items (SD in parentheses)......254

Table 7.24 Results of the omnibus ordinal model for acceptability ratings for NQ construction items 258

Table 7.25 Results of the nested [+definite] ordinal model for acceptability ratings 259

Table 7.26 Results of the L2-only ordinal model for acceptability ratings for NQ construction items with proficiency effect considered......260

Table 7.27 Fixed effects estimates for linear mixed effects model of log-transformed RRTs for items regarding noun-classifier agreement and the interpretive contrast of proper nouns with *-tati* 263

Table 7.28 Comparisons of native Japanese ratings in the target contexts between Pilot AJT ver. 2 and Main study 1 AJT (SD in parentheses)......264

Table 7.29 Comparisons of native Japanese ratings in the control fillers between Pilot AJT ver. 2 and Main study 1 AJT (SD in parentheses)......265

<u>Chapter 8</u>

Table 8.1 Participant information: Main study 2 272

Table 8.2 Scores on cloze test (0–42 points): Main study 2......273

Table 8.3 Mean acceptability ratings by context and by group: Main study 2 *sono* items (SD in parentheses)......278

Table 8.4 Results of the omnibus ordinal model for acceptability ratings for *sono* items 283
Table 8.5 Results of nested models for acceptability ratings for each context of *sono* items 284
Table 8.6 Mean acceptability ratings by context and by group: Main study 2 NQ construction items (SD in parentheses) 285
Table 8.7 Results of the omnibus ordinal model for acceptability ratings for NQ construction items (SD in parentheses) 289
Table 8.8 Results of separate nested ordinal models for acceptability ratings for each condition of NQ construction items 290
Table 8.9 Distribution of response patterns within each group in the [+definite] context 290
Table 8.10 Comparisons of native Japanese acceptability ratings in the target contexts between Pilot ver. 2, Main studies 1 and 2 (SD in parentheses) 293
Table 8.11 Comparisons of native Japanese acceptability ratings in the control fillers between Pilot ver. 2, Main studies 1 and 2 (SD in parentheses) 294

<u>Chapter 9</u>

Table 9.1 Assessment of predictions for each approach 303
Table 9.2 Possible acquisition scenarios for L1-English learners 308

List of figures

Chapter 1

Figure 1.1 Cline of difficulty in feature acquisition (adapted from Cho & Slabakova, 2014, p. 166) 26

Chapter 2

Figure 2.1 Slabakova's (2009) cline of difficulty in feature acquisition........ 50

Figure 2.2 Extended cline of difficulty in feature acquisition (adapted from Cho & Slabakova 2014, p. 166)........ 52

Figure 2.3 Predicted difficulty of acquisition of [+familiar] and [+unique] features in L2 English by L1- Mandarin and Russian speakers (adapted from Tuniyan, 2018, p. 139) 57

Chapter 5

Figure 5.1 Cline of difficulty in feature acquisition (adapted from Cho & Slabakova 2014, p. 166) 162

Chapter 6

Figure 6.1 Mean ratings for NQ constructions items (error bars = SE)........ 176

Figure 6.2 Mean ratings for *sono* items (error bars = SE) 186

Figure 6.3 Mean ratings for NQ construction items (error bars = SE) 188

Chapter 7

Figure 7.1 Histogram of the trimmed RT data 224

Figure 7.2 Mean RRTs (log) for *sono* items in the anaphoric context (error bars = SE) 228

Figure 7.3 Mean RRTs (log) for *sono* items in the non-anaphoric context (error bars = SE) 228

Figure 7.4 Mean RRTs (log) for *sono* items in the bridging context (error bars = SE)

....229
Figure 7.5 Mean RRTs (log) for NQ construction items in the [+definite] context
(error bars = SE)..240
Figure 7.6 Mean RRTs (log) for NQ construction items in the [−definite] context
(error bars = SE)..240
Figure 7.7 Mean acceptability ratings for *sono* items in the anaphoric context (error
bars = SE)..247
Figure 7.8 Mean acceptability ratings for *sono* items in the non-anaphoric context
(error bars = SE)..247
Figure 7.9 Mean acceptability ratings for *sono* items in the bridging context (error
bars = SE)..248
Figure 7.10 Mean acceptability ratings for NQ construction items in the [+definite]
context (error bars = SE)..254
Figure 7.11 Mean acceptability ratings for NQ construction items in the [−definite]
context (error bars = SE)..255
Chapter 8
Figure 8.1 Mean acceptability ratings in the anaphoric context (error bars = SE) ..278
Figure 8.2 Mean acceptability ratings in the non-anaphoric context (error bars = SE)
....279
Figure 8.3 Mean acceptability ratings in the bridging context (error bars = SE) 279
Figure 8.4 Mean acceptability ratings in the [+definite] context (error bars = SE).285
Figure 8.5 Mean acceptability ratings in the [−definite] context (error bars = SE).286
Chapter 9
Figure 9.1 Revised Cline of Difficulty...304
Appendices
Figure Appx.1 Mean acceptability ratings for *sono* items: Pilot AJT ver. 1 (error bars

= SE) 331
Figure Appx.2 Mean RRTs for *sono* items in the anaphoric context: Pilot SPRT ver. 1 (error bars = SE)..331
Figure Appx.3 Mean RRTs for *sono* items in the non-anaphoric context: Pilot SPRT ver.1 (error bars = SE)..332
Figure Appx.4 Mean RRTs for *sono* items in the bridging context: Pilot SPRT ver. 1 (error bars = SE)..332
Figure Appx.5 Mean RRTs for NQ constructions items in the [+definite] context: Pilot SPRT ver. 1 (error bars = SE)...332
Figure Appx.6 Mean RRTs for NQ constructions items in the [−definite] context: Pilot SPRT ver. 1 (error bars = SE)...333
Figure Appx.7 Mean RRTs for *sono* items in the anaphoric context: Pilot SPRT ver. 2 (error bars = SE)..333
Figure Appx.8 Mean RRTs for *sono* items in the non-anaphoric context: Pilot SPRT ver. 2 (error bars = SE)..334
Figure Appx.9 Mean RRTs for *sono* items in the bridging context: Pilot SPRT ver. 2 (error bars = SE)..334
Figure Appx.10 Mean RRTs for NQ construction items in the [+definite] context: Pilot SPRT ver. 2 (error bars = SE)...334
Figure Appx.11 Mean RRTs for NQ construction items in the [−definite] context: Pilot SPRT ver. 2 (error bars = SE)...335
Figure Appx.12 Mean RRTs for noun-classifier agreement items: Pilot SPRT ver. 2 (error bars = SE)..335
Figure Appx.13 Mean RRTs for items regarding collective reading of proper names with -*tati*: Pilot SPRT ver. 2 (error bars = SE)..............................335

Acknowledgements

I wish to thank everyone whose assistance was a milestone in the completion of this thesis.

First and foremost, I would like to express my deepest gratitude to my supervisor, Heather Marsden for her continuous academic and pastoral support throughout my PhD studies. In difficult times when things appeared to be getting out of hand, her insightful guidance never failed to put me back on track. I was emotionally saved countless times by her unwavering confidence in me. It is wholeheartedly appreciated that she never gave up on me when I almost did. I was really fortunate to have Leah Roberts as my TAP member and I am truly grateful to her for invaluable advice and heartfelt encouragement at TAP meetings, and for kindly agreeing to be the internal examiner.

Special thanks go to my supervisors and Linguistics teachers when I was an undergraduate at Chuo University and a graduate at Nanzan University, especially Ryoji Kasai, Tatsuya Suzuki, and Hiroshi Aoyagi. Without their education, guidance and encouragement, my PhD journey would not have started in the first place. I would also like to pay my special regards to Hokari Tomohiro and Mutsumi Ogawa for their help with my PhD application, particularly for the way they used their own expertise to help shape my research proposal.

My PhD studies at the University of York would have not been possible without the scholarship from the Japan Student Services Organization (Graduate Scholarship for Degree-Seeking Students). I am indebted to my sponsor for funding my PhD abroad.

Appreciation is due to my native Japanese consultants, particularly, Kazuki Inoue and Nanami Kawade. They both checked all the pilot versions of the test instruments and provided me with helpful feedback from a layman's point of view, which allowed me to improve the quality of the instruments significantly. I was always amazed by their very keen metalinguistic awareness of the Japanese language. Further thanks go to Samuel Lynas, Harry Carlton and James Forsythe, for trying out the test materials as Japanese learners. Their insightful feedback helped me fine-tune the materials before the actual data collection. I would also like to thank Kook-Hee Gil for kindly translating the consent form and the information sheet into Korean for me.

Needless to say, immense thanks are due to all the participants in my studies. Without them, this thesis would never have been completed. Further thanks are due to the following organisations and individuals for their help with participant recruitment (in no particular order): York Japanese Families' Association, University of York Japanese Society, York St. John University Japanese Society, Durham University Anglo-Japanese Society, Becky Muradás-Taylor, Megumi Bailey, Yumi Nixon, Kazuki Morimoto, Kumi Casey, Mary Murata, Mikiko Inoue, Kaori Nishizawa, Sakie Chiba, Tomoko Miyairi, Ritsuko Koso-Kirk, Kenichi Murakami, Shigenori Wakabayashi, Takayuki Kimura, Kazunori Suzuki, Shunji Inagaki, Mitsuru Doi, Tatsuya Suzuki, Hiroshi Aoyagi, Tack ung Lee, Robert Croker, and the staff of the International Centre at Chuo University (Yumiko Watanabe in particular).

I wish to show my sincere gratitude to the Faculty of Letters in Chuo University for welcoming me as a visiting scholar during my data collection in Japan, for providing a perfect venue for the experiments, and for offering me an opportunity to give a talk on my research.

I would like to thank Takanobu Nakamura for sharing his wealth of knowledge on Japanese numeral quantifiers, which greatly helped me grasp numerous issues pertaining to Japanese numerals. I also acknowledge the contribution of the Japanese language instructors who answered my questions about Japanese language education, particularly, Megumi Bailey, Yumi Nixon, Aki Ito, and Ikue Saito. Additionally, I wish to thank Shayne Sloggett for his helpful advice on statistical analysis.

I would like to extend my deepest gratitude to Shigenori Wakabayashi for his invaluable advice on academic and personal life. I was very fortunate to spend a lot of time with him in person over my PhD years in the UK, during his stay at Cambridge University as a visiting scholar and his research data collection at the University of York. In particular, the road trip he took my family and me on will forever be one of our fondest memories of our time in England.

I would like to express my heartfelt appreciation to my family members: my wife, Sayaka and my children, Yukina and Ryotaro for moving to York together with me for my PhD studies. In particular, Sayaka supported me with her self-less dedication to our family, taking on the bulk of the childcare responsibilities, to allow me to focus on my PhD project. I am also thankful to my parents, Yuya and Noriko, and my grandmother, Kazue, who supported me from Japan, both financially and through their love and patience.

Finally, I would like to dedicate this thesis to my late grandfather, Manjiro, who passed away during my PhD years. He was one of the most intelligent, hard-working, and gentle people that I have ever known. I still miss him greatly and cannot help regretting that I could not go back to Japan while he was still alive. Nonetheless, I hope that he will be proud of me for completing my PhD with the help of all the people above.

Author's declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Chapter 1: Introduction

1.1 Aims and objectives of the thesis

This thesis investigates learnability problems and the role of first language (L1) in the non-native language acquisition of syntax-semantics mappings. The specific objective is to examine the acquisition of definiteness in Japanese as a second language (L2). The L2 acquisition of definiteness has been widely investigated via research into the acquisition of English articles (e.g., Ionin, Ko, & Wexler, 2004; Ionin, Zubizarreta, & Bautista-Maldonaldo. 2008; Ko, Ionin, & Wexler, 2010; Snape, 2008; Thomas, 1989; Trenkic 2007). However, little attention has been paid to definiteness in article-less languages such Japanese and Korean. The present study focuses on two distinctive definiteness-related properties in Japanese: definiteness-marking by the demonstrative, *sono*, and an interpretive contrast between numeral quantifier (NQ) constructions. Japanese learners from two different language backgrounds, namely native speakers of English (i.e., a language with articles) and Korean (i.e., a language without articles) are compared. These linguistic properties and L1-L2 combination were selected in order to empirically test predictions based on recent theoretical accounts of L2 learnability problems: the Feature Reassembly Hypothesis (Lardiere, 2008, 2009) and the cline of difficulty in feature acquisition (Cho & Slabakova, 2014; Slabakova 2009). I implement two experimental data collection methods of different nature, namely an acceptability judgement task (AJT) and a self-paced reading task (SPRT) in order to elicit learners' linguistic knowledge of the target properties. Findings of the study will provide new insights into learnability problems in L2 syntax-semantics mapping and advance our understanding of the nature of L2 knowledge.

1.2 Learning problems in L2 syntax-semantics mappings

The L2 acquisition of syntax-semantics correspondences has been widely investigated with various linguistic phenomena with different L1-L2 combinations and argued to be nontrivial (for a comprehensive review, see Slabakova, 2008). A vast number of researchers have investigated syntax-semantics mismatches, in which universal meanings (e.g., past tense, number, definiteness) are expressed through different functional morphemes in the L1 and L2. Previous studies revealed that mismatches at the syntax-semantics interface could be overcome even for relatively low-proficient learners (e.g., Montrul & Slabakova, 2003; Slabakova, 2003). Moreover, there has been robust evidence that successful L2 syntax-semantics mappings are possible even in poverty-of-the-stimulus situations, where evidence required to learn a given linguistic property is unavailable in the linguistic input (i.e., natural input and classroom instruction) (e.g., Dekydtspotter & Sprouse, 2001; Dekydtspotter, Sprouse, & Anderson, 1997; Marsden, 2008, 2009; Unsworth, 2005). However, a number of studies have found that syntax-semantics interface properties might remain problematic even for advanced learners presumably due to L1 transfer and a scarcity of evidence to motivate changing the L1-based structure (e.g., Gabriele, 2009, 2010; Yuan, 2010). Therefore, although the L2 acquisition of the syntax-semantics interface phenomena appears to be relatively successful and ultimately attainable, further research is necessary to find what conditions make L2 syntax-semantics mappings relatively easier or more difficult (White, 2018).

The Feature Reassembly Hypothesis (FRH) (Lardiere, 2008, 2009) offers a framework for investigating this issue. According to the FRH, the main L2 acquisition task is to (re)assemble formal L2 features in the following two steps. The first is to map features that are assembled on L1 lexical items onto perceived L2 equivalents. The second process involves reconfiguration of feature sets into the way they are represented in the

L2. This reconfiguration task will be successful only if motivated by evidence in the input. Therefore, the general predictions of the FRH for ease or difficulty of acquisition are as follows: with a mismatch between the L1 and the L2 in terms of how features are bundled into lexical items (e.g., a certain feature expressed through different classes of lexical items or combine with different features), acquisition will be harder than a case without such a mismatch due to feature reassembly necessity; and when feature reassembly is needed, the task will be easier when triggering evidence is readily available in the input than when such evidence is scarce or non-existent. A number of recent studies have tested these predictions, yielding supporting evidence (e.g., Cho & Slabakova, 2015; Gil & Marsden, 2013; Hwang & Lardiere, 2013; Su, 2019; Yuan, 2014).

With a view to refining these broad predictions of the FRH, Cho and Slabakova (2014), building on Slabakova (2009), proposed a model that incorporates whether the relevant feature expression is overt or covert. The idea is that features expressed overtly (i.e., with a dedicated morpheme) will be easier to acquire than those represented covertly (i.e., by a word order change, or context) by virtue of the more consistent evidence available for overtly realised features whether in the L1 or the L2. From this perspective, overt realisation of a feature in the L1 is expected to aid the learner in detecting that feature in the L2 more than covert realization would. Combining this with the assumption under the FRH that L2 acquisition is more complicated when feature reassembly is required than when it is not, Cho and Slabakova put forth a cline of difficulty in feature acquisition, as given in Figure 1.1.

Easier to acquire					Harder to acquire
F_{overt} to F_{overt}	F_{overt} to F_{overt}	F_{covert} to F_{overt}	F_{overt} to F_{covert}	F_{covert} to F_{covert}	F_{covert} to F_{covert}
reassembly not required	reassembly required			reassembly not required	reassembly required

Figure 1.1. Cline of difficulty in feature acquisition (adapted from Cho & Slabakova, 2014, p. 166)

On the far-left end of the cline is the easiest task, where a given property is expressed overtly with functional morphemes both in the L1 and L2 and those morphemes have identical feature configurations, thus no reassembly is required. The most difficult task concerns features covertly expressed both in the L1 and L2 but in different ways, thus reassembly is required. Cho and Slabakova found quantitative evidence for the cline from an investigation of definiteness in L2 Russian by L1 speakers of Korean and English. These three languages have overt expressions of definiteness: English overtly realises definiteness with articles; whereas Russian adjectival possessors and Korean case-markers within noun-noun compounds do so as a secondary function of the relevant morphemes. Moreover, Russian and Korean have covert expressions of definiteness by means of word order, although the conditions (i.e., feature configurations) for feature realisation do not match. The results of Cho and Slabakova's felicity judgement task revealed that both the Korean- and English-speaking learners had successfully acquired the Russian overt realization of definiteness through adjectival possessors. However, target-like knowledge of the covert definiteness expression through word order in Russian was observed only with some advanced English-speaking learners but none of the Korean-speaking learners. Cho and Slabakova accounted for the difference between the two groups as a reflection of the greater complexity of the acquisition task facing the Korean speakers: feature reassembly where the relevant feature is realised covertly both in the L1 and the L2 but differently (i.e., the hardest task). They speculated that English speakers, tackling the third hardest task, might have benefited from the overt realization of definiteness in the L2 as a heuristic in acquiring the L2 covert expression.

Cho and Slabakova's (2014) core assumption that features expressed covertly are harder to acquire than those expressed overtly has been partially supported by some recent studies investigating the L2 acquisition of English articles (Tuniyan, 2018) and English kind reference (Köylü, 2019), but the results seem inconsistent (discussed in Chapter 2). Thus, the cline of difficulty requires further empirical validation through testing with a variety of L1-L2 combinations and linguistic properties. Particularly, to my knowledge, the acquisition of a covert feature realised in the same way in the L1 and L2 has not been investigated yet (F_{covert} *to* F_{covert} *no reassembly required* in Figure 1.1). Cho and Slabakova rank this task as the second hardest on the grounds that covert properties are harder to acquire than overt properties and that covert feature realisations in the L1 do not facilitate L2 acquisition as much as overt counterparts. However, this prediction is not uncontroversial particularly because there is no a priori reason for the advantage of no feature reassembly no longer applying in the acquisition of covert feature expressions. Whether this prediction is correct or not remains an open question, which the present study addresses through an investigation of the L2 acquisition of a covert Japanese expression of definiteness by English- and Korean-speaking learners. The acquisition of an overt definiteness property in Japanese is also investigated in order to examine the role of overt vs. covert feature realization more comprehensively. The following research questions are addressed in this thesis.

(1.1) Research Question 1:

Is a covert feature expression more difficult to acquire than an overt feature expression?

(1.2) Research Question 2:

Does the necessity of feature reassembly make the acquisition task more difficult?

(1.3) Research Question 3:

In which situation is the acquisition of a covert feature expression less difficult, (i) when the L1 has a functional morpheme that realizes the feature overtly or (ii) when the L1 has a corresponding covert expression?

In Japanese, NQs, which consist of a numeral and a classifier appropriate for the associated noun, can occur in multiple syntactic positions. However, their interpretation is constrained by the interaction of word order and semantics in terms of definiteness (e.g., Furuya, 2012; Watanabe, 2006): NQs that immediately follow the associated noun can be either definite ([+definite]) or indefinite ([−definite]), whereas those that are separated from the associated noun, called *floating* NQs, can be [−definite] only. This counts as a case of covertly realised definiteness. Korean possesses a corresponding property (Lee, 2013; Shin, 2017) but English does not have floating NQs (e.g., Kobuchi-Philip, 2007) or the semantic constraint. Furthermore, definiteness is realised overtly in English, predominantly through its article system, whereas in Japanese and Korean, definiteness is predominantly expressed covertly. These crosslinguistic differences allow us to investigate the acquisition of the second hardest task on the cline in comparison with the third hardest. The cline of difficulty predicts that the definiteness constraint on Japanese floating NQs will be easier for English-speaking learners, who may benefit from the consistently overt realisation of definiteness in the L1, than for Korean counterparts, whose L1 expresses definiteness predominately covertly. On the other hand, despite the absence of an article system, Japanese can express definiteness overtly by means of the demonstrative *sono*, similarly to the English definite article *the* in some contexts (i.e., anaphoric definite contexts) (e.g., Hoji, Kinsui, Takubo, & Ueyama, 2003; Yoshida, 2011). Korean also has a demonstrative, *ku*, which overtly expresses definiteness in the same way as Japanese *sono* (e.g., Ahn, 2017; Cho, 2017). *Sono* and *ku* share some functions with the English article *the* and the English demonstrative *that* but behave differently in other respects (discussed in Chapter 3). According to the cline of difficulty, the overt expression of definiteness by means of the demonstrative *sono* will be easier to acquire than the covert definiteness property of NQs, irrespective of the L1. However, L1-Korean learners are predicted to acquire the overt property more easily than L1-English learners due to the crosslinguistic differences between the two L1s: Korean-speaking learners do not have to undertake feature reassembly to acquire the definiteness function of *sono*, whereas English-speaking learners do. The thesis presents an experimental investigation of these predictions.

1.3 Organisation of the thesis

This thesis is organized as follows. Chapter 2 reviews previous findings of research into the L2 acquisition of syntax-semantics phenomena and provides theoretical background on the feature reassembly approach to this domain of inquiry. Chapter 3 first introduces the semantics of definiteness along with a brief description of how definiteness is expressed in Japanese, Korean and English. Then the Japanese definiteness properties in question are detailed in comparison with related Korean and English properties, highlighting similarities and differences between the three languages. Chapter 4 provides an overview of L2 acquisition research into definiteness phenomena along with discussions of outstanding issues and their implications for the present thesis. Chapter 5 identifies acquisition tasks and formulates predictions for the research questions. Chapter 6 illustrates the methodology for the experimental research and reports on a series of preliminary studies conducted with native speakers of Japanese. These serve to develop and pilot the test instruments (AJTs and SPRTs), but also, crucially, to provide quantitative validation of the linguistic properties of NQs and *sono* in native Japanese, which have previously been described extensively but not tested experimentally. Chapter 7 presents the first main study with L2 Japanese learners (and native Japanese speakers as a control group) (Main study 1). A follow-up study with other groups from the same target populations (Main study 2) is reported in Chapter 8 which aims to address some methodological issues raised with the results of Main study 1. Chapter 9 summarises the main findings and then discusses their implications. Chapter 10 concludes the thesis by stating its contributions, along with identification of some limitations, and directions for future research.

Chapter 2: Theories of the L2 acquisition of syntax-semantics interface properties and learnability problems

2.1 Introduction

Natural languages are more or less equivalent in their expressive power: no meaning is expressible in one language and not in another (Katz, 1976). In other words, a certain set of semantic primitives is arguably universally available across languages (e.g., Jackendoff, 2002; Ramchand & Svenonius, 2008). However, languages can considerably differ as to how those meanings are expressed. For example, it varies from language to language how meanings are mapped to linguistic forms (i.e., words, phrases, sentences) and whether they are overtly encoded in functional morphology or not (e.g., tense is morphologically marked in Japanese and English but not in Chinese). These crosslinguistic differences at the syntax-semantics interface make form-meaning mappings nontrivial in L2 acquisition and the research in this domain has received a lot of attention.

In what follows, some key observations in the research into L2 syntax-semantics mappings are discussed first while briefly reviewing several representative studies. I will then introduce theoretical frameworks that this study adopts, mainly the Feature Reassembly Hypothesis (Lardiere, 2008, 2009) and the cline of difficulty in feature acquisition (Cho & Slabakova, 2014; Slabakova, 2009) along with previous findings relevant to these proposals. The chapter ends with a summary of outstanding issues within these feature-based approaches to L2 acquisition some of which the present study aims to address.

2.2 L2 acquisition at the syntax-semantics interface

Slabakova (2008) identifies two main types of learning challenges in the recent research on the L2 syntax-semantics interface. One concerns syntax-semantics mismatches, where a universal meaning is mapped onto different forms in the L1 and the L2. The other type involves poverty-of-the-stimulus situations, in which the linguistic input available for the learner is purportedly insufficient for a successful acquisition of a given property. In this section, I will briefly review five representative studies: two for the first type, namely Ionin, Montrul, and Crivos (2013) and Gabriele (2010); and three for the second, namely Unsworth (2005), Marsden (2009), and Okuma (2019), thereby illustrating challenging aspects of the L2 syntax-semantics interface. These studies are chosen for their relevance to the present study as follows. First, Okuma (2019) is the only previous study focusing on the L2 interpretation of Japanese numeral quantifier (NQ) constructions (but an interpretive property different from the one tested in the present thesis). Furthermore, some (Marsden, 2009; Unsworth, 2005) concern the scope of quantifiers and others computation of meanings at the noun phrase and verb phrase levels (Gabriele, 2010; Ionin et.al., 2013), which are properties relevant to those investigated in the present study (discussed in Chapter 3).

2.2.1 Syntax-semantics mismatches

In a typical syntax-semantics mismatch, similar forms and meanings exist in both the L1 and the L2 but are misaligned (Slabakova, 2016, p. 311). For example, consider the interpretation of definite plural noun phrases (NPs) (e.g., *the tigers*) and bare plural NPs (e.g., *tigers*) in English and Spanish as investigated by Ionin et al. (2013). An NP has a generic meaning when it refers to the whole class of individuals denoted by the noun, whereas it has a specific/non-generic reading when it refers to particular things or persons that satisfy the denotation. In Spanish, definite plurals can have both generic and specific readings (2.2), whereas they only have a specific reading in English (2.1). Generic readings can be expressed with bare plurals in English (2.3) but bare plurals are ungrammatical in pre-verbal subject position in Spanish (2.4).

(2.1) <u>The tigers</u> eat meat. * generic reading, ✓specific reading

(2.2) Los tigres comen carne. ✓generic, ✓specific reading

the-PL tigers eat meat

'The tiger ear meat.'

(2.3) Tigers eat meat. ✓generic reading, *specific reading

(2.4) * Tigres comen carne.

Ionin et al. (2013, p. 485)

The form-meaning mismatch between English and Spanish can be summarised as in Table 2.1. The two superficially similar sentences with definite plurals in (2.1) and (2.2) have different meanings: in Spanish, they can refer to both all tigers as a kind and some specific tigers in the context, but their English equivalents only have the latter interpretation.

Table 2.1 Syntax-semantics mappings in English and Spanish plural NPs in subject position

	Generic	Specific
English	bare plurals	definite plurals
Spanish	definite plurals	definite plurals

Ionin et al. (2013) exploited this crosslinguistic difference to investigate bidirectional L2 learnability between the two languages (i.e., L1-English → L2-Spanish and L1-Spanish → L2-English), using a Truth-Value Judgement Task (TVJT) and an Acceptability Judgement Task (AJT). The TVJT was designed to examine which reading (generic or specific) learners prefer for definite plurals in L2-English and Spanish. Participants read stories accompanied by pictures and judged test sentences as true or false. The purpose of the AJT was to test how L2 learners of English and Spanish judge definite and bare plurals in generic vs. specific contexts. In the experiment, participants had to rate each test sentence as a continuation of the preceding story on a 4-point scale. The results suggested that L1 transfer occurred in both directions at lower proficiency levels. In the TVJT, L1-English L2-Spanish learners preferred the specific reading of definite plurals more than native Spanish speakers in contexts that favoured the generic reading, whereas L1-Spanish L2-English learners tended to lean towards the generic reading of definite plurals which is unavailable in English. Moreover, it was found from the AJT that in the generic condition, L1-English L2-Spanish learners rated the definite plurals lower and bare plurals higher than native Spanish speakers, although they were nativelike in the specific condition. Ionin et al. argue that the difficulty in the generic contexts is due to L1 transfer: the mismatch between English and Spanish in form-meaning mapping of genericity. Importantly, however, target-like performance was attested at higher proficiency levels despite the L1 transfer. These findings suggest that learners may be influenced by the L1 in the beginning but with proficiency are able to retreat from it to attain native-like interpretations.

Gabriele (2010) examined another syntax-semantics mismatch involved in L1-English speakers' acquisition of telicity in L2-Japanese. In both English and Japanese, sentences with a mass noun can denote events that are either telic (i.e., with an endpoint) or atelic (i.e., without an endpoint). For example, a drinking-juice event can have either have a clear end point (i.e., telic) or not (i.e., atelic) in English (2.5a) and Japanese (2.5b) alike. That is, the English and Japanese sentences in (2.5) are both possible when Sam drank a specific quantity of juice (e.g., a glass of juice) (telic reading) as well as when the quantity of juice is irrelevant or unspecified (atelic reading).

(2.5) a. Sam drank juice. ✓telic, ✓atelic

b. Sam-wa jyuusu-o nomimasita. ✓telic, ✓atelic

Sam-TOP juice-ACC drank

However, there is a mismatch between these languages in terms of interpretation of count nouns. In Japanese, a language without obligatory singular/plural and count/mass morphology, telicity is largely determined by context. For example, the verb phrase *tegami-o kakimasita* 'wrote letter' in (2.6) can be interpreted as telic 'wrote a/some/the letter(s)' or atelic 'wrote letters'. In English, however, telicity needs to be distinguished morphosyntactically through singular/plural and count/mass markings. For example, verb phrases such as *write a/the letter(s)* denote bounded events (i.e., telic), whereas verb phrases such as *write letters* denote a repeated action without a clear endpoint (i.e., atelic).

(2.6) Sam-wa tegami-o kakimasita.

Sam-TOP letter-ACC wrote

'Sam wrote a/the/some letter(s) on his birthday.'

(Gabriele, 2010, p. 385)

Gabriele (2010) conducted two studies with intermediate and advanced Japanese learners with L1-English in order to examine their interpretation of telicity in their L2-Japanese. Interpretational tasks were administered to participants in which they looked at pictures while listening to short stories in Japanese with two different ending patterns (the event was completed (i.e., telic) or was unfinished (i.e., atelic)), and then judged the compatibility of the test sentence with the preceding context on a 5-point scale. The results showed that many learners in both groups exclusively assigned a telic reading to bare count nouns unlike native Japanese speakers, although their interpretation of bare mass nouns was generally native-like. Gabriele attributed this asymmetry between the noun types to transfer of L1 nominal semantics (i.e., the boundedness of count nouns in English). She proposed that the recovery from negative transfer (i.e., the mismatch in the boundedness of count nouns between English and Japanese) becomes a great challenge to L2 learners when the L1 makes semantic distinctions through morphosyntax but in the L2 the relevant meaning is often determined contextually (p. 402). Nevertheless, there were still learners in each group who exhibited fully target-like interpretation regardless of noun type. Therefore, this property seems ultimately acquirable.

2.2.2 Poverty of stimulus at the syntax-semantics interface

Poverty of the stimulus (POS) is often found at the syntax-semantics interface in situations in which "two related minimally different sentences differ in available interpretations" (Slabakova, 2016, p. 299), as illustrated in Table 2.2.

Table 2.2 Syntax-semantics mappings in a POS situation

	Meaning 1	Meaning 2
Sentence 1	✔	✔
Sentence 2	✔	✘

Knowledge of the lack of meaning 2 for sentence 2 is typically not motivated by information available to learners (i.e., input, L1, or L2 classroom instruction). This can cause a learnability problem because theoretically, absence of Sentence 2 in contexts where meaning 2 is favoured does not provide definitive evidence for the incompatibility of Sentence 2 with meaning 2. To illustrate a case in point, consider a pair of sentences whose interpretative property was investigated by Unsworth (2005). In Dutch, objects can occur in two different positions relative to quantified NPs. In (2.7a), the object noun, *een app* 'a monkey' is in the canonical position. On the other hand, in (2.7b), the object noun is moved over the adverbial quantified NP, *twee keer* 'twice' by the syntactic operation known as scrambling, which results in an interesting interpretive effect.

(2.7) a. Het meisuje heft twee keer een aap gekieteld
The girl has two times a monkey tickled

b. Het meisuje heft een aap$_i$ twee keer e_i gekieteld
The girl has a monkey two times e_i tickled

'The girl has twice tickled a monkey.'

(Adapted from Unsworth 2005, p. 298)

The sentence in (2.7a) has two potential interpretations: (i) the girl tickled two different monkeys or (ii) the same monkey twice, whereas the minimally different sentence in (2.7b) can have the latter interpretation only, with the object, *een aap* taken outside of the scope of the quantified NP, *twee keer*. This interpretive difference is assumed to be very difficult to acquire because learners are not likely to encounter information that scrambled sentences cannot have a different-objects reading, though they will find evidence that both different-objects and same-object readings are possible for non-scrambled sentences and the same-object reading for scrambled sentences, hence a POS situation (non-scrambled and scrambled sentences corresponding to sentences 1 and 2, and same-object and different-objects readings to meaning 1 and 2, respectively, in Table 2.2).

Unsworth (2005) examined this subtle interpretive property in the L2 Dutch of a group of child and adult native English speakers. A TVJT was administered in which participants were presented with stories accompanied with matching pictures and then a puppet described the event depicted in those pictures using sentences with either an unscrambled or scrambled object. Participants had to judge whether what the puppet said was true or false. For analysis, the participants were divided into two groups by proficiency: intermediate and advanced. Unsworth found that in the non-scrambled condition, both proficiency-level groups accepted the two types of interpretation at high accuracy. In the scrambled condition, on the other hand, the intermediate group successfully rejected the different-objects interpretation only about 20 % of the time, whereas the advanced group exhibited target-like judgment at about 80 % accuracy as in the non-scrambled condition. The performance of the advanced learners suggests that the POS property is acquirable presumably because of being part of humans' innate linguistic knowledge (i.e., Universal Grammar, henceforth UG). [1] The generally inaccurate judgement of the intermediate learners means that they might not have yet acquired the target syntactic operation (i.e., scrambling), which does not exist in the L1 (English). However, once they have acquired it, the interpretive property represented by scope relations seems to become automatically activated.

Marsden (2009) investigated a different POS property, namely distributive quantifiers in L2 Japanese of L1-English and L1-Korean learners. In English, sentences with quantifiers such as *every(one)* and *some(one)* suggest that some meanings result from covert movements of the arguments. For example, the sentence in (2.8) has two potential meanings.

(2.8) Someone read every book.

---

[1] For a general overview of what UG means in relation to SLA, the reader is referred to textbooks such as White (2003) and Hawkins (2019).

One meaning is that there was one person, who read multiple books (i.e., one person and many books). Under the other interpretation, the sentence is true when for each of multiple books, there was some person who read that book (i.e., many persons and many books). In the latter, the quantified object (*every book*) is assumed to take wide scope over the subject (*someone*), as a result of a covert syntactic movement. However, in Japanese and Korean, such a covert movement is not available, and Japanese and Korean equivalents to the English sentence in (2.8) lack the object wide scope construal (2.9).

(2.9) a. Japanese:	Dareka-ga	dono	hon-mo	yonda.
b. Korean:	Nwukwunka-ka	enu	chayk-ina	ilkessta.
	Someone-NOM	every	book	read
	‘Someone read every book.’			

Marsden (2009) argues that the acquisition of the absence of object wide interpretation in Japanese SOV sentences constitutes a POS situation for native speakers of English because neither positive input nor the L1 provides reliable evidence for the target knowledge and it is not typically covered in L2 language instruction. By contrast, Korean speakers may acquire this property as a facilitative effect of the L1. Marsden tested knowledge of this property in L2 Japanese learners with L1-English or L1-Korean background by using a sentence-picture rating task. In the experiment, participants were shown pictures either depicting a subject-wide scope situation or an object-wide scope situation. For each trial, the picture was presented alone followed by visual and oral presentation of a test sentence to judge for its compatibility with the picture description. The results confirmed the facilitative L1 effect: the L1-Korean speakers tended to reject object-wide scope whereas the L1-English speakers tended to accept it. However, it was also found from an individual analysis that half of the advanced English-speaking learners consistently rejected the object wide reading of Japanese sentences as in (2.9), despite their less target-like judgement at group level. Marsden took these L1-English learners' successful acquisition as evidence that the universal syntax-semantics computation is operative in their L2 acquisition of Japanese. As to why L1-English learners were slower to acquire the property, she proposed that distinct semantic properties of the L1 and L2 universal quantifiers required English-speaking learners to reconfigure the lexical properties of the L1 quantifier (*every*) to match the target L2 quantifier (*dono…mo*) and this resulted in the delayed acquisition. On the other hand, Korean-speaking learners acquired the target property of the L2 quantifier due to positive transfer of the L1 quantifier (*enu…(i)na*) with the same feature specification.

Finally, Okuma (2019) has recently explored another interesting POS phenomenon in L2 acquisition of Japanese by L1-English speakers. As briefly mentioned in Chapter 1 and detailed in the next chapter, Japanese numerals must combine with semantically matching classifiers and can appear in multiple constructions, including the post-nominal construction (2.10a), where the combination of a numeral and a classifier (henceforth, simply NQ) *ni-ko* 'two-CL' appears between the associated noun *ringo* 'apple' and the case marker *-o* '-ACC'; and the floating construction, where the NQ appears after the case marker (i.e., adverbial position) (2.10b).

(2.10) a. Post-nominal NQ

Taroo-ga	[ringo ni-ko]-o	tabeta.
Taroo-NOM	[apple two-CL]-ACC	ate

'Taroo ate (the) two apples.'

b. Floating NQ

Taroo-ga	[ringo]-o	ni-ko	tabeta.
Taroo-NOM	[apple]-ACC	two-CL	ate

‘Taroo ate two apples.’

Furthermore, it is argued that post-nominal NQs have both distributive and collective interpretations whereas floating NQs cannot have a collective interpretation (Ishii, 1999; Nakanishi, 2007). For example, (2.11a) means either that three students worked individually to each submit a separate piece of homework (i.e., distributive interpretation), or that three students worked together to submit one piece of homework (i.e., collective interpretation). By contrast, (2.11b) only allows the former interpretation.

(2.11) a. [Gakusei san-nin]-ga kyoo syukudai-o dasita.
[student three-CL]-NOM today homework-ACC submitted
‘Three students submitted homework today.’
(Post-nominal NQ: ✓distributive ✓collective)

b. [Gakusei]-ga kyoo san-nin syukudai-o dasita.
[student]-NOM today three-CL homework-ACC submitted
‘Three students submitted homework today.’
(Floating NQ: ✓distributive ✘collective)
(Okuma, 2019, p. 497)

Okuma investigated whether English-speaking L2 learners of Japanese could acquire this semantic restriction on floating NQs. According to Okuma, the corresponding property does not exist in learners’ L1 (English), nor is the semantic restriction on floating NQs taught in L2 Japanese language classrooms. Therefore, English-speaking learners need to acquire it based on natural L2 input. However, since there seems to be no negative evidence, such error correction or instruction whereby information is provided to indicate the incompatibility of floating NQs with the collective interpretation, this situation might pose a POS problem. Okuma examined whether such a learnability problem involved in L2 acquisition of Japanese NQs could be overcome presumably as a consequence of UG (i.e., a universal syntax-semantics interface) operating in L2 acquisition, as suggested in previous studies, such as those cited in this section). A TVJT was undertaken by 22 native Japanese controls and 18 English-speaking learners of Japanese with intermediate to advanced proficiency. The results showed that although the L2 learners did not reliably distinguish between the collective and distributive interpretation of floating NQs at the group level, four out of the eighteen learners made a target-like distinction. Based on these results, Okuma argues that successful acquisition of the semantic constraint on Japanese floating NQs is possible despite the relevant learnability problem.

2.2.3 Main findings of L2 research on the syntax-semantics interface

The findings of the previous studies reviewed above point to the following characteristics of L2 syntax-semantics mappings. Whether syntax-semantics mismatches or POS situations, L2 interpretive properties seem ultimately acquirable. However, L1 transfer (either positive or negative) may lead L2 learners to take different acquisition paths.

With regard to syntax-semantics mismatches, Ionin et al. (2013) showed that L2 learners were generally successful in overcoming L1 influence attributed to the form-meaning mismatch. Similar results have been obtained by many other researchers investigating different linguistic phenomena with different L1-L2 combinations (e.g., Gabriele, 2008; Montrul & Slabakova, 2003; Slabakova, 2003). On the other hand, Gabriele (2010) suggested a syntax-semantics mismatch may remain problematic for some learners even with advanced proficiency due to L1 negative transfer. Other studies (e.g., Gabriele, 2009; Yuan, 2010) also have provided evidence for such persistent difficulty of recovering from negative L1 influence. Some bidirectional studies (e.g., Gabriele, 2005) suggest that with a given form-meaning mismatch between two languages, L2 acquisition tends to be slower and less successful when negative evidence is required to attain the target property than when positive evidence is sufficient to do so.

Essentially similar findings have been reported from studies testing POS situations. Unsworth (2005) suggested that novel L2 constructions and their interpretive properties are acquirable despite the POS. Other studies such as Dekydtspotter and Sprouse (2001) and Dekydtspotter, Sprouse, and Anderson (1997) came to the same conclusion with different linguistic properties in L2 French, although in those studies, target-like performance was observed even with intermediate learners as well as advanced learners. Okuma (2019) provides further evidence for this argumentation: even though the L1 (i.e., English) lacks the target interpretive L2 property (i.e., the distributivity constraint on floating NQs) and the situation constitutes a POS problem, at least some learners can acquire that L2 property.

These findings generally support the argument that the universal compositional semantic computation made available by UG is operative not only in the L1 but also in the L2. With this domain-specific parsing mechanisms, POS semantic properties seem to come for free, once L2 learners have acquired the lexical properties of functional morphemes and efficient morphological decomposition relevant to the target construction. In this sense, morphology can be seen as a bottleneck in L2 acquisition as opposed to syntax-semantics computation, which appears to flow relatively smoothly (Slabakova, 2008). Furthermore, as documented in Marsden (2009), whether the L1 has a morpheme with the same lexical properties as the target L2 morpheme seems to play a role. When the L1 has a corresponding morpheme, facilitation is expected. By contrast, acquisition may be delayed when L1 and L2 equivalent morphemes do not match in lexical specification. Testing a different POS property with L2 Japanese learners with L1-English, Korean, and Chinese, Marsden (2008) also documented negative L1 transfer due to different specifications of the relevant L1 and L2 morphemes.

In summary, linguistic properties at the L2 syntax-semantics interface seem overall ultimately acquirable, although relative ease or difficulty of acquisition may depend on the target property and the L1-L2 combination. As detailed in the next section, more recently, the question of what conditions make L2 acquisition at the syntax-semantics interface relatively easier or more difficult has been addressed, within the feature-based contrastive framework, which the present thesis adopts.

2.3 Feature-based contrastive approaches to L2 acquisition

2.3.1 Current generative approaches to learnability problems in L2 acquisition

In addition to syntax-semantics mismatches and POS problems, there have been different sources of persistent difficulty proposed in recent generative SLA research. For example, the Interpretability Hypothesis (Hawkins & Hattori, 2006; Tsimpli & Dimitrakopoulou, 2007) is a proposal based on the notion of interpretable vs. uninterpretable feature contrast in the current generative grammatical theory, the Minimalist Program (e.g., Chomsky, 1995, 2000). Interpretable features are features that make fundamental contributions to meaning (e.g., number, gender and definiteness on nouns), whereas uninterpretable features are those relevant only to grammaticality/morphosyntax (e.g., case, agreement). This hypothesis holds that uninterpretable features not selected by the L1 are no longer accessible (thus acquirable) in the L2 after a critical period. Another influential generative account for L2 learnability is the Interface Hypothesis (Sorace, 2006; Sorace & Filiaci, 2006), which predicts greater difficulty for L2 learners when having to integrate linguistic information across different domains (e.g., syntax and semantics, syntax and phonology) than when focusing on one specific domain (e.g., 'narrow' syntax). In a more recent version of this hypothesis (e.g., Sorace, 2011; Sorace & Serratice, 2009; Tsimpli & Sorace, 2006), those properties at external interfaces, whereby internal components of the grammar (i.e., syntax/morphology/semantics/phonology) interact with external components (i.e., discourse/pragmatics), are considered more challenging than those related to internal interfaces, whereby integration of components takes place within the linguistic-internal system.

Although these hypotheses have been empirically supported to some extent (e.g., Hawkins & Hattori, 2006; Tsimpli & Dimitrakopoulou, 2007 for the Interpretability Hypothesis; Belletti, Bennati, & Sorace, 2007; Sorace & Filiaci, 2006 for the Interface Hypothesis), they also have been called into question as evidence has accumulated against them, namely that the kind of properties predicted to be persistently hard to acquire even for highly proficient learners (i.e., phenomena related to uninterpretable features or external interfaces) may not be necessarily so (e.g., Prevost & White, 2001; Rule & Marsden, 2006 for the Interpretability Hypothesis; Ivanov, 2009; Slabakova & Ivanov, 2011 for the Interface Hypothesis). Such mixed results underscore the complexity of identifying areas and conditions where L2 acquisition is inherently difficult or easy. The present thesis addresses this problem from different perspectives, namely the Feature Reassembly Hypothesis (Lardiere, 2008, 2009) and the cline of difficulty in feature acquisition (Cho & Slabakova, 2014; Slabakova, 2009), as discussed below.[2]

2.3.2 The Feature Reassembly Hypothesis

Acknowledging potential difficulties involved in the L2 acquisition of features, Lardiere

[2] Note that the present thesis does not test the Interpretability Hypothesis or the Interface Hypothesis because the target definiteness properties involve only interpretable features and they concern external interfaces. That is, the Interpretability Hypothesis does not offer concrete predictions for the acquisition of interpretable features, whereas the Interface Hypothesis, in essence, predicts difficulty of external interface properties in relation to internal interface properties, which the thesis does not examine.

(2008, 2009) proposed the Feature Reassembly Hypothesis (FRH). According to the FRH, L2 acquisition progresses as features bundled on L1 lexical items are mapped onto their closest L2 counterparts based on functional and semantic similarities. The inventory of features potentially realised by natural languages is assumed to remain universally available to L2 learners, regardless of whether or not they are interpretable and how late the onset of the L2 acquisition is. However, languages differ in terms of which features map onto overt morphemes and how features are combined onto those morphemes, as postulated within the Minimalist framework. Lardiere argues that difficulty may arise particularly when there is a mismatch in feature specification between the L1 and the L2 equivalent lexical items because the learner needs to reassemble the L1-based feature configuration into the target representation based on the linguistic input. That is, the main prediction of the FRH is that L2 acquisition will become more difficult when learners have to reassemble the L1-based features into a new target configuration compared to when they do not have to do so. Lardiere also posits that the key factor for ultimate attainment in L2 acquisition is whether the learner has access to input that motivates the relevant feature reassembly: when some reassembly is necessary but triggering evidence is hard to detect or even non-existent, successful feature reassembly may be delayed or never occur. The FRH builds on the Full Transfer/Full Access hypothesis (Schwartz & Sprouse, 1996), according to which L2 learners transfer all the L1 properties and restructure them into the L2 specifications by availing themselves of positive evidence in the L2 input and innate linguistic knowledge provided by UG.[3] However, the novelty of the FRH lies in its potential to specify a given acquisition task more precisely by making use of the notion of mapping and (re)assembly of lexical features between the L1 and the

[3] Note also that the FRH is compatible with other proposals assuming full L1 lexical transfer, such as Sprouse (2006) and Stringer (2010). However, I focus on the FRH since the cline of difficulty, another theoretical proposal to be examined in this thesis is framed within it.

L2.

To illustrate how acquisition tasks can be formulated within the FRH, let us consider the L2 acquisition of plural markings in Mandarin Chinese (henceforth, Mandarin) and English bi-directionally, as discussed in Lardiere (2009). The Mandarin plural marker, *-men* represents not only the [+plural] but also the [+definite] and [+human] features (Li, 1999), unlike the English plural marker, *-s*, which encodes only the [+plural] feature. Put differently, English *-s* is underspecified for the [±human] and [±definite] features. The distinct feature specifications of the two plural morphemes are summarised in Table 2.3.

Table 2.3 Lexical features of English and Mandarin plural markers

	English *-s*	Mandarin *-men*
[plural]	+	+
[human]	±	+
[definite]	±	+

Learners are predicted to initially treat the L2 plural marker as an equivalent of the L1 plural marker based on perceived functional similarities. However, the subsequent feature reassembly task differs according to the direction of learning: L1-English L2-Mandarin learners need to constrain the distribution of plural marking from [±human, ±definite] to [+human, +definite], whereas L1-Mandarin L2-English learners need to extend the feature distribution from [−human, −definite] to [±human, ±definite]. Furthermore, the latter may not be as challenging as the former because learners can rely on positive evidence in the input: instantiations of plural marking *-s* on [−definite] and [−human] nominals. By contrast, to achieve the constraining task, learners may require negative evidence that *-men* is not compatible with [−human] or [−definite]. However, such information is unlikely to be available in natural input, (similarly to poverty-of-stimulus situations), thus this task may be more difficult.[4] Furthermore, what is subject to reassembly also includes conditioning factors for feature expression such as whether overt realisation of the feature is obligatory/optional in what morphosyntactic/semantic/phonological/discoursal environments. For example, although plural marking is in principle optional in Mandarin unlike in English, it is prohibited under some conditions in Mandarin (e.g., *-men* cannot be used with a quantifier). Lardiere argues that to figure out all these crosslinguistic differences and successfully conduct the relevant feature reassembly based on detectable evidence is far from trivial.

2.3.3 Previous studies testing the FRH

The main proposal of the FRH regarding what makes L2 acquisition more difficult has been actively tested by a number of researchers with a variety of linguistic phenomena and L1-L2 combinations. Some have confirmed that the more complex the feature reassembly, the more apparently difficult the acquisition task, through investigation of the acquisition of L2 morphemes that have different feature specifications from functionally similar L1 morphemes (e.g., Choi & Lardiere, 2005; Shimanskaya & Slabakova, 2015; Spinner, 2013). Other studies have provided evidence for the argument that a task that involves more feature reassembly means additional difficulty, through a comparison of cases where reassembly is not necessary with cases where reassembly is necessary; and a comparison of cases where less reassembly vs. more reassembly is necessary (e.g., Cho, 2017; Cho & Slabakova, 2015; Gil & Marsden, 2013; Hwang & Lardiere, 2013). Furthermore, studies testing the contrast between relevant-input-available and relevant-input-unavailable cases both across and within languages (e.g., Gil & Marsden, 2013; Su,

[4] However, in reality, the deletion of the [−human] feature on Mandarin *-men* by English speakers does not seem so problematic. According to Su (2019), the fact that *-men* can be used only with [+human] nouns is covered in some Mandarin textbooks and this may constitute the relevant evidence and thus facilitate the relevant feature reassembly.

2019; Yuan, 2014) found that the latter tends to be more problematic, as predicted by the FRH. Importantly, some studies suggest that even when L1-driven feature mapping is expected to make the subsequent reassembly considerably challenging, successful acquisition seems possible at least for learners at advanced stages (e.g., Domínguez, Arche, & Myles, 2017; Gil & Marsden, 2013; Lee & Lardiere, 2019). These findings generally demonstrate that the FRH is indeed a testable and promising account of L2 acquisition.

However, researchers also have raised questions in relation to the FRH. Particularly, White (2009) takes issue with the predictive power of the FRH. She argues that it remains unclear which types of reassembly are more problematic than others, whether a certain L1-L2 combination presents a greater challenge and whether all features are equally difficult to reassemble. Several proposals have been put forth to address these questions. Hwang and Lardiere (2013) argue that the more complex (phonological, morphosyntactic, semantic, pragmatic, and discourse) conditions under which a feature is expressed are and the more they differ between the L1 and the L2, the harder it becomes to acquire the relevant property. Particularly, they predict that the more dependent a feature is on other features in the feature hierarchy, the more difficult it is to acquire. Mai and Yuan (2016), on the other hand, maintain that feature reassembly becomes more complicated and difficult to accomplish when it involves "cross-domain" reassembly (e.g., from prosody to syntax) compared to when it takes place within a single linguistic domain. Another attempt particularly relevant to this thesis is the cline of difficulty in feature acquisition proposed by Slabakova (2009) and elaborated by Cho and Slabakova (2014), which is discussed in detail in the following section.

2.3.4 Cline of difficulty in feature acquisition

Addressing the predictive power problem with the FRH pointed out by White (2009), Slabakova (2009) proposed that the relative ease/difficulty of feature reassembly varies according to whether relevant features are realised overtly or covertly and whether the realization of the features is similar or different in the L1 and the L2. Slabakova (2009) draws on Ramchand and Svenonius's (2008) speculation that in L1 acquisition, features realised overtly are acquired earlier than those realised covertly because a grammaticalized morpheme should provide more constant input for feature value tracking than discourse/context.[5] Combining Ramchand and Svenonius's idea with the FRH, Slabakova postulated a cline of difficulty in L2 acquisition of functional features as in Figure 2.1 (adapted from Slabakova, 2009, p. 321).

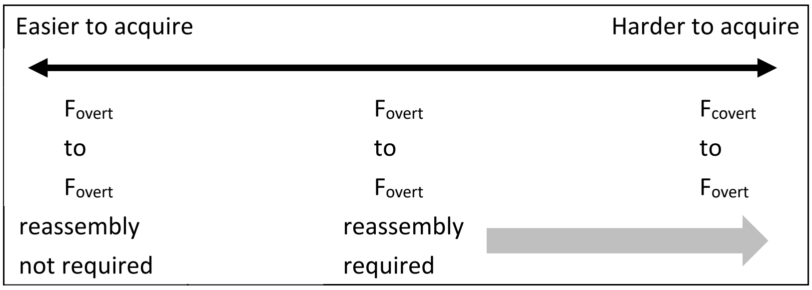


Figure 2.1 Slabakova's (2009) cline of difficulty in feature acquisition

The left-most point represents the case where a give feature (F) is represented by an overt morpheme in the L1 and an overt morpheme in the L2, and both the L1 and L2 morphemes have the same syntactic and semantic features, so no feature reassembly is required. This category is considered the easiest. One level harder is the case at the middle point where

[5] For justification of their proposal, Ramchand and Svenonius (2008) pointed out asymmetric acquisition rates of definite articles in L1 English and Norwegian. According to Ramchand and Svenonius, the meaning of definiteness consists of multiple semantic primitives including familiarity and specificity; and English and Norwegian differ in how they express these components of definiteness: familiarity is overtly realised with a morpheme (pre-nominal determiner) in both languages, whereas specificity is distinguished morphologically through a suffix in Norwegian but expressed covertly through context in English. The researchers argue that this crosslinguistic difference can account for the experimental observations that L1-Noweigian children appear to be able to acquire the definite suffix at the age of 2(Anderssen, 2007), whereas L1-English children seem to have difficulty using the definite article properly until the age of 4 (Schaeffer & Matthewson, 2005).

a given feature is encoded overtly in both the L1 and the L2 but differently (e.g., through different morphemes or through similar morphemes but with distinct feature configurations). The right-most point is the hardest task, where a given feature is realised by an overt morpheme in the L2 but distinguished covertly through discourse/context in the L1. Note that this case inherently involves reassembly, in terms of manner of realisation from covert to overt, although in the present thesis (and relevant previous studies), it is simply labelled as F_{covert} *to* F_{overt} (the same applies to the F_{overt} *to* F_{covert} category). This proposal can be regarded as a step forward from the FRH in terms of increased predictive power: the addition of this new dimension of overt vs. covert feature realisation offers more detailed predictions for relative difficulty/ease of acquisition of a given feature expression with a specific L1-L2 combination.

Building on Slabakova's (2009) proposal, Cho and Slabakova (2014, based on Cho, 2012) presented an extended cline as in Figure 2.2.[6] Specifically, they proposed that features expressed covertly (F_{covert}) are harder to acquire in the L2 than those expressed overtly (F_{overt}), following the same logic of overt vs. covert realisation as Ramchand and Svenonius (2008) used for their proposal with L1 acquisition.

---

[6] Cho and Slabakova (2014) do not cite Cho (2012), though the experimental data they report appear to be the same data reported in Cho (2012). The latter also tested the cline in Figure 2.2 with slightly different theoretical assumptions. Henceforth, following Cho and Slabakova (2014), I do not refer to Cho (2012).

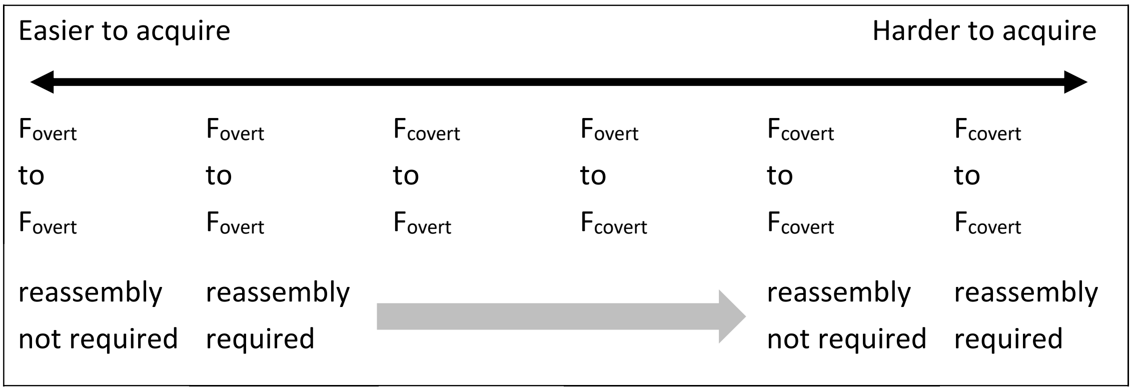


Figure 2.2 Extended cline of difficulty in feature acquisition (adapted from Cho & Slabakova 2014, p. 166)

In addition, Cho and Slabakova further clarified the notion of covert features (F_{covert}) by defining them as features determined via "discourse tracking as well as some inconsistent (non-uniform) signals that require discourse observation, including word order changes" (Cho & Slabakova, 2014, p. 166). In this sense, "covert" applies to any non-morpholexical manifestation of a feature.

Cho and Slabakova (2014) empirically tested some predictions of the revised cline with L2 Russian learners with L1-English or Korean by investigating their knowledge of overt and covert expressions of definiteness in Russian.[7] Russian does not have dedicated morphemes to overtly distinguish definiteness. Definiteness is expressed predominantly through context in Russian. Korean is similar in this respect, unlike English, a language with dedicated morphemes (i.e., articles) to systematically express the feature overtly. Russian can still encode definiteness overtly through possessor modifiers. Specifically, whereas nominal possessors can be interpreted as either [+definite] or [−definite] (2.10) depending on context, adjectival possessors are restricted to a [−definite] interpretation (2.11) (both examples are from Cho & Slabakova, 2014, p.

[7] Cho and Slabakova (2014, p.161) adopted an informal definition of definiteness as follow: "a nominal is definite when there is a presupposition of its referent being unique in the domain of discourse, where unique can be established through previous mention or world knowledge."

167).

(2.10)	Za	dverju	slyšalsja	golos	ženščiny.
	behind	door	heard	voice-NOUN.NOM	woman-NOUN.GEN

‘The voice of a/the woman was heard behind the door.’

(2.11)	Za	dverju	slyšalsja	ženskij	golos
	behind	door	heard	woman-ADJ.NOM	voice-NOUN.NOM

‘A woman’s voice was heard behind the door.’

According to Cho and Slabakova, Korean also has an overt way to express definiteness through case-marking in noun-noun compound constructions. Furthermore, Russian can additionally express definiteness covertly, by means of word order. In Russian, definite NPs are usually placed in pre-verbal position and indefinite NPs in post-verbal position (2.12). In particular, bare indefinite nouns cannot occur pre-verbally (2.13).

(2.12)	Na	stole [+definite]	stoja-la	lampa [−def].
	On	desk	stand-PAST	lamp

‘A lamp was on the desk / there was a lamp on the desk. ’

(Cho & Slabakova, 2014, p. 169)

(2.13)	Na	stole [+definite]	lampa [+definite]	stoja-la.
	On	desk	lamp	stand-PAST

‘The lamp was standing on the desk/ # on a desk.’

(Adapted from Cho & Slabakova, 2014, p. 169)

Cho and Slabakova point out that Korean also has a covert way to express definiteness through word order: object nouns must be interpreted as [+definite] in the OSV order. But what makes Korean different from Russian is that all nouns must occur pre-verbally in Korean. That is, unlike in Russian, post-verbal nouns are ungrammatical in Korean. English does not mark definiteness covertly through word order change.

Based on these crosslinguistic analyses, the extended cline of difficulty (the cline of difficulty, henceforth) offers the following predictions. First, because both English and Korean have overt morphological means to express definiteness (i.e., articles, and case marking in noun-noun compound constructions, respectively), the acquisition task for the Russian overt definiteness expression is the second easiest for both L1 groups: *F_{overt} to F_{overt} reassembly required*. As to the covert distinction of definiteness in Russian, the cline of difficulty assigns the third hardest task, *F_{overt} to F_{covert}* for English-speaking learners; and the hardest task, *F_{covert} to F_{covert} reassembly required* for Korean-speaking learners. This is based on the crosslinguistic fact that Korean covertly realises definiteness through word order change but in a different way to Russian, whereas English has dedicated morphemes (i.e., articles) to overtly encode definiteness. In sum, both L1-English and Korean learners will acquire the overt L2 Russian property more easily or earlier than the covert property, and English-speaking learners will acquire the covert property more easily or earlier than Korean-speaking counterparts.

Cho and Slabakova (2014) tested these predictions using a felicity judgement task. Participants were instructed to read short passages and rate on a 5-point Likert scale how felicitous the target sentence was as a description of each story. The results were generally in favour of the predictions. On the one hand, both English- and Korean-speaking learners performed better on the overt property than the covert property, in line with the prediction that a covert expression of a feature is more difficult than an overt expression, even when some reassembly is required for the latter. In terms of the covert property, on the other hand, whereas some English-speaking advanced learners responded in a target-like way, target-like performance was not observed with the Korean-speaking learners even at advanced levels, which is also compatible with the cline. The researchers speculated that the overt realization of a feature in the L1 may have helped the English-speaking learners to acquire the covert expression of the feature in the L2 whereas reassembling a feature that is covertly realised both in the L1 and L2 in different ways can cause great difficulty to an extent that is impossible for Korean-speaking learners to overcome.

The potential effect of an overt L1 feature on acquiring covert expression of that feature in the L2 is a key claim of Cho and Slabakova (2014) that has not explicitly been tested; and this is a focus of the present thesis. Cho and Slabakova's (2014) findings allow at least two interpretations regarding the role of the L1 for Korean-speaking learners. One is that covertly realised features in the L1 are never facilitative in the acquisition of corresponding L2 features. This is indeed what the cline (Figure 2.2) implies, although Cho and Slabakova do not state it explicitly. This suggests that even if the identical covert property exists in the L1 (i.e., *F_{covert} to F_{covert} reassembly not required*), acquisition of that property may be more difficult than in the case of *F_{overt} to F_{covert}*, where learners may benefit from overt realisation of the relevant L1 feature. The other possibility is that it may be the feature reassembly rather than the covert realisation of the feature in both the L1 and the L2 that makes the *F_{covert} to F_{covert} reassembly required* task unattainably difficult. This seems to be a more natural explanation from the perspective of the FRH, which maintains that the more reassembly is necessary, the harder the acquisition task becomes, whether the relevant feature is realised overtly or covertly. In a sense, this is a question about which advantage is stronger for L2 acquisition of a covert feature expression, having a dedicated morpheme for the relevant feature in the L1 or having a corresponding covert feature expression. The present thesis addresses this question.

2.3.5 Previous studies testing the cline of difficulty

There are, to my knowledge, two recent studies that tested predictions based on the cline of difficulty: Tuniyan (2018) and Köylü (2019). However, these studies had mixed results as shown below.

Tuniyan (2018) tested three points on the cline: *F_{overt} to F_{overt} reassembly not required*, *F_{overt} to F_{overt} reassembly required* and *F_{covert} to F_{overt}*, by investigating the L2 acquisition of English article semantics with native speakers of Mandarin (two levels: intermediate and advanced) and Russian (three levels: beginner, intermediate, and advanced). English overtly encodes two different types of definiteness, namely familiarity and uniqueness with its articles: the features [+familiar] (i.e., indication of a familiar discourse referent for both the speaker and the hearer) and [+unique] (i.e., presupposition of a unique referent in a given situation or through general knowledge) are both overtly marked with the definite article, *the* (e.g., Schwarz, 2009, 2013). The feature [+familiar] can also be encoded with the demonstrative, *that.* However, because the way familiarity is computed differs between *the* and *that*, there are contexts where one is more appropriate than the other.[8] On the other hand, Mandarin and Russian, both of which are languages without an article system, do not obligatorily express familiarity or uniqueness morphologically. However, these languages can mark the feature [+familiar] overtly with their demonstratives (i.e., Mandarin *na* and Russian *etot*) in the same way as English *that*, whereas they must realise the feature [+unique] covertly through context (i.e., bare nouns). Therefore, relative ease or difficulty for the L2 acquisition of overt realisation of [+familiar] and [+unique] definiteness in L2 English by L1 Mandarin and Russian

[8] Put briefly, for *the* to be felicitous, a relevant familiar referent must be the most salient in the context, whereas for *that*, the referent needs to be immediately salient (i.e., most recently mentioned). This subtle semantic difference between *the* and *that* is relevant to the present thesis, too, and is discussed in some detail in Chapter 3.

speakers were predicted as in Figure 2.3.[9]

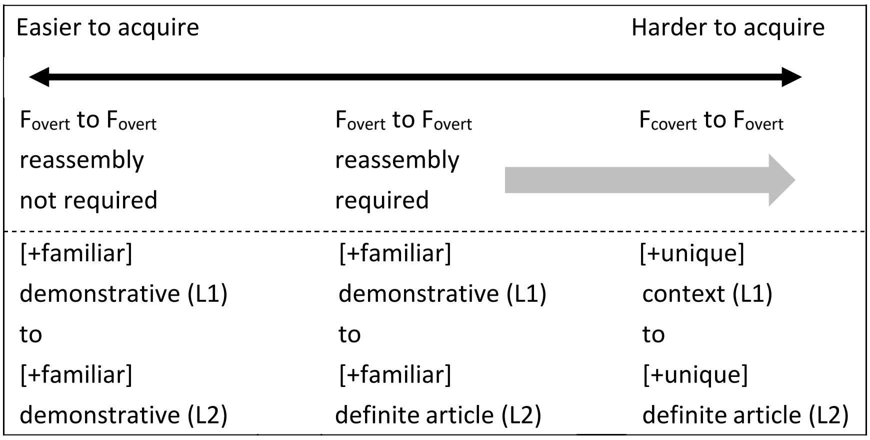


Figure 2.3 Predicted difficulty of acquisition of [+familiar] and [+unique] features in L2 English by L1-Mandarin and Russian speakers (adapted from Tuniyan, 2018, p. 139)

The easiest acquisition task is the mapping of the feature [+familiar] from L1 demonstrative to L2 demonstrative. Because of the correspondence between the L1 and L2 demonstratives in terms of how the feature is computed, no reassembly is required (i.e., F_{overt} *to* F_{overt} *reassembly not required*). Less easy is the task of mapping of the [+familiar] feature from L1 demonstrative to L2 definite article: in this case, the conditioning environment of feature realization needs reassembly between L1 demonstrative and L2 definite article (i.e., F_{overt} *to* F_{overt} *reassembly required*). The greatest difficulty is expected for the [+unique] feature mapping from context to the definite article (i.e., F_{covert} *to* F_{overt}) due to the covertness of the feature in the L1.

In order to test the prediction above, Tuniyan (2018) investigated the L2 interpretation and use of English articles by means of an AJT and a written sentence production task, respectively. In the AJT, participants were asked to read short stories and

[9] Tuniyan also investigated overt expression of indefiniteness with the indefinite article *a* in English, which will be briefly explained in Chapter 4, where previous studies on the L2 acquisition of definiteness are reviewed.

decide whether the test sentence with the target NP following each story was acceptable or not. In the written sentence production task, the participants had to read the beginning of a story and continue the story by making sentences using lexical words that were provided, but adding functional words such as articles and conjunctions as necessary. The task was designed to elicit an article (*the* or *a*) for the target NP that fit each story. Tuniyan found suggestive (but not statistically significant) evidence for the contrast of *F_{overt} to F_{overt} reassembly not required* vs. *F_{overt} to F_{overt} reassembly required* and *F_{covert} to F_{overt}* in the AJT performance of L1-Russian learners. Specifically, while the advanced L1-Russian group performed in a target-like way across the contexts, the intermediate group did so only in the *F_{overt} to F_{overt} reassembly not required* context. The beginners were not target-like in any contexts. This developmental pattern is compatible with the prediction (Figure 2.3). However, as to the L1-Mandarin learners, such a contrast was not observed: the AJT data from both intermediate and advanced L1-Mandarin groups showed target-like performance, regardless of the context type.[10] On the other hand, the results of the written sentence production task, which compared *F_{overt} to F_{overt} reassembly not required* and *F_{covert} to F_{overt}* situations, suggested that the L2 learner groups were generally native-like, regardless of overtness of the relevant feature in the L1, except for the L1-Russian beginner group, who omitted articles about 90 % of the time across the contexts and showed little sign of the relevant knowledge. Therefore, the prediction was only partially borne out. The L1-Russian learners' AJT results indicated that as expected, the *F_{overt} to F_{overt} reassembly not required* task is the easiest and that some combination of feature reassembly necessity and covert realization of the relevant L1 feature makes the task harder. However, it was left open whether a *F_{covert} to F_{overt}* mapping is more problematic

[10] Tuniyan (2018, p. 209) argues that the contrast between the intermediate learner groups may be attributed to the fact that the Chinese intermediate learners were statistically more proficient than the Russian intermediate learners.

than an F_{overt} *to* F_{overt} mapping with feature reassembly acquired, unlike Cho and Slabakova's (2014) cline suggesting that is the case. Furthermore, the absence of the overt vs. covert contrast in the production data did not support the prediction.

Köylü (2019) mainly examined whether acquisition of a feature that is covertly expressed through context in the L2 but overtly with a functional morpheme in the L1 is more difficult than the opposite direction (Slabakova, 2009) through an investigation of the acquisition of kind-referring NPs in L2 English by native speakers of Arabic, Chinese, and Turkish. Köylü argues that kind reference is a product of an interaction of three features: (i) the [−domain restriction] ([−dr]) feature, which is encoded on an overt or a covert D(eterminer), (ii) the [+set] feature, which is realised with the head of a functional projection in the nominal domain, Set Phrase, and (iii) [−exception tolerance] ([−et]), which was overtly lexicalised with kind-selecting predicates (e.g., *extinct*, *rare*, *common*). Table 2.4 summarises crosslinguistic variation between the four languages in terms of what NP forms can refer to kinds and how each feature is realised. Overt determiners that encode [−dr] are all shown as *the* for simplicity. "∅" represents a covert determiner for the [−dr] feature or absence of overt plural marking on the NP (all plural markers are presented as *-s*) embedded in the Set Phrase that hosts the [+set] feature. No variation exists in the way the [−et] feature is encoded through kind predicates across the languages. Köylü's predictions for learning tasks and their relative ease/difficulty are presented in (2.14) (learning tasks are described in parentheses for each situation). In short, whenever a covert vs. overt contrast is identified between the L1 and the L2 with a given feature, the task was expected to become harder (in line with the FRH), and L1-overt L2-covert situations are assumed to be more difficult than L1-covert L2-overt situations (in line with the cline of difficulty) and tasks with no feature reassembly necessary were expected to be the easiest (in line with the FRH).

Table 2.4 Crosslinguistic variation in kind-referring count NPs in the four languages tested by Köylü (2019)

		[−domain restriction]	[+set]	[−exception tolerance]
English	Definite singulars	*the*	∅	kind predicate
	Bare plurals	∅	N-*s*	kind predicate
Turkish	Bare singulars	∅	∅	kind predicate
	Bare plurals	∅	N-*s*	kind predicate
Arabic	Definite singulars	*the*	∅	kind predicate
	Definite plurals	*the*	N-*s*	kind predicate
Chinese	Bare singulars	∅	∅	kind predicate

(2.14) Predictions for learning tasks (adapted from Köylü, 2019, p. 203)

a. The easiest

- Arabic learners acquiring kind reference with definite singulars

  (No reassembly required.)

- Turkish learners acquiring kind reference with bare plurals

  (No reassembly required.)

b. Intermediate difficulty

- Turkish learners acquiring kind reference with definite singulars

  (Need to map [−dr] from ∅ to *the*)

- Chinese learners acquiring kind reference with definite singulars

  (Need to map [−dr] from ∅ to *the*)

- Chinese learners acquiring kind reference with bare plurals

  (Need to map [+set] from ∅ to a Set Phrase with overt plural marking-*s*)

c. The most difficult

- Arabic learners acquiring kind reference with bare plurals

  (Need to map [−dr] from *the* to ∅)

  (Based on Köylü, 2019, pp. 142–144)

Köylü (2019) administered a fill-in-the-gap task and an AJT to test the predication. In the fill-in-the-gap task, participants saw a context sentence and subsequently an incomplete sentence with a picture. The task was to complete the sentence using the context and the picture as clues. It was designed to elicit the target NP forms in each context. In the AJT, participants were presented with a short sentence followed by 5 minimally different continuations (i.e., a bare singular noun, indefinite singular, a definite singular, a bare plural, and a definite plural). They rated each continuation on a 4-point scale. The results of both tasks suggested that the L2 English learners had difficulty with definite singulars but not with bare plurals, irrespective of the L1 background. The researcher also reported that the Arabic speakers were more native-like in definite singulars than the Chinese and Turkish speakers, although they were still significantly different from native speakers. In contrast, all the L1 groups were statistically indistinguishable from the native controls with bare plurals. Therefore, the results seem, again, mixed: although the Arabic speakers' higher accuracy compared to the other L1 groups with singular definites was compatible with the prediction, their more target-like performance with bare plurals was clearly not. Furthermore, the predicted acquisition order between the learner groups for bare plural NPs were not attested. Instead, universally successful performance was observed.

Note that Köylü does not seem to exactly follow Cho and Slabakova's (2014) cline of difficulty because in his prediction, covertness of a feature does not pose a particular problem when no reassembly is necessary. That is, *F_{overt} to F_{overt} no reassembly required* and *F_{covert} to F_{covert} no reassembly required* tasks seem to be treated as equally easy, though the latter is predicted to be harder to acquire on the revised cline (Figure 2.2).[11] However, the results do not constitute any clear-cut evidence for either version of the cline or the FRH. Furthermore, since kind reference, as Köylü defines it, involves a combination of three different features, with relevant nominal constructions (e.g., definite singulars, bare plurals) encoding those features overtly and covertly in different ways across languages, it seems methodologically difficult to effectively evaluate the role of overt vs. covert realisation of each feature.

---

[11] Taking for example the acquisition of the [−dr] feature for bare plurals by L1-Turkish learners and that for definite singulars by L1-Arabic learners, although both tasks involve no reassembly, the former seems to be rated as more difficult on the revised cline because it is a covert-to-covert mapping.

2.4 Remaining questions about feature contrastive approaches to L2 acquisition

In this chapter, the FRH has been shown to be an appropriate approach to explore L2 learnability at the syntax-semantics interface. As discussed in the earlier sections, it has been supported by a number of studies that set out to test it, but it can also account for findings of studies that pre-date it (such as Marsden, 2009 and others). Nevertheless, for improvement of its predictive power, proposals about what makes feature reassembly difficult are in order. In this respect, the cline of difficulty appears to be one promising candidate. However, whereas the majority of previous studies within the FRH (cited above) concerned L2 acquisition of features overtly realised through morpholexical items and produced generally supporting evidence, the overt vs. covert contrast subsumed under the cline of difficulty needs further empirical validation. Specifically, there are at least four outstanding questions as below, about the relation between the two key factors in the cline of difficulty model: overt vs. covert feature realisation and necessity of reassembly.

(2.15) Question 1:

When no feature reassembly is required, is an F_{overt} *to* F_{overt} task easier than F_{covert} *to* F_{covert}?

(2.16) Question 2:

When the L2 feature is covert, is it easier to acquire it when the L1 feature is overt (F_{overt} *to* F_{covert}), or when the L1 feature is covert in the same way as in the L2 (F_{covert} *to* F_{covert} *reassembly not required*)?

(2.17) Question 3:

Is a reassembly task between overt expressions (*F_{overt} to F_{overt} reassembly required*) easier than a reassembly from a covert to overt expression (*F_{covert} to F_{overt}*)?

(2.18) Question 4:

In which situation is it easier to acquire a feature, when the feature is expressed overtly in the L1 or in the L2?

Question 1 corresponds to the contrast between *F_{overt} to F_{overt} reassembly not required* and *F_{covert} to F_{covert} reassembly not required* on the cline in Figure 2.2. This comparison is important in terms of examining the role of overt vs. covert contrast. The cline assumes that an overt expression of a feature is easier to acquire than a covert expression of that feature. However, this assumption needs further corroboration, given the conflicting previous findings above (e.g., generally successful acquisition of kind reference through bare plural NPs vs. generally incomplete acquisition of kind reference through singular definite NPs in Köylü (2019)).

As to Question 2, under the cline of difficulty (Figure 2.2), covert realization of a feature is predicted to lead to acquisition difficulty, even when no feature reassembly is required. Importantly, the cline suggests that having an overt functional morpheme to distinguish the relevant feature in the L1 will ease the L2 acquisition of a covert expression of the same feature to a greater degree than having a corresponding covert means of feature realisation in the L1. But, as yet, this prediction has not been tested.

According to the cline of difficulty, the answer to Question 3 is yes. However, as shown above, Tuniyan (2018) tested this prediction but did not find concrete evidence for it. Thus, the relative difficulty/ease between *F_{overt} to F_{overt} reassembly required* and

F_{covert} *to* F_{overt} is worth further testing with different linguistic properties and/or different language combinations.

Finally, Question 4 concerns relative degree of facilitation by overt realisation in the L1 vs. the L2. The cline of difficulty assumes that overt realisation in the L2 leads to a greater facilitative effect than in the L1—a F_{covert} *to* F_{overt} task is predicted to be easier than a F_{overt} *to* F_{covert} task. Köylü (2019) tested this contrast and the results were not supporting: F_{covert} *to* F_{overt} mappings turned out to be more problematic than a F_{overt} *to* F_{covert} (i.e., Arabic-speaking learners acquiring English singular definites vs. bare plurals). This suggests that there may be cases where the facilitative effect of overt realisation in the L1 is stronger than that in the L2. Therefore, further studies testing F_{covert} *to* F_{overt} vs. F_{overt} *to* F_{covert} would be helpful in addressing this question.

This thesis deals with the first two questions about the cline of difficulty along with the FRH's main prediction that feature reassembly makes acquisition harder, through an investigation of L2 acquisition of overt and covert expressions of definiteness in Japanese, which are detailed in the next chapter.

Chapter 3: Semantics of definiteness and its expressions in English, Japanese, and Korean

3.1 Introduction

In this chapter, the semantics of definiteness is discussed followed by an overview of crosslinguistic variation in the realisation of definiteness with a focus on the languages related to this study: English, Japanese, and Korean. Then Japanese definiteness phenomena that the present study focuses on and corresponding or related phenomena in English and Korean are discussed with a survey of representative theoretical accounts for their properties. Comparisons are to be made often between Japanese/Korean, on the one hand, and English, on the other, because of similarities between Japanese and Korean, and differences between these languages and English regarding the properties under investigation. In the final section, key properties of the target phenomena are summarised in terms of what is agreed upon and disputed. Crosslinguistic analyses to be presented in this chapter are largely exploited in later chapters, where I will formulate specific acquisition tasks for learners with each L1 background, and predictions based on the feature-based L2 acquisition theories reviewed in the previous chapter.

3.2 What is definiteness?

Definiteness is a semantic concept that concerns whether a referent is identifiable in the domain of discourse or context. One of the most influential semantic accounts of the feature is the presupposition analysis originally proposed by Frege (1960) and Strawson (1950), and further developed by Heim (1991).[12] In this approach, a noun is [+definite]

[12] Frege (1960) is an English translation of the original work in 1892 written in German titled "*Uber Sinn und Bedeutung*".

when there exists a presupposed unique entity. For example, for the sentence in (3.1) to have a truth value, a unique king of France must exist. If there exists no king or more than one king, then it is neither true nor false (i.e., presupposition failure).

(3.1) John met the king of France yesterday.

The domain for the computation of uniqueness of a referent is relativised to content. For example, the sentence in (3.2) is true even though there obviously exist multiple cats in the world that we live in, as long as there is a unique cat that is salient in "some contextually given domain" (Ionin 2003, p. 34 based on Heim & Wexler, 2000).

(3.2) The cat is eating fish.

Furthermore, uniqueness can be established not only through previous discourse (e.g., the previous mention of a referent) (3.3) but also mutual world knowledge (3.4). In (3.3), the definite nouns, *the motorbike* and *the car* are felicitous because the previous mentions of *a car* and *a motorbike* allow the speaker and the hearer to uniquely identify the referents. One the other hand, for the sentence in (3.4) to be truthfully uttered, the speaker and hearer do not necessarily have to be talking about a contextually salient winner: the uniqueness presupposition can be satisfied by our common knowledge that a race typically has only one winner. When the presupposition is not fulfilled, the nominal is indefinite, marked with the indefinite article *a*, as with *a car*, *a motorbike* and *a prize* in (3.3) and (3.4).

(3.3) I bought a car and a motorbike last year. <u>The motorbike</u> was more expensive than <u>the car</u>.

(3.4) <u>The winner of this race</u> will receive a prize.

This presupposition analysis of definiteness is generalisable to plural nominals by assuming that definite nouns refer to "a greatest element, where, a greatest element of a set M is an element of M which has all other elements of M as parts" (i.e., the maximality presupposition) (Ionin, 2003, p. 35 based on Heim, 1991, p. 22). In the case of singulars, the greatest element of the set described by the noun is one. For example, the singular definite noun, *the cat* (3.5a) is presupposed to denote the set that has only one member, whereas for the plural counterpart, *the cats* (3.5b), there must be a presupposed set of cats that contains two or more members and those must be all members of the set. Plural nouns can be marked as indefinite with the quantifier *some* or the null article *ø* when the maximality presupposition is not established (e.g., *I saw some/ø students*).

(3.5) a. The cat is asleep.

b. The cats are asleep.

(Ionin, 2003, p. 35)

Therefore, In English, the definite article *the* can be felicitously used when the uniqueness/maximality presupposition is established either through a previous mention of a referent or through mutual world knowledge. Conventionally, the kind of definiteness established through a previous mention of a referent (i.e., linguistic antecedent) is known as "familiarity/anaphoricity" (henceforth, anaphoricity for simplicity) (e.g., Hawkins, 1978; Roberts, 2002) and that licenced by situation/context or mutual world knowledge as "uniqueness." These characterisations have been the two most influential analyses of definiteness in the literature.[13] Recently, however, it has been argued that these two types

[13] For example, Heim (1982) and Kamp (1981) characterise definiteness with familiarity, whereas

of definites are indeed independent universal meanings and distinguished in many other languages with and without articles (e.g., Schwartz, 2009, 2013). In the present study, I assume with Schwartz that English encodes the two types of definiteness with one lexical item, the article *the*, whereas some languages have linguistic means to distinguish them, including the other two languages relevant to the study: Japanese and Korean. Specifically, these languages distinguish the two types of definiteness by marking anaphoricity with demonstratives while leaving unique definite nouns in their bare form, as described in the next section. Importantly for the purpose of this study, looking at crosslinguistic variation in definiteness expressions between English and Japanese/Korean from Schwartz's perspective will allow us to form specific acquisition tasks within the relevant feature contrastive approaches to L2, as demonstrated later in Chapter 5.

3.3 Two types of definiteness: anaphoricity and uniqueness

Before looking at crosslinguistic expressions of the two types of definiteness, let us briefly see some typical cases of each type using English examples. Starting with anaphoricity, in (3.6), *the book* is anaphoric because it refers back to a previously mentioned discourse referent, *a book*.

(3.6) Anaphoric definite

John bought a book and a magazine. **The book** was expensive.

(Adapted from Schwartz, 2009, p. 3)

On the other hand, unique definite expressions are licensed with non-linguistic context that enables the hearer to identify the referent. Consider some English examples of unique

---

Hawkins (1978) and Heim (1991) adopts the uniqueness view.

definiteness from Hawkins (1978), where definiteness is established through situational information rather than the presence of an antecedent.

(3.7) a. Visible situation use

Pass me the bucket, please.

b. Immediate situation use

Beware of the dog (as a sign).

c. Larger situation use

the president (uttered in the US), the Prime Minister (uttered in the UK), the sun, the moon

(3.7a) can be truthfully uttered when the hearer can unambiguously identify the referent of the definite expression, *the bucket* in the relevant visible situation. However, the referent does not have to be visible to be unique: (3.7b) is felicitous as a sentence to inform the reader that there exists a unique dog in the immediate situation. (3.7c) exemplifies cases where uniqueness is established in larger situations such as countries and the world. For example, the definite article can be used with *president* in a situation of the country, the US based on the knowledge shared by the interlocutors that there is only one president in a country. In these examples, the referent is deemed uniquely identifiable in the situation without a linguistic antecedent.

Furthermore, in so-called "bridging" uses (Clark, 1975), definite expressions can relate back to the previous context more indirectly. There are many different types of bridging attested in the literature (e.g., Charolles, 1999; Matsui, 2000), but here I focus on two: one is established through a "producer-product" relation (e.g., *author–book*, *painter–painting*) and the other through a "part-whole" relation (e.g., *fridge–crisper*, *house–living room*, *bike–bike handle*) (Schwartz, 2009, p. 54) because of their relevance to the anaphoric vs. unique contrast, as shown below. (3.8) is a case of product-producer bridging, where the referent is not the definite expression (i.e., *the author*) itself but a unique entity clearly related to an indefinite expression in the previous discourse (i.e., *a book*). Because of its dependency on previous linguistic context, this type of bridging can be considered a special case of anaphoric definiteness.

(3.8) John bought a book today. **The author** is French.

On the other hand, in part-whole bridging as in (3.9), definiteness of the expression depends on the general knowledge shared by the discourse participants in a given situation (i.e., driving situation) that there should be a uniquely identifiable referent denoted by the noun (i.e., *the steering wheel*). Uniqueness, rather than anaphoricity, plays a role here.

(3.9) John was driving down the street. **The steering wheel** was cold.

Although some languages such as English and Italian express both anaphoric and unique definites with a single article, other languages such as German and Fering (a Germanic language) systematically distinguish them in various ways (e.g., Schwartz, 2009, 2013, 2019). For example, Fering uses distinct articles for the two types of definiteness: the weak definite article *a* is used in unique contexts, whereas the strong definite article *di* is used in anaphoric contexts, as illustrated in (3.10).

(3.10) Two definite articles in Fering

a. Unique definite

Ik	skal	deel	tu	a/*di	kuupmaan.
I	must	down	to	the$_{weak}$/the$_{strong}$	grocer

'I have to go down to the grocer.'

b. Anaphoric definite

Oki	hee	an	hingst	keeft.	*A/Di	hingst	haaltet.
Oki	has	a	horse	bought	the$_{weak}$/the$_{strong}$	horse	limps

'Oki has bought a horse. The horse limps.'

(Ebert 1971, p. 161, cited in Schwartz, 2013, p. 538)

Interestingly, Schwartz (2009, 2013) notes that the two types of bridging shown above are related to the strong vs. weak definite contrast. For example, in German, the contrast between these bridging definites is mapped onto two different forms of article. Specifically, weak (i.e., unique) definiteness is realised as contraction of the definite article and an immediately following preposition (3.11a) whereas such contraction cannot occur for strong (i.e., anaphoric) definiteness (3.11b) (Schwartz, 2009, 2013).

(3.11) a.

Der	**Kühlschrank**	war	so	groß,	dass	der	Kürbis
the	fridge	was	so	big	that	the	pumpkin

problemlos	**im/#in dem**	**gemüsefach**	untergebracht
without.a.problem	in_the$_{weak}$/in the$_{strong}$	crisper	stowed

werden	konnte
be	could

'The fridge was so big that the pumpkin could easily be stowed *in the crisper*.'

b.

Das	**Theaterstück**	missfiel	dem	Kritiker	so	sehr,	das
the	play	displeased	the	critic	so	much	that

er	in	seiner	Besprechung	kein	gutes	Haar
he	in	his	Review	no	good	hair

#**am/an dem**	**Author**	ließ.
on_ the_{weak}/in the_{strong}	author	left

'The play displeased the critics so much that he tore the author to pieces in his review.'

(Schwartz, 2009, pp. 52–53)

These crosslinguistic examples constitute evidence for the anaphoric vs. unique (non-anaphoric) split in the analysis of definiteness.

3.4 Crosslinguistic expressions of definiteness (English, Japanese, and Korean)

The main variation lies in whether languages have functional morphemes to systematically distinguish definiteness, that is, whether they have articles or not. However, there are a number of different ways to express the feature (overtly or covertly) as shown below.

3.4.1 Definiteness in English

As already seen in the earlier sections, English distinguishes definiteness morphologically through its article system: the definite article *the* encodes [+definite] for both singular and plural nouns; and to mark [−definite], the indefinite article *a* is used for singular nouns and a null article for plural nouns (i.e., bare plurals). Although definiteness is overtly realised predominately with these articles in English, it can be expressed with other lexical items such as demonstratives and possessive pronouns as well. However, definiteness expressed by articles and other lexical items are not completely equivalent. Particularly relevant to this study are similarities and differences between the definite article, *the* and the demonstrative, *that*. It has been observed that anaphoricity can be overtly encoded with *that*, which shares the same semantic core as the definite article, namely the presupposition of a uniquely identifiable entity in the discourse that satisfies the property denoted by the noun (Roberts, 2002; Wolter, 2006). For example, in English, *that* can mark anaphoricity like *the*, referring to a previously mentioned discourse referent as shown in (3.12).

(3.12) A woman came onto the stage. Then <u>the/that woman</u> started singing and dancing.

(Adapted from Ionin, Baek, Kim, Ko, & Wexler, 2012, p. 75)

However, *the* and *that* differ in terms of where to presuppose the referent to be uniquely identified in the discourse. Specifically, whereas a definite expression needs to find its referent in the whole discourse, a demonstrative expression does so in a further restricted ('salient' in Roberts's terms) set of discourse referents. Turning to the example in (3.13), the woman (i.e., *another woman*) in the second sentence becomes salient by the immediate prior mention. In this situation, *that* can uniquely select the woman as its referent whereas *the* is infelicitous because it cannot single out its referent with there being two relevant discourse referents.

(3.13) A woman entered from stage left. Another woman entered from stage right.

That/ #the woman was carrying a basket of flowers.

(Adapted from Wolter, 2006, p. 4 cited in Ionin et al., 2012, p. 73)

The and *that* also differ in terms of the (in)felicity in less direct anaphoric contexts, namely anaphoric bridging contexts. For example, anaphoric bridging via a product-producer relation can be established with *the* but not *that* (3.14).

(3.14) John bought a book. *The*/#*that* author is French.

(Irani, 2019, p. 123)

Furthermore, *that* cannot express non-anaphoric or unique definiteness unlike *the* (3.15a). Naturally, this extends to contexts where uniqueness is established more indirectly, such as part-whole bridging (3.15b).

(3.15) a. Larger situation uniqueness

the/#*that* sun, *the*/#*that* Prime Minister

b. Bridging uniqueness via a part-whole relation

I bought a house. *The*/# *that* roof was leaking.

(Sahkai, 2015, p. 218)

These distinct properties of the definite article *the* and the demonstrative *that* are considered again later in section 3.5, where Japanese and Korean demonstratives, *sono* and *ku* are compared to English *the* and *that* in terms of their definiteness marking.

3.4.2 Definiteness in article-less languages, Japanese and Korean

In languages without articles such as Japanese and Korean, nouns are bare by default and mostly ambiguous in terms of number and definiteness, as illustrated in (3.16). The nouns *gakusei* and *haksayng* are still interpreted as either definite or indefinite according to the context, but this feature is not overtly marked.

(3.16) a. Gakusei-o mita. (Japanese)

b. Haksayng-ul poassta. (Korean)

student-ACC saw

'(I) saw the/a/some/ø student(s).'

These languages can also express definiteness with post-nominal topic markers, Japanese *wa* and Korean *nun*. However, their cardinal function is to indicate old/given information (i.e., topic) rather than definiteness. On the other hand, new information (i.e., focus) is often marked with nominative markers, Japanese *ga* and Korean *ka*. Topic-marked nominals tend to be definite by being given information whereas nominative-marked nominals, since they are new information, are often indefinite (Heycock, 2008; Kuno, 1973; Kuroda, 1972; Lee, 2001).[14]

However, the correlation between those morphemes and definiteness does not seem very reliable. For example, the pair of sentences in (3.17) from Shibatani (1990) show that the choice between *ga* and *wa* does not necessarily correlate with definiteness.

(3.17) a. Hi-ga noboru.

sun-NOM rise

'The sun rises.'

---

[14] According to Lee and Shimoji (2016), Japanese *wa* and Korean *nun* express distinct notions of new/old information. They argue that in Korean, *nun* marks "episode-old" entities and it can refer back only to an entity previously mentioned in the current episode. In contrast, Japanese *wa* encodes "hearer-old" entities: it often refers to an entity that is not directly mentioned but inferable in the relevant discourse or part of the shared knowledge of the interlocutors. This means that Korean *nun* is distributed more restrictedly than Japanese *wa*, and Korean *ka* can appear in some contexts where *wa* is used but not vice versa.

b. Hi-wa noboru.

sun-TOP rise

'The sun rises.'

(Shibatani, 1990, p. 262)

Both sentences express similar propositions. Crucially, both *hi ga* (3.17a) and *hi wa* (3.17b) are considered definite (i.e., *the sun*) as reflected in the English translation. However, Shibatani argues that they are semantically different and occur in distinct contexts. The sentence in (3.17a) with *ga* is typically used to describe events or states with an implication of some sort of surprise for the discovery or witnessing of an event or state. In contrast, the sentence in (3.17b) indicates a more context-free expression, precisely because its fundamental function is that of conveying a generic statement (Shibatani, 1990, p. 263).

Moreover, it has been often reported that in addition to topic-hood, Japanese *wa* and Korean *nun* can express contrasts, in which case they can be suffixed to focused elements, as illustrated for Japanese in (3.18). As the translation suggests, *wa*-marked nouns can be either definite or indefinite.

(3.18) John-ga pai-*wa* tabeta ga (keeki-*wa* tabenakatta).
John-NOM pie-TOP ate but cake-TOP ate-NEG
'John ate (the) pie, but he didn't eat (the) cake.'

(Heycock, 2008, p. 55)

Another well-known observation is that subject nouns in embedded clauses are marked with nominative markers (Japanese *ga*, Korean *ka*) by default regardless of informational value and they must have a contrastive reading (as opposed to a topic reading) if marked with *wa* or *nun* (e.g., Heycock, 2008; Kuroda, 2005; Vermeulen, 2012). That is, topicality expressed with *wa* and *nun* is argued to be a root phenomenon, except for in the complement clause of some attitude predicates such as *regret*.

Furthermore, information structure seems to interact with word order. It has been proposed that (non-contrastive) topic noun phrases need to be licensed in clause-initial position in Japanese and Korean (e.g., Vermeulen, 2012). The following example dialogues illustrate the point for Japanese (3.19–3.20) and Korean (3.21–3.22) (both dialogues are adapted from Vermeulen, 2012, pp. 86–87). In replying to the requests in (3.19) and (3.21), the object nouns (which are marked with *wa*, *ano boosi-wa* in (3.20a) and with *nun*, *ku moca-nun* in (3.22a)) must occupy clause initial position. The reply is infelicitous when these nominals appear in non-clause initial positions in (3.20b) and (3.22b).

(3.19) Ano boosi-nituite nanika osiete-kudasai.
that hat-about something tell-give
'Tell me something about that hat.'

(3.20) a. Ano boosi-wa$_i$ John-ga kinoo e$_i$ katta.
that hat-TOP John-NOM yesterday bought
b. #John-ga ano boosi-wa kinoo katta.
John-NOM that hat-TOP yesterday bought
'John bought that hat.'

(3.21) Ku moca-etyayhayse mal-hay-po-a.
that hat-about tell-do-try-IMPERATIVE
'Tell me about that hat.'

(3.22) a Ku moca-nun$_i$ John-i ecey e_i sasse.

that hat-TOP John-NOM yesterday bought

b # John-i ku moca-nun ecey sasse.

John-NOM that hat-TOP yesterday bought

'John bought that hat.'

All these observations taken together, the correspondences between topic markers and definiteness and between nominative markers and indefiniteness do not seem very consistent in either Japanese or Korean. However, there is a consistent relationship in both languages between demonstratives and definiteness.[15] The overt marking of anaphoric definiteness by Japanese and Korean demonstratives, *sono* and *ku* will be discussed in the next section as a target property of this study. Another example of realisation of definiteness in Japanese and Korean is a semantic constraint on numeral quantifier (NQ) constructions, which is detailed in section 3.6 as another property under investigation.

3.5 Overt marking of anaphoric definiteness by Japanese *sono* (in comparison to Korean *ku* and English *the* and *that*)

In Japanese, Korean, and many other languages without an article system (e.g., Chinese and Thai), the two types of definiteness are distinguished by marking anaphoricity with demonstratives and realising unique definites in the form of bare nouns (e.g., Ahn, 2017; Jenks, 2015, 2018; Schwartz, 2009, 2013). Cross-linguistically, demonstratives have two different uses: exophoric/deictic use, where the referent is identified by means of demonstration (e.g., pointing at something) in the speech context; and anaphoric/non-

[15] Certain other lexical items, such as possessives and pronouns are also consistently definite, but these are beyond the scope of the present thesis and are not discussed further.

deictic use, where a demonstrative expression refers to a linguistic antecedent in the previous discourse. In article-less languages such as Japanese and Korean, demonstratives cover the anaphoricity-marking function of definite articles. Japanese and Korean both have three-way demonstrative systems as shown in Table 3.1 (based on Ahn, 2017; Hoji, Kinsui, Takubo, & Ueyama, 2003).

Table. 3.1 Comparison of demonstratives in Japanese, Korean, and English

Japanese	Korean	English	Location of referent in deictic use
kono	*i*	*this*	near the speaker
sono	*ku*	*that*	closer to the hearer
ano	*ce*	*that (over there)*	at a distance from both the speaker and the hearer

It has been observed that only the medial demonstratives, Japanese *sono* and Korean *ku* are productively used to mark anaphoricity, with the other demonstratives (i.e., Japanese: *kono* and *ano*; Korean: *i* and *ce*) primarily used deictically (e.g., Hoji, et al., 2003; Yoshida, 2011 for Japanese, Chang, 2009; Cho, 2017 for Korean). Importantly, *sono* and *ku* require a linguistic antecedent (e.g., Hoji et al., 2003; Ahn, 2017), thus they are necessarily anaphoric like the German strong article. These anaphoric demonstratives are not only used when the referent is previously mentioned directly like *the* and *that* (3.23) but also when anaphoricity is established indirectly through, for example, product-producer bridging unlike *that* (3.24) (e.g., Mohri & Isse, 2018; Tsutsumi, 2012 for Japanese, Ahn, 2017; Cho, 2017 for Korean) (*?(x)* means that the sentence sounds more natural with *x*).

(3.23) Direct anaphoric use (b and c are adapted from Ahn, 2017, p. 40)

a. Japanese

Hon	i-ssatu-o	katta.	?(sono)	hon-wa	takaka-tta.
book	one-CL-ACC	bought.	SONO	book-TOP	expensive.was

'I bought a book. The book was expensive.'

b. Korean

Cecyk	han-kwen-ul	sass-ta.	?(ku)	cheyk-un	pissass-ta.
book	one-CL-ACC	bought.	KU	book-TOP	expensive.was

'I bought a book. The book was expensive.

c. English

I bought a book. The/that book was expensive.[16]

(3.24) Anaphoric Bridging: producer-product relation

a. Japanese

Watasi-wa	kyoo	syosetu-o	i-ssatu	katta.
I-NOM	today	novel-ACC	one-CL	bought

Sono	tyosya-wa	furansujin-da.
SONO	author-TOP	French-DECL

'I bought a novel today. The author is French'

16 It has been pointed out in the literature that even though both *the* and *that* are in principle possible in contexts like (23c), *that* is more marked than *the*. For example, Roberts (2002) proposes that a demonstrative implicates a contrast set. That is, a demonstrative suggests that the relevant property attributed to the referent is not satisfied with other members of the relevant contrast set. This means that in the absence of a contrast set, there is no compelling reason for using *that*, thus *the* is more natural than *that*.

b. Korean (adapted from Kang, 2021, p. 316)

Nay-ka	onul	sosel-ul	han-kwen	sassta.
I-NOM	today	novel-ACC	one-CL	bought

Ku	ceca-nun	phulangsuin-i-ta.
KU	author-TOP	French-be- DECL.

'I bought a novel today. The author is French'

c. English

I bought a novel today. *The*/#*that* author is French.

However, because of their anaphoric nature, Japanese *sono* and Korean *ku* are infelicitous in non-anaphoric/unique definite contexts, where the presupposition of uniqueness is established by means of pragmatic knowledge about the situation (3.25), including bridging via a part-whole relation (e.g., Kaneko, 2014; Mohri & Isse, 2018 for Japanese, Chang, 2009; Cho, 2017 for Korean). Thus, in these contexts, *sono* and *ku* are more like English *that* than *the* (Ahn, 2017; Wolter, 2006) (3.26).

(3.25) Situation uniqueness

a. Japanese

Sakuya	(#sono)	tuki-ga	totemo	akaruka-tta.
last.night	SONO	moon-NOM	very	bright-PAST

'The moon was very bright last night.'

b. Korean

Eceypam	(#ku)	tal-i	acwu	palk-assta.
Last.night	KU	moon-NOM	very	bright-PAST

'The moon was very bright last night.'

c. English

The/#that moon was very bright last night.

(b and c were adapted from Kang, 2021, p. 318)

(3.26) Bridging uniqueness: part-whole relation

a. Japanese

Kekkonsiki-ni	itta.	(#sono)	sinpu-ga	ao-o	kiteita.
wedding-to	went	SONO	bride-NOM	blue-ACC	wore

'I went to a wedding. The bride wore blue.'

b. Korean (Cho, 2017, p. 372)

Gyelhonsik-ey	kassta.	(#ku)	sinpu-ka	phalansayk-ul	ipessta
wedding-to	went	KU	bride-NOM	blue-ACC	wore

'I went to a wedding. The bride wore blue.'

c. English (Lyons, 1999, p. 7)

I've just been to the wedding. The bride wore blue.

Note that use of anaphoric demonstratives is infelicitous and bare nominals must be used in non-anaphoric definite contexts. However, in anaphoric contexts, as Cho (2017, p. 372) points out, use of anaphoric demonstratives in article-less languages such as Japanese and Korean are, though strongly favoured, not obligatory. Although the presence/absence of demonstratives does not affect well-formedness of the sentence (unlike the English determiners, lack of which can result in ungrammaticality), bare nominals are often avoided because they can be ambiguous in many ways (e.g., definite, indefinite, generic, and non-generic) whereas use of demonstratives can unambiguously establish anaphoric relationships. However, researchers hold different views about the degree of preference for *sono*-marked nouns over bare nouns and the degree of acceptability of bare nouns in anaphoric bridging contexts. For example, Tsutsumi (2012) argues that use of *sono* is strongly favoured and bare nouns are rather unnatural. However, Mohri and Isse (2018) report that some of their native Japanese consultants seem more permissive of bare nouns in bridging anaphoric contexts than in direct anaphoric contexts.[17]

In sum, Japanese and Korean, despite their lack of articles, seem to distinguish the two types of definites, anaphoricity and uniqueness by marking the former with demonstratives, *sono* and *ku*. Furthermore, these anaphoric demonstratives show both similarities to and differences from English determiners, *the* and *that*: *sono* and *ku* mark anaphoricity similarly to *the* and *that* in direct anaphoric contexts (3.23); however, in bridging anaphoric contexts (3.24), whereas *the*, *sono*, *ku* are felicitous, *that* is not; in non-anaphoric (unique definite) contexts whether situational (3.25) or bridging (3.26), *sono* and *ku* are considered unnatural like *that*, in contrast to *the*. These cross-linguistic differences are summarised in Table 3.2. What remains an empirical question is to what extent overt anaphoric marking is preferred over a covert counterpart (i.e., bare nouns) in bridging contexts.

Table 3.2 Crosslinguistic differences between English determiners and Japanese and Korean demonstratives

		Anaphoric		Unique	
		Direct	Bridging	Situational	Bridging
English	*the*	✓	✓	✓	✓
	that	✓	#	#	#
Japanese	*sono*	✓	✓	#	#
Korean	*ku*	✓	✓	#	#

[17] This asymmetry may be accounted for with Nemoto's (2015) proposal that the felicity of anaphoric definite reading of bare nominals depends on the presence of competing referents. For example, in anaphoric bridging contexts involving product-producer relations, the producer of a product is likely to be unique hence the absence of competition with other potential individuals.

3.6 Covert marking of definiteness: Japanese numeral quantifiers (NQs) and semantic constraint in comparison with the Korean and English NQs

In this section, the other property under investigation, namely a covert realisation of definiteness through word order in NQ constructions is discussed in comparison to relevant Korean and English properties to demonstrate similarities and differences between the three languages. However, before looking into the definiteness property, let us start with other essential properties of numeral classifiers.

3.6.1 Numeral classifier constructions and their essential properties

Japanese and Korean are classifier languages, in which nouns normally cannot combine with a numeral without a classifier (a lexical item representing a unit of measurement) (3.27). Classifiers are bound morphemes that must be appropriate to the semantic type of the noun: human nouns require human classifiers, animate nouns require animate classifiers, and inanimate require inanimate classifiers (3.28). Unlike Japanese and Korean, English nouns with numerals do not require classifiers, as is evident in the translations in (3.28).

(3.27) a. Japanese

gakusei san *(nin)

student 3 (CL)

b. Korean

haksayng sey *(myeng)

student 3 (CL)

(3.28) Japanese/Korean classifiers

a. Human classifier

gakusei	san-nin	(J)apanese
haksayng	sey-myeng	(K)orean
student	3-CL	

'three students'

b. Animate classifier

inu	san-biki	(J)
kangoci	sey-mali	(K)
dog	3-CL	

'three dogs'

c. Inanimate classifier

ringo	san-ko	(J)
sakwa	sey-kay	(K)
apple	3-CL	

'three apples'

One important characteristic of Japanese and Korean is that NQs can appear in different syntactic positions: the pre-nominal position (before the associated noun) (3.29a, 3.30a), the post-nominal position (between the associated noun and the case particle) (3.29b, 3.30b), and the floating (adverbial) position (after the case particle) (3.29c and 3.30c).[18] In English, unlike Japanese and Korean, only a few quantifiers such as, *all*, *both,* and *each* can float (3.31) (Kobuchi-Philip, 2007). English numerals are allowed only in the pre-

[18] Note that the term "float(ing)" conventionally refers to quantifiers separated from the associated noun phrase. The present thesis follows this convention without subscribing to any specific syntactic account. For a comprehensive summary of syntactic derivation accounts of Japanese NQs, see Nakanishi (2008).

nominal position (3.32a). English approximations of the post-nominal and floating constructions are ill-formed (3.32b).

(3.29) Japanese

a. Pre-nominal

Taroo-ga	[ni-ko-no	ringo]-o	tabeta.
Taroo-NOM	[two-CL-GEN	apple]-ACC	ate

'Taroo ate two apples.'

b. Post-nominal

Taroo-ga	[ringo	ni-ko]-o	tabeta.
Taroo-NOM	[apple	two-CL]-ACC	ate

'Taroo ate two apples.'

c. Floating

Taroo-ga	[ringo]-o	ni-ko	tabeta.
Taroo-NOM	[apple]-ACC	two-CL	ate

'Taroo ate two apples.'

(3.30) Korean (adapted from Shin, 2017)

a. Pre-nominal

Cheli-ka	[twu-kay-uy	sakwa]-lul	mekessta.
Cheli-NOM	[two-CL-GEN	apple]-ACC	ate

'Cheli ate two apples.'

b. Post-nominal

Cheli-ka	[sakwa	twu-kay]-lul	mekessta.
Cheli-NOM	[apple	two-CL]-ACC	ate

'Cheli ate two apples.'

c. Floating

Cheli-ka	[sakwa]-lul	twu-kay	mekessta.
Cheli-NOM	[apple]-ACC	two-CL	ate

'Cheli ate two apples.'

(3.31) a. [All students] walked.

b. [The students] all walked.

(3.32) a. John ate two apples. (pre-nominal)

b. *John ate apples two. (post-nominal/floating)

The three types of construction are often used to express the same proposition, as indicated by the English translations of the Japanese and Korean examples in (3.29–3.30). Post-nominal and floating NQs are, in particular, known to have striking similarities in meaning. For example, post-nominal and floating NQs are both subject to the monotonicity constraint (Nakanishi, 2007; see Shin, 2009 for Korean). The monotonicity constraint states that a measure function expressed by a measure phrase must track part-whole relations to the entity that the measure phrase applies to (Schwarzschild, 2002, p. 226). For example, the volume measurement of *three litres of (the) water* is monotonous because it tracks part-whole relations to the entity water: if the quantity of water is three litres, every proper subpart of the water has a volume less than three litres. In contrast, the temperature measurement is non-monotonic. For example, if the temperature of the water is 50 degrees Celsius, its proper subparts will not necessarily have lower temperatures. It has been observed that all types of NQs are allowed for monotonous measurement (3.33) whereas for non-monotonous measures, post-nominal and floating NQs cannot be used (3.34).

(3.33) Monotonous measurement

a. mizu san-rittoru (post-nominal)

water 3-litre

b. mizu-ga/o san-rittoru (floating)

water-NOM/ACC 3-litre

c. (cf.) san-rittoru-no mizu (pre-nominal)

3-litre-GEN water

(3.34) Non-monotonous measurement

a. *mizu san-do (post-nominal)

water 3-degree

b. *mizu-ga/o san-do (floating)

water-NOM/ACC 3-degree

c. (cf.) san-do-no mizu (pre-nominal)

3-degree-GEN water

However, several semantic differences between post-nominal and floating NQs have been observed. For example, it has been pointed out that the floating construction permits a distributive reading but not a collective reading (e.g., Ishii, 1999; Kitagawa & Kuroda, 1992; Nakanishi, 2007) (as briefly introduced in the previous chapter). To illustrate, the post-nominal NQ in (3.35a) allows for the collective reading where three boys built one boat together as well as the distributive reading where each of the three boys built a boat (i.e., three boats in total). However, the floating counterpart in (3.35b) has only the distributive reading.

(3.35) a. Post-nominal (✓collective ✓distributive)

Otokonoko	san-nin-ga	kinoo	booto-o	tukutta.
Boy	3-CL-NOM	yesterday	boat-ACC	made

‘Three boys built a boat yesterday.’

b. Floating (?? collective ✓distributive)

Otokonoko-ga	kinoo	san-nin	booto-o	tukutta.
boy-NOM	yesterday	3-CL	boat-ACC	made

(Nakanishi, 2007, p. 58)

Another widely observed difference between post-nominal and floating NQ constructions is that the floating construction is incompatible with I(ndividual)-level predicates (i.e., predicates denoting permanent properties), although compatible with S(tage)-level predicates (i.e., predicates denoting events or temporary properties); on the other hand, post-nominal constructions can be used with both types of predicate (e.g., Fukushima 1991; Nakanishi, 2007 for Japanese; Shin, 2017 for Korean). In (3.36), the S-level predicate, *byooki-dearu* ‘be sick’ can occur with both types of NQ. In contrast, (3.37) shows that the I-level predicate, *kasikoi* ‘be smart’ is compatible only with the post-nominal variant.

(3.36) S-level predicate

a. Kono	kurasu-de	**gakusei**	**san-nin**-ga	byooki-dearu	(post-nominal)
this	class-in	student	three-CL-NOM	sick	

‘Three students are sick in this class.’

b. **Gakusei**-ga	kono	kurasu-de	**san-nin**	byooki-dearu.	(floating)
student-NOM	this	class-in	three-CL	sick.	

(3.37) I-level predicate

a. Kono kurasu-de [**gakusei san-nin**]-ga kasikoi. (post-nominal)

this class-in [student three-CL]-NOM smart

'Three students are smart in this class.'

b. ??**Gakusei**-ga kono kurasu-de **san-nin** kasikoi. (floating)

student-NOM this class-in three-CL smart

(Nakanishi, 2007, p. 56)

Finally, there is an interesting contrast observed between post-nominal and floating NQs in terms of interpretation of definiteness, which is discussed in the following section.

3.6.2 Definiteness constraint on NQs

It has been pointed out that whereas post-nominal NQs are compatible with both [+definite] and [−definite] interpretations, floating NQs can have only a [−definite] interpretation (e.g., Furuya, 2012; Watanabe, 2006) (3.38). The same contrast has been reported for the Korean counterparts as well (e.g., Lee, 2013; Shin, 2017) (3.39).[19]

(3.38) a. Taroo-wa kinoo hon san-satu-o yonda.

Taroo-TOP yesterday book 3-CL-ACC read

'Taroo read three books/the three books yesterday.'

b. Taroo-wa kinoo hon-o san-satu yonda

Taroo-TOP yesterday book-ACC 3-CL read

'Taroo read three books/# the three books yesterday.'

---

19 Note that the type of definiteness relevant here is that of the maximality presupposition (see section 3.2).

(3.39) a. Cheli-nun [cacenke twu-tay]-lul sa-lyeko ha-n-ta

Cheli-TOP [bicycle two-CL]-ACC buy-intending do-PRE-DEC

'Cheli intends to buy two bicycles/the two bicycles.'

b. Cheli-nun [cacenke]-lul twu-tay sa-lyeko ha-n-ta

Cheli-TOP [bicycle]-ACC two-CL buy-intending do-PRE-DEC

'Cheli intends to buy two bicycles/#the two bicycles.'

(Shin, 2017, p. 21)

However, although researchers agree that post-nominal NQs are felicitous and floating NQs are not in definite contexts, they have different views about the precise range of acceptability of each type of NQ in indefinite contexts with respect to specificity. Specificity is a concept closely related to definiteness that subdivides indefinite nominals. Among several different definitions of specificity proposed in the literature (see von Heusinger, 2011), the one relevant here concerns presupposition of existence, more specifically, partitive specificity or membership of a previously mentioned set (Deising, 1992; Enç, 1991). Some researchers propose syntactic accounts (Furuya, 2012; Ochi & Huang, 2014; Watanabe, 2006), whereas others provide semantic explanations (e.g., Kobayashi & Yoshida, 2001; Nakanishi, 2007; Shin, 2017), each using different theoretical tools. However, for convenience, I will outline these different accounts by dividing them into three groups according to the prediction they make. Table 3.3 schematises three distinct predictions about interpretation of each type of NQ. These will be elucidated below. (Note that the (un)acceptability indicated in the table is taken from the different studies to be outlined in the next few sections and that these studies all rely on introspective data—an issue that I will return to later in section 3.6.2.5.)

Table 3.3 Three different predictions for interpretation of post-nominal and floating NQs[20]

Context	NQ type	Prediction A	Prediction B	Prediction C
[−definite, −specific]	Post-nominal	✓	*	✓
	floating	✓	✓	✓
[−definite, +specific]	Post-nominal	✓	✓	✓
	floating	*	*	✓
[+definite, +specific]	Post-nominal	✓	✓	✓
	floating	*	*	*

Although those groups of studies are largely in agreement (unshaded area), their predictions vary in [−definite, −specific] and [−definite, +specific] contexts regarding the acceptability of post-nominal and floating NQs, respectively (shaded area). Researchers in the Prediction B group (e.g., Downing, 1996; Huang & Ochi, 2014) judge that only floating NQs are felicitous in [−specific] contexts such as (3.40) (where the preamble ensures a [–definite, –specific] interpretation of ‘three books’), whereas those categorised as Prediction A (e.g., Furuya, 2012, Watanabe, 2006) argue that both constructions are possible. For [+specific] contexts (3.41–3.42), the Prediction A group argues that only the post-nominal construction is felicitous; but the Prediction C group (e.g., Kobayashi & Yoshida, 2001; Nakanishi, 2007; Shin, 2017) predicts felicitous occurrence of both constructions in [−definite, +specific] contexts (3.42).

(3.40) [−definite, −specific]

PREAMBLE: Taroo does online shopping almost every day.

Kinoo-wa	{ hon san-satu-o vs. hon-o san-satu }	kaimasita.
yesterday-TOP	{ book 3-CL-ACC vs. book-ACC 3-CL }	bought

[20] Note that under the presuppositionality definition of specificity, the feature combination of [+definite, −specific] is impossible because [+definite] entails [+specific].

‘He bought three books yesterday.’

(3.41) [+definite, +specific]

PREAMBLE: Taroo has two little sisters.

Sensyu	{ imooto huta-ri-o vs. imooto-o huta-ri }	
last.week	{ little.sister 2-CL-ACC vs. little.sister-ACC 2-CL }	

yuuenti-ni	tureteikimasita.
amusement.park-to	took

‘He took the two sisters to an amusement park last week.’

(3.42) [−definite, +specific]

PREAMBLE: Hanako got two PCs and one printer from a friend.

{ pasokon iti-dai-o vs. pasokon-o iti-dai }	sigoto yooni	tsukau-koto	ni	
{ PC 1-CL-ACC vs. PC-ACC 1-CL }	for.work	using	that	

simasita.
decided

‘She decided to use one PC for work.’

Korean NQ constructions also have been provided different accounts whose predictions can be categorised as Prediction A (Kim, 2005; Lee, 2013) and Prediction C (Shin, 2017).[21]

---

[21] However, it is possible to apply Prediction B given the structural similarities and the proposed definiteness effect between Japanese and Korean.

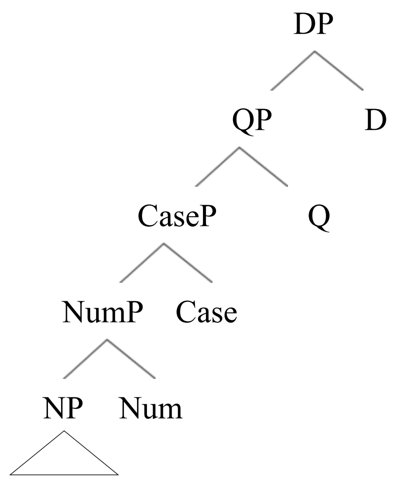
In what follows, I briefly overview representative accounts from each prediction group, using Japanese examples for simplicity.

3.6.2.1 Prediction A: movement analysis

For the Prediction A group, I consider syntactic accounts proposed by Watanabe (2006) and Furuya (2012). These two studies share the assumption that floating NQs are syntactically derived from their corresponding post-nominal NQs by means of movement. However, they differ in terms of what element moves (i.e., NP or NQ movement) and why an indefinite interpretation is forced on floating NQs.

Watanabe (2006) accounts for the structural and interpretive variation in NQ constructions assuming multiple layers of functional projections for Japanese nominals as presented in (3.43).

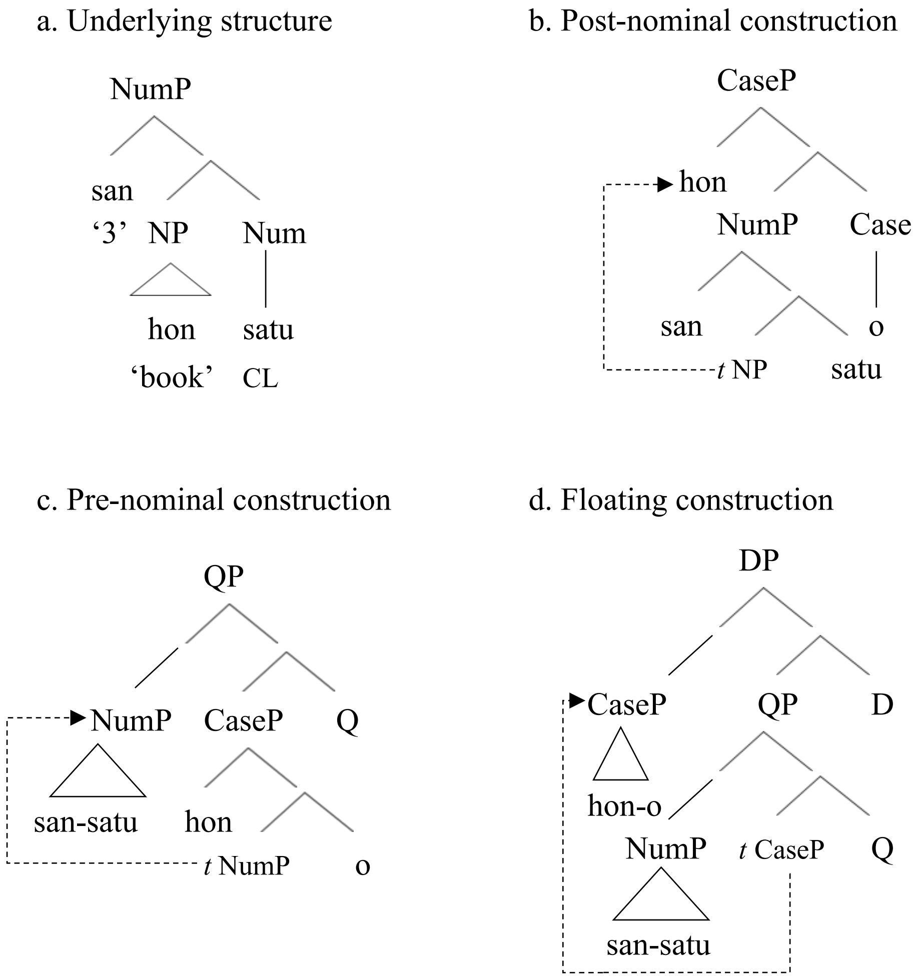
(3.43) Functional projections of nominals in Japanese



(Adapted from Watanabe, 2006, p. 252)

Based on this structure, Watanabe proposes derivational paths for NQ constructions as follows (3.44). Firstly, a classifier (CL) heads NumP taking NP as its complement and a numeral as its specifier (3.44a). Then, after this structure merges with the Case head, NP is obligatorily raised to [Spec, CaseP] for Case feature agreement between N and Case, resulting in the post-nominal NQ construction (3.44b). If NumP moves to the specifier of QP and the derivation ends here, the pre-nominal NQ structure obtains (3.44c). This movement is for the agreement between Num and Q in the feature related to the mass/count distinction. Furthermore, CaseP is raised to [Spec, DP] for an agreement reason between Case and D in terms of specificity as well as case feature, which results in the configuration for floating NQs (3.44d). Watanabe argues that this final movement of CaseP to Specifier of non-specific D forces a non-specific and indefinite reading on floating NQs. Furthermore, according to Watanabe, in the post-nominal NQ construction, both specific and nonspecific readings are possible because the host NP is not required to overtly raise to [Spec, DP] hence underspecified for specificity.

(3.44) Derivation of NQ constructions (based on Watanabe, 2006)



However, one problem to this analysis is that, as Watanabe himself points out, it does not address a well-documented observation that floating NQs cannot occur with NPs with demonstratives. To illustrate, the DP in the pair of sentences in (3.45) contains the demonstrative *kono* 'this', whereby the floating construction (3.45b) is unacceptable under the same interpretation as its corresponding post-nominal construction (3.45a).[22] This phenomenon can be accounted for by an alternative movement analysis proposed by Furuya (2012), which is reviewed next.

---

22 Furuya (2012) notes that the floating NQ construction still allows for a partitive interpretation. That is, (3.45b) is possible for the interpretation that "(I) read three copies of this book."

(3.45) a. Kono hon san-satu-o yonda. (post-nominal)

this book 3-CL-ACC read

'(I) read these three books.'

b. *Kono hon-o san-satu yonda. (floating)

this book-ACC 3-CL-ACC read

Similarly to Watanabe (2006), Furuya assumes that floating NQs are syntactically derived from their post-nominal counterparts. In her analysis, however, it is the NQ rather than the associated NP that is extracted from the host nominal structure. Furthermore, she proposes that the extraction of the NQ is prohibited when the host structure is specific/definite as follows (here I use "definite/specific" because Furuya does not distinguish between definiteness and specificity). Following Campbell (1996), Furuya argues that when the nominal structure (DP) that hosts the NQ and its associated noun is specific/definite, [Spec, DP] is filled with a null referential quantifier or an explicit definite expressions such as a demonstrative. With the further assumption of DP constituting a phase, a derivational unit in the Minimalist syntax (e.g., Chomsky, 2000), the researcher proposes that those elements occupying [Spec, DP] block the extraction of the NQ. That is, taking the NQ out of the DP without going through [Spec, DP] as an escape hatch results in ungrammaticality because such an operation violates the Phase Impenetrability Condition (PIC) in (3.46).

(3.46) Phase Impenetrability Condition (PIC)

Material within a phase XP is not accessible to operations at ZP (the next phase) unless it is within the edge domain of XP.

(Furuya, 2012, p. 40 based on Chomsky, 2004, p. 108)

Accordingly, (3.45b) is ungrammatical because the demonstrative *kono* occupies [Spec, DP], thereby preventing the NQ from floating away from the DP as illustrated in (3.47).

(3.47) a. [$_{DP}$ demonstrative [$_{D'}$ D [NP…[NQ]]

b. *[NQ$_i$ [$_{DP}$ demonstrative [$_{D'}$ D [NP…[t_i]]

(Adapted from Furuya, 2012, p. 34)

In contrast, the NQ can either stay in or move out of the associated nominal structure when the whole DP receives an indefinite/nonspecific interpretation with nothing present in [Spec, DP] (3.48a). However, once a definite/specific interpretation is established for the associated noun phrase through context, a null operator arises and as a result, NQ floating is no longer legitimate (3.48b).

(3.48) a. NQ$_i$ [$_{DP}$ [$_{NP}$ NP t_i]] (indefinite/nonspecific)

b. *NQ [$_{DP}$ null operator [$_{NP}$ NP t_i]] (definite/specific)

(Adapted from Furuya, 2012, p. 41)

Furthermore, under Furuya's analysis, non-floating NQs (i.e., post-nominal NQs) are structurally ambiguous between indefinite/nonspecific and definite/specific constructions as shown in (3.49). In other words, they can be interpreted as either indefinite/nonspecific or definite/specific.

(3.49) a. [$_{DP}$ [$_{NP}$ NP NQ]] (indefinite/nonspecific)

b. [$_{DP}$ null operator [$_{NP}$ NP NQ]] (definite/specific)

Although Furuya does not distinguish definiteness from specificity, she assumes that this is a specificity effect.[23] That is, NQ floating from a specific indefinite DP is predicted to be also impossible. Consequently, the DP associated with a floating NQ must be [−definite, −specific] as in Watanabe's approach. Kim (2005) also proposed a movement account similar to Furuya (2012) to explain the definiteness effect on floating NQs in Korean. In short, studies in this group argue that post-nominal NQs are felicitous regardless of definiteness and specificity whereas floating NQs must be nonspecific, thus infelicitous in [±definite, +specific] contexts.

3.6.2.2 Prediction B: size analysis

As to Prediction B, I focus on Huang and Ochi (2014) because other studies in this category (e.g., Downing, 1996) are, to my knowledge, observational rather than theoretical. Huang and Ochi's approach can be considered another movement analysis as it assumes the same underlying representation for post-nominal and floating NQs as Watanabe (2006). However, it differs from Watanabe in a crucial way, which results in a different prediction for the interpretation of those types of NQ: Huang and Ochi claim that specificity is determined by the size of the host nominal structure. Specifically, under the assumption that "a specific indefinite nominal has a larger structure than an non-

---

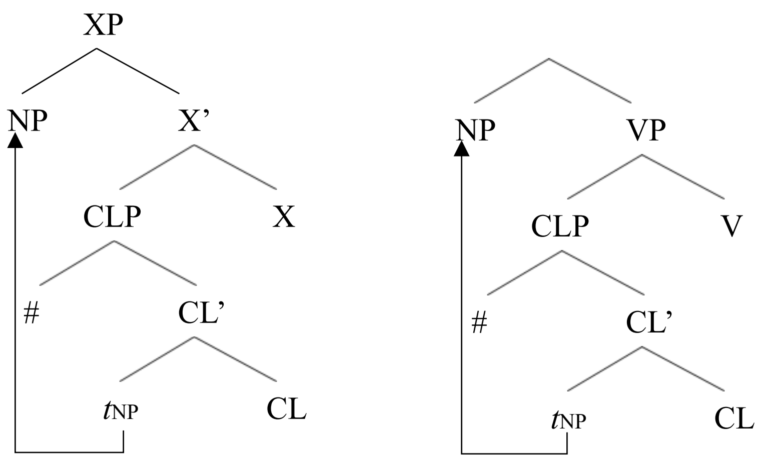
[23] Specificity effects have been observed in other languages including English. As illustrated in (i), it has been observed that in English, an internal element (*wh*-word in this case) can be extracted when the whole nominal structure is indefinite/nonspecific (ia) but cannot when the host nominal is definite/specific as in (ib-c).

(i) a. Who$_i$ did John read [a story of t_i]?
  b. *Who$_i$ did Fred read [the stories of t_i]?
  c. *Who$_i$ did Mary steal [that picture of t_i]?
   (Enç, 1991)

specific indefinite nominal" (Huang & Ochi, 2014, p. 60), they postulate distinct syntactic representations for each NQ construction as illustrated in (3.50).

(3.50) Structural Difference between post-nominal and floating NC constructions (based on Huang & Ochi, 2014)

a. Post-nominal b. Floating



Both types of NQ start off as a classifier (CL) head that selects an NP as its complement. Huang and Ochi follow Watanabe in assuming that CL heads the Classifier Phrase (CLP) and the number phrase (#) is situated in [Spec, CLP]. The NP complement moves within the nominal structure, eventually surfacing as either a post-nominal or floating NQ. The main difference from Watanabe's analysis is the presence of the functional projection, XP above CLP in the post-nominal construction. Huang and Ochi maintain that when XP is projected, the whole nominal phrase must be interpreted as specific. That is, the post-nominal NQ is structurally specific due to the presence of XP, whereas the floating NQ nonspecific due to the absence thereof. Their motivation for the movement in the derivation of each structure is based on the following assumption (3.51).

(3.51) N needs to be visible (i.e., accessible) from outside the extended nominal domain (for the purpose of selection and/or Case)

(Huang & Ochi, 2014, p. 67)

In the case of the post-nominal construction, the NP has to move to the edge of the nominal projection (i.e. spec, XP) in order to be accessible to heads outside the nominal domain (e.g., T, v) and the whole XP serves an argument. On the other hand, even when X does not project above CLP as in (3.50b), (3.51) means that the NP is still required to be outside of CLP. Therefore, it is forced to move into the V domain, resulting in the floating construction.[24] In sum, the prediction on this account is that post-nominal NQs have to be specific (they are infelicitous in [−specific, −definite] contexts) whereas floating NQs have to be non-specific (they are infelicitous in [+specific, ±definite] contexts).

3.6.2.3 Prediction C: floating-NQs-as-adverbial analysis

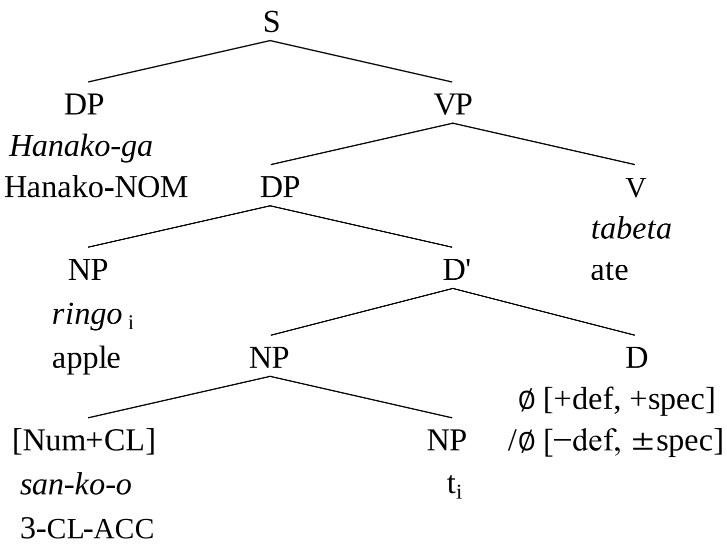
Lastly, the prediction C group consists of studies analysing floating NQs as adverbials and not as elements that are derivatively related to their corresponding post-nominal NQs (e.g., Kobayashi & Yoshida, 2001; Kobuchi-Philp, 2007; Nakanishi, 2007; Shin 2017). In such an approach, the interpretive difference between the two NQ types arises from a difference in composition of the two structures, and they are not interdependent syntactically or semantically. Therefore, two independent different questions arise: why floating NQs are constrained to indefinite contexts and why post-nominal NQs are not. In this section, I will discuss these questions mainly based on Shin (2017) because it is, to my knowledge, the only adverbial account that explicitly addresses the semantic

[24] NP cannot adjoin to the CLP because such a movement is too short, violating the Anti-locality constraint, which prohibits movements within the same phrase.

contrast between post-nominal and floating NQs in terms of definiteness and specificity. However, since Shin's main focus is on Korean NQs, I will also draw on studies of Japanese NQs such as Kobayashi and Yoshida (2001) and Nakanishi (2007), because of their relevance and similarities to Shin's approach.

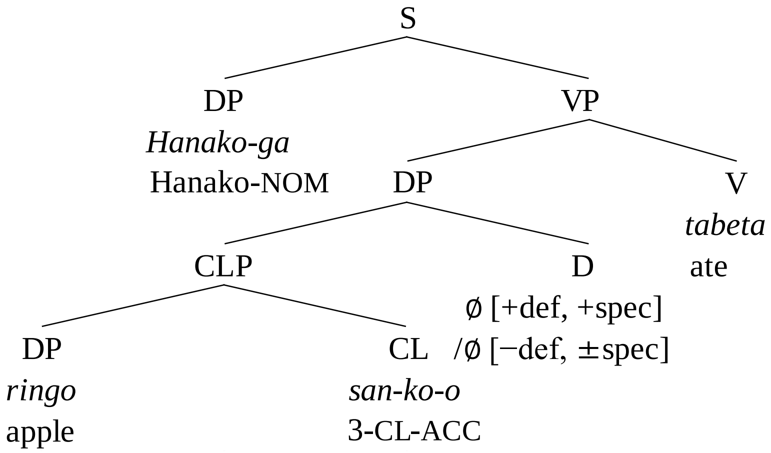
Let us consider the interpretation of post-nominal NQs first. According to the semantic accounts of properties of post-nominal and floating NQs by Nakanishi (2007) and Shin (2017), in a post-nominal NQ structure, the NQ is argued to merge with the associated NP, quantifying individuals denoted by the NP within the nominal structure. On the one hand, Nakanishi assumes that the NQ is adjoined to the NP, and the NP obligatorily moves to [Spec, DP], as illustrated in (3.52).

(3.52) Post-nominal NQ structure (based on Nakanishi, 2007)



Shin, on the other hand, postulates the structure in (3.53), where the numeral classifier, as the head of CLP, takes the associated nominal as its complement.

(3.53) Post-nominal NQ structure (based on Shin, 2017)

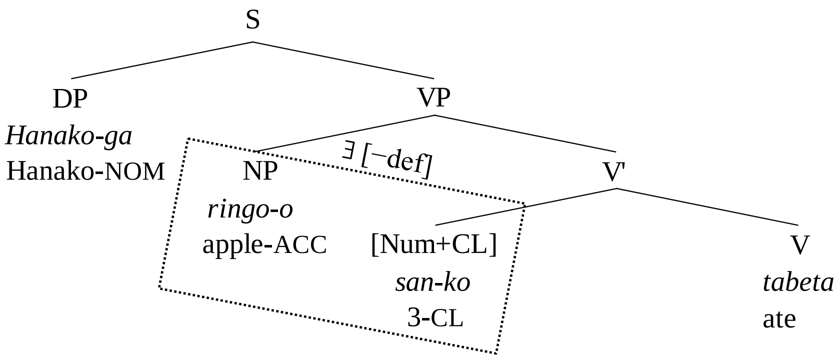


Despite the difference in the syntactic status and position of the NQ, both Nakanishi and Shin locate the NQ and its associated nominal within the DP complement of V. In these structures, the head D (of the uppermost DP) can be either [+definite, +specific] or [−definite, ±specific], depending on the context.[25]

In the floating NQ structure, adverbial analyses in general (e.g., Kobayashi & Yoshida, 2001; Nakanishi, 2007; Shin, 2017) assume that the NQ combines first with the verb, as an adverbial modifier. The associated noun (the direct object for a transitive verb) is merged subsequently, as the internal argument of the verbal predicate, resulting in structures where the NQ and the NP do not form a constituent. Further, when the NP merges with the intermediate V' node containing the NQ, the existential operator is introduced into the semantic representation, to bind a variable related to the internal argument of the verbal predicate. This forces a [−definite] interpretation of the combination of the NQ and NP, as illustrated in (3.54).

[25] Which structure of Nakanishi's (2007) or Shin's (2017) is more appropriate is beyond the scope of the present thesis. However, the choice does not affect arguments in the thesis.

(3.54) Floating NQ structure in the adverbial approach



The semantic representation for the whole sentence is given in (3.55), which is a simplified version of what is presented in Kobayashi and Yoshida (2001) and Shin (2017).[26]

(3.55) ⟦Hanako-ga ringo-o san-ko tabeta⟧ =

$$\exists z\ [z \leq \text{APPLE} \wedge \text{INAN.OBJ}\ (z) \wedge |z| = 3 \wedge [\text{ATE}\ (\text{HANAKO}, z)]$$

Note that the variable z corresponds to the internal argument of the verb and is a subpart of the set denoted by the NP (i.e., APPLE). The exact semantic notation of the set denoted by the NP depends on the context. If there is a previously mentioned (=presupposed) set, the NP is bound by the iota operator (i.e., ιAPPLE), which yields a [−definite, +specific] interpretation (i.e., three (of the) apples). In this case, the semantic representation says that there exists z such that z is a subpart of a contextually salient set of apples and is an inanimate object and the cardinality of z is three and Hanako ate z. If there is no such salient set, the NP refers to a kind or the whole set of what it denotes (i.e., all apples in

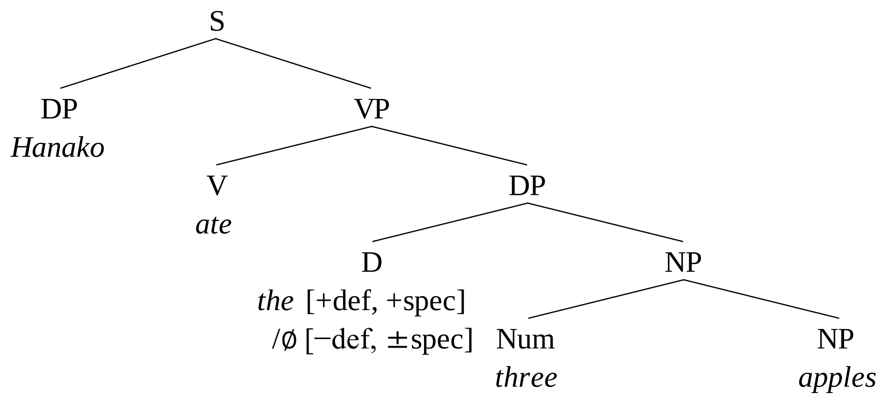
[26] There are as many different semantic representations as there are adverbial accounts. However, they generally assume that the internal argument of the verb is bound by the existential quantifier as in (3.55).

the world) and it is prefixed by the nominalisation operator (i.e., ∩APPLE) (Chierchia, 1998), which results in a [−definite, −specific] interpretation. Crucially, in either case, the sum of individuals introduced by the floating NQ is a subpart of the set denoted by the associated NP, thus remains indefinite. Therefore, floating NQs quantify individuals like their post-nominal counterparts, but they do this through their composition with a verbal predicate, rather than within a DP, which leads to the definiteness constraint in Japanese and Korean.

3.6.2.4 English NQ construction

As to English, NQs are generally assumed to be NP modifiers that adjoin to an NP like adjectives (Krifka, 1999; Verkuyl, 1981). Their definiteness and specificity values are straightforwardly determined by the head D that the combination of the numeral classifier and the NP merges with, like Japanese/Korean post-nominal NQs, as illustrated in (3.56).

(3.56) English numeral construction



3.6.2.5 Which account is the most plausible?

As shown above, different accounts have been proposed regarding the constraint on floating NQs whereby they cannot be definite. Aside from this definiteness constraint, it has been argued that floating NQs are restricted to a distributive reading (as opposed to a collective reading) and incompatible with I-level predicates (as opposed to S-level predicates); however, post-nominal NQs are free of such restrictions. These three constraints are not necessarily expected under either syntactic derivational accounts such as Furuya (2012) and Watanabe (2006) (Prediction A) or the size-difference account by Huang and Ochi (2014) (Prediction B). Adverbial accounts (e.g., Kobayashi & Yoshida, 2001, Nakanishi, 2007; Shin, 2017) (Prediction C) could naturally account for such semantic effects by assuming that post-nominal NQs quantify directly over associate nouns whereas floating NQs first quantify over predicates and then indirectly nouns. For example, researchers taking adverbial approaches (e.g., Nakanishi, 2007; Kobuchi-Philip, 2007; Shin, 2017) explain the incompatibility of floating NQs with I-level predicates (see section 3.6.1) as a consequence of their modifying events denoted by predicates. That is, floating NQs can occur with S-level predicates, which denote events, but not with I-level predicates, which essentially denote temporary states rather than events. Under Kratzer's (1995) assumption that S-level predicates have event arguments whereas I-level predicates lack them, the incompatibility of I-level predicates with floating NQs can be explained by postulating that floating NQs require event arguments. Another problem with the movement approaches (i.e., any accounts assuming a derivational relation between post-nominal and floating NQs) is that there are some quantifiers such as *sorezore* 'each' (3.57), which appear in the floating construction (3.57b) but lack their corresponding post-nominal construction (3.57a) (for more examples, see Nakanishi, 2007, pp. 127–131). That is, if the floating construction is derived from its post-nominal counterpart, the post-nominal construction should also be acceptable.

(3.57) a.	*[Kotosi	sotugyoosuru	gakusei	sorezore]-wa	sangatu-kara
	[this.year	graduate	student	each]-TOP	March-from

syuusyokusaki-de kensyuu-o ukeru

new.job-at training-ACC get

‘Each student who is graduating this year will get training at a new job from March.’

b. Kotosi sotugyoosuru gakusei-wa sangatu-kara sorezore

this.year graduate student-TOP March-from each

syuusyokusaki-de kensyuu-o ukeru

new.job-at training-ACC get

(Nakanishi, 2007, p. 127; (b) is originally from Inoue, 1978, p. 180)

Therefore, the adverbial approach (Prediction C) seems to offer a better coverage of the observed properties of NQ constructions presented above. However, this does not guarantee that Prediction C is a better choice for the definiteness/specificity constraint than the other two predictions. This problem of varied linguistic intuition is nontrivial because any reasonable argumentation becomes unreliable if the data that it is based on are not substantiated by objective evidence. Thus, an experimental study is warranted, to evaluate which of the three different predictions best captures actual linguistic behaviour.

3.7 Summary: properties of overt and covert realisations of definiteness in Japanese by the demonstrative *sono* and numeral classifier constructions

In this chapter, the concept of definiteness and its realisation in English, Japanese, and Korean were overviewed followed by a summary of various proposals about the definiteness-related Japanese properties of linguistic phenomena that this study investigates, namely the overt and covert realisations of definiteness by the demonstrative *sono* and NQ constructions. Before concluding the chapter, let me reiterate the points of agreement and disagreement of previous research.

(3.58) Properties of the demonstrative *sono*

Agreed:

- The use of *sono* is preferred in direct anaphoric contexts over bare forms.
- The use of *sono* is infelicitous in non-anaphoric (i.e., unique) definite contexts.
- The use of *sono* is felicitous in bridging (anaphoric) definite contexts.

Disputed:

- Whether the use of *sono* is preferred over bare forms in bridging contexts.

(3.59) Properties of NQ constructions

Agreed:

- Floating NQs are allowed only in [−definite] contexts.
- Post-nominal NQs are compatible with [+definite] contexts.

Disputed:

- Whether Floating NQs are acceptable in [+specific, −definite] contexts.
- Whether post-nominal NQs are acceptable in [−specific, −definite] contexts.

It has been widely assumed that *sono* marks anaphoric definiteness thus it is incompatible with non-anaphoric (unique) definiteness and although not obligatory, its use is preferred in direct anaphoric contexts. However, introspective judgements from the theoretical research (i.e., non-experimental data) suggest that such a preference may be subject to individual variation among native speakers in bridging anaphoric contexts: some may strongly prefer the use of *sono* over bare nominals in bridging (product-producer relation) contexts, whereas others may consider bare nominals quite acceptable.

The potential contrast between (direct) anaphoric and bridging contexts (i.e., preference for *sono* NPs over bare NPs in anaphoric contexts vs. no such preference in bridging contexts) may not be trivial because it could affect the learnability of the target property in L2 acquisition. For example, the more *sono* is preferred above bare nominals, the easier it may be to learn the feature specification of *sono* because learners should be more likely to find instantiations of *sono* in the relevant context.

As to the properties of the NQ constructions, despite the agreement about the incompatibility of floating NQs with indefinite nonspecific contexts and the compatibility of post-nominal NQs with definite contexts, their precise interpretive properties remain under dispute regarding specificity. If floating NQs are not allowed for a [+specific, −definite] interpretation, then this means that the semantic effect is attributed to non-specificity rather than indefiniteness (Prediction A). Alternatively, if they are compatible with this interpretation, it can be attributed to indefiniteness (Prediction C). If post-nominal NQs turn out to be unacceptable in [−specific, −definite] contexts, then they are must be restricted to [+specific] contexts (Prediction B). Whether the semantic constraint on the NQs is about definiteness or specificity needs formal evaluation before articulating concrete acquisition tasks and predictions within feature-based approaches to L2 acquisition.

Therefore, it is crucial to examine whether the proposed properties are observable in an experimental setting, and if so, which description is the most accurate. One cause of disagreement regarding the properties in question is presumably methodological: researchers build their arguments essentially on informally collected native language intuitions (including their own), incurring the problems of a limited amount of data and potential expert bias (Sprouse, 2015). In order to mitigate such problems, I conducted a series of studies with native Japanese speakers using experimental techniques, including presentation of multiple test sentences and recruitment of non-linguist judges, which are reported in Chapter 6.

Chapter 4: Definiteness in second language

4.1 Introduction

L2 acquisition of definiteness properties has been explored for over the past few decades predominantly in the form of English article acquisition, with many interesting findings. It has been widely observed that L2 English learners from article-less L1 backgrounds tend to have more persistent problems in acquiring English articles compared to those learners whose L1s have articles (e.g., Hawkins et al., 2006; Ionin, Zubizarreta, & Maldonado, 2008; Snape, 2008). For example, Article-less L1 speakers tend to omit L2 articles in obligatory contexts (e.g., Ionin, Ko, & Wexler, 2004; Master, 1997; Trenkic, 2007; Wakabayashi, 1998). They are also known to make substitution errors (e.g., overuse of *the* in contexts where *a* is more appropriate) . However, substitution errors do not seem completely random, but rather they have been argued to be triggered by semantic universals relevant to natural language article choice such as specificity (e.g., Ionin et al., 2004) and presuppositionality (e.g., Ko, Ionin, & Wexler, 2010). Furthermore, it has recently been pointed out that the semantics of L1 demonstratives also affect L2 article acquisition. Specifically, it has been proposed that partially overlapping distributions of L1 demonstratives and L2 definite articles lead learners to misuse or misinterpret L2 articles in contexts where demonstratives are acceptable but not the definite article (e.g., Ionin, Baek, Kim, Ko, & Wexler, 2012). Currently, researchers are beginning to explore L2 acquisition of definiteness properties within the framework of the Feature Reassembly Hypothesis (FRH) (Lardiere, 2008, 2009). For example, some propose to reinterpret misuse of articles attributed to the effect of the semantic universal, presuppositionality as that of L1 demonstrative semantics (Tuniyan, 2018) whereas others investigate a new theoretically interesting learnability contrast between different definite contexts (anaphoric vs. non-anaphoric definiteness) (e.g., Cho, 2017; Feng, 2019). However, in stark contrast to this array of research on L2 acquisition of definiteness in English, little is known about acquisition in the opposite direction: the L2 acquisition of definiteness in article-less languages by article L1 speakers. One of few studies in this domain is Cho and Slabakova (2014), who investigated the acquisition of L2 Russian definiteness expressions by L1 English and Korean speakers. They identified potentially challenging nature of covert expressions of definiteness as opposed to overt counterparts (as reviewed in Chapter 2). Another rare example is Crosthwaite, Yeung, Bai, Lu, and Bae (2018), who examined the acquisition of bridging reference in L2 Mandarin by an article L1 (English) and article-less L1s (Japanese and Korean) speakers. The results suggest effects of L1 article semantics on the acquisition of L2 definiteness marking with demonstratives in article-less languages.

In this chapter, an overview of the research into L2 acquisition of definiteness is provided. Firstly, previous findings on the L2 acquisition of English article semantics will be reviewed. Starting with early studies in the 80's (e.g., Hubners, 1983; Thomas, 1989), this section traces the theoretical development in this research domain, from those studies based on semantic universals pioneered by Ionin et al. (2004) to recent works within the framework of the FRH (e.g., Cho, 2017; Tuniyan, 2018). Then, the acquisition of definiteness in article-less L2s, particularly the acquisition of definiteness marking by means of demonstratives is considered, based primarily on the findings of Crosthwaite et al. (2018). The chapter concludes with discussions of some outstanding questions and their implications for the present study.

4.2 L2 acquisition of English article semantics

4.2.1 Role of semantic universals in L2 article choice

Non-target-like article choice by L2 learners is often attributed to mental representations that are distinct from those of native speakers'. Such an analysis dates back to Huebner (1983), the first to investigate the L2 acquisition of English articles in different semantic contexts, employing Bickerton's (1981) Language Bioprogram Hypothesis. This hypothesis states that article use in natural languages is constrained by two semantic universals: specificity, which concerns whether there is a specific referent in the mind of the speaker's ([±SR]); and hearer's knowledge, which concerns whether a specific referent is in the mind of not only the speaker's but also the hearer's ([±HK]). Huebner tested four different contexts by crossing the two semantic universals: specific definite ([+SR, +HK]), specific indefinite ([+SR, −HK]), nonspecific indefinite ([−SR, −HK]), and generic ([−SR, +HK]), as exemplified in (4.1) (Thomas, 1989, p. 337).

(4.1) Hubners' classification of English article use

a. [+SR, +HK] : specific definite (*the* is used)

Chris approached me carrying a dog. **The dog** jumped down and started barking.

b. [+SR, −HK] : specific indefinite (*a* or ∅ is used)

Chris approached me carrying **a dog**.

c. [−SR, −HK] : nonspecific indefinite (*a* is used)

I guess I should buy **a new car**.

d. [−SR, +HK] : generic (*a*, ∅, or *the* is used)

A paper clip comes in handy.

These examples show that the English definite article is used only in [+HK] contexts. It can be understood that [+SR] asserts that a specific referent exists in the actual world and [−SR] indicates the absence of such a referent, and crucially, [±SR] does not play a role in the English article system. Huebner reported that a Hmong speaker, the target in his longitudinal case study, overused the definite article in indefinite specific ([+SR, −HK]) contexts. Similar errors were observed by a number of researchers who examined the four types of context in (4.1) (e.g., Master, 1987; Parrish, 1987; Thomas, 1989). Thomas (1989) proposed that L2 learners incorrectly associate the feature [+SR] with the definite article rather than [+HK]; and as a result, they misuse the definite article in indefinite specific ([+SR, −HK]) contexts by initially basing their article use on specificity because sensitivity to the [±SR] distinction is innate (Bickerton, 1981).

Building on the above findings, Ionin et al. (2004) proposed a more principled explanation as to why L2 learners make such substitution errors. Specifically, the researchers proposed a semantic UG parameter, the Article Choice Parameter (ACP), which states that languages with two articles distinguish their articles based on either definiteness or specificity. They adopted the informal definitions for each semantic feature in (4.2).

(4.2) If a Determiner Phrase (DP) of the form [D NP] is…

a. [+definite], then the speaker and hearer presuppose the existence of a unique individual in the set denoted by the NP.

b. [+specific], then the speaker intends to refer to a unique individual in the set denoted by the NP and considers this individual to possess some noteworthy property.

(Ionin et al., 2004, p. 5)

Their motivation for the ACP is the crosslinguistic variation that in languages like English, articles are sensitive to definiteness, whereas languages like Samoan distinguish articles based on specificity (Mosel & Hovdhaugen, 1992), as illustrated in (4.3). That is, English uses *the* with definite DPs and *a* with indefinite DPs, regardless of specificity. On the other hand, Samoan marks specific DPs with *le* and nonspecific DPs with *se*, not distinguishing definiteness with the articles.

(4.3) Crosslinguistic variation in natural language article systems

a. Definiteness-based system (English)

<table>
<tr><td></td><td>[+definite]</td><td>[−definite]</td></tr>
<tr><td>[+specific]</td><td rowspan="2">the</td><td rowspan="2">a</td></tr>
<tr><td>[−specific]</td></tr>
</table>

b. Specificity-based system (Samoan)

<table>
<tr><td></td><td>[+definite]</td><td>[−definite]</td></tr>
<tr><td>[+specific]</td><td colspan="2">le</td></tr>
<tr><td>[−specific]</td><td colspan="2">se</td></tr>
</table>

Based on the ACP and the Fluctuation Hypothesis (FH) in (4.4), Ionin et al. postulated that in acquiring a two-way article system, L2 learners would fluctuate between the two settings of the ACP until they fix the parameter value to the one adopted by the target language based on the input.

(4.4) The Fluctuation Hypothesis

a. L2 learners have full access to UG principles and parameter-settings.

b. L2 learners fluctuate between different parameter-settings until the input leads them to set the parameter to the appropriate value.

(Ionin et al., 2004, p. 16)

More specifically, they predicted that in acquiring English articles, L2 learners from article-less L1 backgrounds would go through a stage where they are sensitive to both specificity and definiteness until they have learned that the English article system adopts the definiteness setting. That is, *the* will be used for both specific DPs and definite DPs whereas *a* for both nonspecific DPs and indefinite DPs, resulting in incorrect article choice in [+specific, −definite] contexts and [−specific, +definite] contexts, as summarised in Table 4.1.

Table 4.1 Predictions for L2 English article choice (based on Ionin et al., 2004, pp. 18–19)

	[+definite]	[−definite]
[+specific]	Target use of *the*	Overuse of *the*
[−specific]	Overuse of *a*	Target use of *a*

Ionin et al. (2004) tested these predictions with native speakers of article-less languages, Russian and Korean, using a forced-elicitation test. Participants were asked to read short dialogues and choose one article appropriate to the context, out of three options of *a*, *the* or null (--). Examples of the four relevant semantic contexts are shown in (4.5–4.8) (the correct articles are underlined, and the target DPs are in bold. The same is applied to similar sample stimuli throughout the chapter).

(4.5) [+definite, +specific]

Kathy: My daughter Jeannie loves that new comic strip about Super Mouse.

Elise: Well, she is in luck! Tomorrow, I'm having lunch with (a, <u>the</u>, --) **creator of this comic strip**—he is an old friend of mine. So I can get his autograph for Jeannie!

(4.6) [+definite, −specific]

Bill: I'm looking for Erik. Is he home?

Rick: Yes, but he's on the phone. It's an important business matter. He is talking to (a, <u>the</u>, --) **owner of his company**! I don't know who that person is—but I know that this conversation is important to Erik.

(4.7) [−definite, +specific]

Meeting on a street

Roberta: Hi, William! It's nice to see you again. I didn't know that you were in Boston.

William: I am here for a week. I am visiting (<u>a</u>, the, --) **friend from college**—his name is Sam Brown, and he lives in Cambridge now.

(4.8) [−definite, −specific]

Chris: I need to find your roommate Jonathan right away.

Clara: He is not here—he went to New York.

Chris: Really? In what part of New York is he staying?

Clara: I don't really know, He is staying with (<u>a</u>, the, --) **friend**—but he didn't tell me who that is. He didn't leave me any phone number or address.

Note that each two definite and indefinite contexts differ in terms of specificity, which is operationalised by showing whether the speaker explicitly states some noteworthy property of the referent ([+specific]) (4.5 and 4.7) or not ([−specific]) (4.6 and 4.8).

The results suggested that, as predicted, both Russian- and Korean-speaking L2 English learners overused *a* in [+definite, −specific] contexts and overused *the* in [−definite, +specific] contexts but showed target-like article use in the other two contexts.

Ionin et al., (2004) took this as supporting evidence for the ACP and the FH, although individual results showed that about 40 % of the L2 learners (27/70) behaved in unpredicted ways (i.e., they showed sensitivity only to specificity, or apparently random patterns). Ionin et al. argue that the effect of specificity can persist even in advanced stages, because DPs are often ambiguous between (in) definiteness and (non)specificity— "given the subtlety of the discourse triggers related to speaker and hearer knowledge, generalizing from them is likely to be a fairly long and difficult process" (Ionin et al., 2004, p. 52).

Ionin et al.'s (2004) proposal provides a better empirical coverage than the traditional view of specificity adopted by Huebner (1983) (and the others) as existence in the actual world. That is, based on the existence account, overuse of *the* is predicted both in the indefinite contexts exemplified by (4.7) and (4.8), where the existence of the referent of the target NP is asserted. However, Ionin et al. observed that the L2 learners overused *the* only in contexts considered "specific" by their definition as in (4.7). A number of studies have replicated Ionin et al's findings that L2 learners whose L1 does not have an article system show fluctuation in English article use, affected by both definiteness and specificity (e.g., Ionin, Zubizarreta, & Maldonado, 2008; Snape, 2008).

However, several researchers have contested the ACP and the FH by questioning, for example, the motivation for a UG parameter specific to articles (Hawkins et al., 2006) and the operationalisation of specificity (Trenkic, 2008). On the other hand, Tryzna (2009) reported new evidence about Samoan articles, according to which Samoan distinguishes its articles based on specificity only in indefinite contexts; that is, the specific article, *le* is used in nonspecific definite contexts as well as definite specific and indefinite specific contexts (the nonspecific article *se* is used only in indefinite nonspecific contexts). This calls into question the ACP's prediction for overuse of *a* in [+definite, −specific] contexts.

Furthermore, it has been proposed that another semantic universal, namely presuppositionality also triggers misuse of the definite article in indefinite contexts (e.g., Ko, Ionin, & Wexler, 2010; Ko, Perovic, Ionin, & Wexler, 2008). Pressupositionality can be defined informally as in (4.9).

(4.9) If a Determiner Phrase (DP) of the form [D NP] [+presuppositional], then the speaker and hearer presuppose the existence of at least one individual in the set denoted by the NP.
(Based on Ko et al., 2010, p. 120, which is in turn based on Deising, 1992 and Enç, 1991)

Note that the difference between presuppositionality and definiteness is that presuppositionality marks only the presupposition of existence of a referent but definiteness marks the presupposition of existence of a *unique* referent (Ko et al., 2010, p. 120). There are two ways to establish presuppositionality in indefinite contexts: (i) by previous mention of a set that the referent (denoted by the NP) is a member of, or (ii) by mutual world knowledge. The first subtype of presuppositionality is called "partitivity". To illustrate, *a puppy* in (4.10a) is partitive/presuppositional because its referent is a member of a set introduced in the previous discourse, *five puppies*. In contrast, *a comic book* in (4.10b) is non-partitive/non-presuppositional due to lack of such a previously mentioned set.

(4.10) a. [+partitive/+presuppositional, −definite]
This pet shop had five puppies and seven kittens. Finally, John chose <u>a puppy</u>.
b. [−partitive/−presuppositional, −definite]
Kevin had to memorize two stories and three poems from his textbook. But he

spent the whole evening reading <u>a comic book</u>.

(Ko et al., 2008, p. 121)

As shown in these examples, English articles are not governed by partitivity/presuppositionality. However, a number of studies have reported that L2 learners tend to overuse *the* in [+partitive/+presuppositional, −definite] contexts but not in [−partitive/−presuppositional, −definite] contexts, and importantly, the same is true even when specificity is controlled for (e.g., Ko et al., 2008; Ko et al., 2010). Specifically, it has been found from forced elicitation tasks similar to Ionin et al.'s (2004) that with the target DP held [−definite, +specific], learners incorrectly use *the* significantly more often in [+partitive] contexts (4.11) than [−partitive] contexts (4.12). Based on these findings and similar observations found with children acquiring English as their L1(e.g., Maratsos, 1976), presuppositionality has been argued to be another semantic universal that plays a role in article choice (e.g., Ko et al., 2008; Ko et al., 2010).

(4.11) [−definite, +partitive/+presuppositional, +specific]

Molly: So what did your guest Mr. Svenson do over the weekend?

Jamie: Well, he went to see our local softball team play. He had a good time. Afterwards, he met (<u>a</u>, the, --) **player**—she was very nice and friendly. And she played really well!

(4.12) [−definite, −partitive/−presuppositional, +specific]

Jennifer: Hello, Helen? This is Jennifer!

Helen: Hi, Jennifer! It's wonderful to hear from you. I suppose you want to talk to my sister?

Jennifer: Yes, I haven't spoken to her in years! I'd like to talk to her now if possible.

Helen: I'm very sorry, but she doesn't have time to talk right now. She is meeting with (a, the, --) **very important client from Seattle**. He is quite rich, and she really wants to get his business for our company! She'll call you back later.

It has also been found that the L2 learners are affected by presuppositionality in the form of partitivity even when prior mention of the relevant set is made implicitly (4.13) just as much as when the same noun is repeated (e.g., (4.10a)). (In (4.13), presuppositionality of *a player* is established implicitly with a preceding group noun, *the Boston Red Sox team*.[27])

(4.13) Implicit partitive context

Jane: Your friend Lucy looks really excited. What's going on?

Mary: Well, last Sunday was a really a big day for her. She went to the airport to see her mother off, and ran into the Boston Red Sox team. You know what? She was very lucky—she got an autograph from (a, the, --) **player**. And afterwards, she met some friends at the airport! What a day!

Additionally, Ko et al. (2010) demonstrated that what appears to be the effect of partitivity indeed comes from presuppositionality. Specifically, they showed that the frequency of target-like use of *the* was not significantly different between [+definite, +partitive] contexts, where presuppositionality was established through a previous mention set of which the referent of the target NP was a member (4.14), and [+definite, −partitive]

[27] Yang and Ionin (2009), who tested L2 interpretation of the English articles by intermediate Chinese-speaking learners of English, also reported similar results from an acceptability judgement task: namely, that the learners tend to infelicitously accept *the* in partitive indefinite contexts. However, it was also suggested that implicit partitivity elicited more non-target acceptance of *the* than explicitly partitivity. This may be related to the difficulty involved in bridging contexts, which is discussed later.

contexts, where presuppositionality was satisfied through mutual world knowledge (4.15).

(4.14) [+definite, +partitive]

Sally: I heard that your daughter Karen is a big fan of the Chicago Bears team!

Roger: Yes, she is. She went to Chicago to see them play. And she got a signature from (a, the, --) **head coach**. I have no idea who that is, but Karen was really happy.

(4.15) [+definite, −partitive]

Husband: So who should we invite to dinner this Saturday night?

Wife: How about Alex and Kate?

Husband: No, that won't work. Kate won't be in town—her company needs her to fly west on an assignment. She is meeting with (a, the, --) **governor of Oregon**—you know, I can't remember who that is.

The presupposition of existence of *coach* in (4.14) is established through a prior mention of a set, *the Chicago Bears team*, hence [+partitive] (and [+presuppositional]). In (4.15), on the other hand, the existence presupposition of *governor* is satisfied by the mutual world knowledge that there exists a unique governor in the State of Oregon (*governor* is not [+partitive] but still [+presuppositional]). That is, if learners associated *the* with partitivity rather than just presuppositionality, they would have used *the* even more (4.14) but the results of Ko et al. suggested that they did not. Similarly to specificity, it has been argued that presuppositionality may have a prolonged influence on L2 article choice, due to considerably overlapping distributions of definiteness and presuppositionality as well as a lack of evidence that allows learners to clearly distinguish them. Specifically, if learners associate *the* with presuppositionality, they will correctly use *the* in definite contexts because all definite DPs are [+presuppositional], which can result in difficulty in abandoning their incorrect assumption that *the* marks presuppositionality (Ko et al., 2008, p. 127).

To sum up this section, it has been proposed in the previous research that such semantic universals as definiteness, specificity, and presuppositionality are made available by UG as candidate features for article distinction in natural languages; consequently, if learners do not have an article system in the L1, their L2 article choice is influenced by those features until they have learned which feature is relevant for the target article system. Although it may still be disputed whether there exists a semantic parameter (such as the ACP) to govern articles, it seems that L2 article semantics is ultimately attainable, and a key to successful acquisition is the kind of input that helps learners to distinguish between those competing features. In other words, the main problem may not lie in the acquisition of relevant features per se but rather in the difficulty in teasing them apart from one another and correctly associating the target feature with L2 articles.

However, L2 learners seem to face a further obstacle in acquiring target-like L2 article choice: subtle differences in meaning and function between definite articles and demonstratives, which is discussed next.

4.2.2 Influence of L1 demonstratives on L2 article semantics

As discussed in Chapter 3, many article-less languages have demonstratives that semantically and functionally overlap with –but do not exactly match– definite articles. Not surprisingly, some studies suggest that L1 demonstrative semantics transfers onto L2 definite articles, resulting in non-target-like use and interpretation. For example, Robertson (2000) found from his collaborative communication task that L1-Chinese (an article-less language) L2-English learners overused demonstratives (i.e., *this* and *that*) in contexts where the definite article, *the* is more appropriate as in (4.16).

(4.16) *This square size* is eight cm, er . . .

(Robertson, 2000, p. 71)

Subsequently, Ionin et al. (2012) systematically investigated the L2 interpretation of the definite article *the* vs. the demonstrative *that* by native speakers of Korean (another article-less language). As discussed in Chapter 3, *the* and *that* share the core semantics of uniqueness (definiteness), but they slightly differ in terms of how uniqueness must be established: *the* requires the referent to be unique in the whole discourse whereas *that* must be unique in the immediately salient situation. Ionin et al. predicted that due to the absence of L1 articles, Korean-speaking learners of English would initially map the semantics of the L1 demonstrative (*ku*) onto the L2 definite article, applying the immediately-salient condition rather than the whole-discourse condition to the computation of uniqueness with *the*. That is, they were predicted to infelicitously accept or use the definite article (*the*) when the demonstrative (*that*) is more appropriate in the L2. The researchers tested the prediction using two tasks: a forced elicitation production task and a picture-based comprehension task. In the production task, participants were asked to read a story and fill a gap by choosing the most appropriate answer from four options: *the*, *that*, *a*, and *one*. Three experimental conditions were tested, as exemplified in (4.17). When the referent described by the target noun is unique in the whole discourse as in (4.17a) and (4.17b), *the* is felicitous and preferred over *that*. The demonstrative, *that* is allowed if the referent is salient in the immediate discourse as in (4.17a) and (4.17c), and favoured over *the*, when uniqueness is not established in the whole discourse (4.17c).

(4.17) a. Unique and salient: both *the* and *that* possible, but *the* preferred

Betsy was staying at a hotel, and didn't have anything to read. It was too early to go to bed. So she went to a bookstore, and bought a magazine. Then she came back to her hotel and read ____ magazine. She enjoyed it a lot.

b. Unique and non-salient: *the* preferred over *that*

Vicky was getting ready for a long train trip, and she wanted something to read on her trip, so she went to the library, and got out a book and a new magazine, and packed them in her bag. The next day, Vicky got on the train. She found her seat and sat down. Then, she read ____ book. It was really interesting.

c. Non-unique: *that* preferred over *the*

Richard went to a bookstore and bought two books to read. One of the books turned out to be long and boring. The other book had a really exciting storyline. So, Richard finished ____ book. He read it in just one night.

(Ionin et al., 2012, pp. 79–80)

The results suggested that L2 learners correctly chose the definite article when it is an appropriate choice, namely in the unique and salient condition (4.17a), and the unique and non-salient condition (4.17b). However, learners, particularly in the lower proficiency group, incorrectly produced *the* when the definite article is not felicitous, namely in the non-unique category (4.17c). Based on these results, the researchers proposed that L2 learners have not fully acquired the appropriate discourse condition in terms of uniqueness (definiteness) for the definite article (i.e., uniqueness in the whole discourse) and as a result, they still apply the condition for the demonstrative (i.e., uniqueness in the immediately salient discourse). They also found that the L1-Korean L2-English learners were not sensitive to the distinction between the definite article and the demonstrative in comprehension. In the picture-based comprehension task, the participants viewed pictures of objects and drew geometric shapes on the objects. Each item started with a sentence that named objects followed by three lines that instructed participants to draw geometry shapes, as shown in (4.18). Items were presented in two conditions with the difference in the second instruction: the target NP was preceded by the demonstrative, *those*, or by the definite article, *the*.

(4.18) Ionin et al.'s (2012) picture-based comprehension task

Here are six pens and six balloons.

1. Please draw arrows above two balloons.
2. Now, please draw triangles around { those/the } balloons.
3. Now, please draw stars on two pens.

(Adapted from Ionin et al., 2012, p. 86)

In the demonstrative condition, native English controls responded to the command by drawing triangles around the two objects above which arrows had been drawn following the previous command, and so did the L2 learners. However, in the definite condition, whereas the native control group drew triangles around the two objects about 60 % of the time and the all six objects otherwise, the L2 learners almost exclusively exhibited the two-objects interpretation, regardless of proficiency (intermediate or advanced). Given that the L1 demonstrative *ku* only allows for the two-objects interpretation like English *those* and unlike English *the*, which is also compatible with the all-objects interpretation, Ionin et al. attributed the L2 learners' non-native-like interpretive preference to the transfer of the L1 demonstrative semantics onto the L2 article.

In sum, the results of Ionin et al. (2012) suggest that L2 learners from article-less L1 backgrounds are influenced by the L1 demonstrative semantics in acquiring L2 articles. More precisely, what is subject to transfer seems to be the discourse-based condition of how the core meaning of uniqueness (definiteness) is computed according to the context.

4.2.3 L2 acquisition of definiteness expressions within the Feature Reassembly Hypothesis

In recent years, researchers have started to apply the FRH to the L2 acquisition of definiteness phenomena (e.g., Cho, 2017; Cho & Slabakova, 2014; Tuniyan, 2018). In this section, I will review representative studies that investigate L2 English article semantics within the FRH.

Cho (2017) is the first study to adopt the FRH for the investigation of L2 English article acquisition. As detailed in Chapter 2, according to the FRH, L2 acquisition proceeds as learners map features of L1 lexical items onto L2 counterparts based on perceived similarities. Crucially, difficulties may arise when there are mismatches in feature specification between relevant L1 and L2 items. In such cases, some feature reassembly must be done by means of available input. Cho targeted L1-Korean L2-English learners, considering the following crosslinguistic difference: both the Korean demonstrative, *ku* and the English definite article, *the* express [+definite] but they differ in terms of whether the referent needs a linguistic antecedent ([+anaphoric]) or not ([−anaphoric]) (as detailed in Chapter 3). Following the FRH, Cho predicted that due to the similarity in definiteness marking, Korean-speaking English learners would first infelicitously map the features of the L1 demonstrative, *ku*, [+definite, +anaphoric] onto the L2 definite article, *the*. Consequently, learners were expected to be more native-like in contexts where both *ku* and *the* are felicitous ([+definite, +anaphoric]) than contexts where *the* is felicitous but *ku* is not ([+definite, −anaphoric]). However, with increased input and proficiency, learners would be able to adjust the initially mapped incorrect feature set to the target representation ([+definite, ±anaphoric]).

In order to test these predictions, Cho (2017) administered an acceptability judgement task, where Korean-speaking English learners with two different proficiency levels (intermediate and advanced) read pairs of sentences and rated the acceptability of the target sentence as a continuation of the first, using a 4-point scale. There were four experimental conditions: three different anaphoric contexts and one non-anaphoric context, as exemplified in (4.19–4.22) (Cho, 2017, p. 376). In direct anaphoric contexts (4.19), the second-mention referent (*cake*) should be in the form of definite NP (*the cake*). The same goes for taxonomic anaphoric contexts (4.20), in which the antecedent (*a dessert*) is mentioned by means of a different NP (*the cake*). Anaphoric reference can be more indirect in bridging definite contexts (4.21), in which the definite expression in the second sentence refers back to the implicit antecedent (*a cake* is an implicit argument of the verbal predicate, *baked*). Finally, non-anaphoric bridging definite contexts are where the relevant referent does not have an antecedent, but its definite interpretation can be established situationally. For example, *birthday* in (4.22) sets up a situation that allows for bridging definite inference for *the cake* through the general world knowledge that there usually exists one unique cake involved in a birthday. Therefore, *the cake* is more acceptable than *a cake*, which implies that there was, though implausibly, more than one cake associated with the birthday.

(4.19) Direct anaphoric definite context

a. Jackie made <u>a cake</u> for the party. She served <u>the cake</u> with coffee and tea.

b. Kevin ordered <u>a cake</u> from the grocery store. #He went to pick up <u>a cake</u> but it was not ready.

(4.20) Taxonomic anaphoric definite context (different head noun antecedents)

a. Lydia's family purchased a dessert. They ate the cake after dinner.

b. Marianne and her daughters shared a dessert. # They enjoyed a cake.

(4.21) Anaphoric bridging definite context (implicit antecedents)

a. Tori baked for her office this morning. Her co-workers enjoyed the cake.

b. Rachel baked for her husband. #He enjoyed a cake.

(4.22) Non-anaphoric bridging definite context (no antecedent)

a. It was Sophie's first birthday. She smashed the cake with her hands.

b. Patrick celebrated his birthday with his friends. #They enjoyed a cake.

The results showed that the intermediate group rated definite NPs significantly higher than indefinite NPs in the three anaphoric definite contexts but not in the non-anaphoric definite contexts, suggesting the predicted L1 influence. However, the advanced group rated definite NPs higher than indefinite NPs in non-bridging (direct) anaphoric contexts but not bridging contexts (either anaphoric or non-anaphoric). Cho argues that this suggests that the advanced learners have correctly reassembled the features of the definite article yet have trouble in "accommodating unmentioned propositions for bridging definites" (Cho, 2017, p. 367). That is, to license bridging definites, the hearer needs to accommodate "the implied link between the bridging description and its anchor (i.e., the element that the bridging description is related to)" (Cho, 2017, p. 379). Cho speculates that the non-target-like performance of the advanced learners resulted from learners' accommodation of context, which made the use of indefinites less unnatural.[28]

Feng (2019) conducted a replication study of Cho (2017) with Mandarin-speaking intermediate and advanced learners of English. Like Korean, Mandarin does not have an article system, but, according to Feng, its demonstrative, *nei* expresses the same feature composition as the English definite article, *the* ([+definite, ±anaphoric]), allowing for anaphoric and non-anaphoric bridging (4.21–4.22). Feng conducted an AJT with the identical set of stimuli used in Cho and reported that the intermediate group had difficulty in non-anaphoric bridging, whereas the advanced group showed overall native-like judgements. Feng argues that the asymmetry between the Korean-speaking learners in Cho's study and Mandarin-speaking cohort in her study is due to the crosslinguistic difference that the Mandarin demonstrative has the same feature specification as the English article unlike the Korean counterpart. That is, Mandarin speakers may have acquired the properties of *the* faster because they do not have to do feature reassembly as opposed to Korean speakers, who need to do so. As to why the intermediate learners had difficulty in non-anaphoric bridging contexts, Feng proposes that establishing bridging reference without a potential antecedent (i.e., non-anaphoric) but only with pragmatic knowledge at the semantics-pragmatics interface is challenging to L2 learners even when L1 positive transfer is expected.

A next relevant study is Tuniyan (2018), who investigated L2 English article acquisition by L1 Mandarin and Russian speakers (two L1 Mandarin groups: intermediate and advanced, and three L1 Russian groups: beginning, intermediate, and advanced).[29] A novelty of Tuniyan's work is that she attempts to account for the presuppositionality

---

[28] This means that those advanced learners may have overaccommodated contexts. Although it may be debatable whether this is completely non-target-like, Cho (2017) considers it non-target-like because the native control group (the target) did not show such a behaviour.

[29] Recall that some findings of Tuniyan (2018) were already discussed in the relation to the cline of difficulty in Chapter 2. Here I will mainly focus on the aspects that have not been discussed yet.

effect on L2 article choice documented in the previous studies (e.g., Ko et al., 2010) as feature transfer from the L1 demonstrative onto the L2 definite article. Like Cho (2017), which follows Schwartz (2009), Tuniyan assumes that the English definite article encodes two types of definiteness: familiarity ([familiar]) and uniqueness ([unique]) ([+definite, +anaphoric] and [+definite, −anaphoric] in Cho, respectively). Furthermore, Tuniyan proposes that these two types of definiteness covary with the meaning of anaphoricity, as presented in (4.23) and (4.24).

(4.23) Different meanings of definiteness (Tuniyan, 2018, p. 93)

a. Familiarity: An NP is familiar if the hearer already has the mental representation of the intended referent through the previously mentioned most salient direct antecedent (*anaphoric familiarity*) or through the presence of the perceptually most salient antecedent (*non-anaphoric familiarity*).

b. Uniqueness: An NP is unique if a unique referent for the hearer exists in a given situation based on the unique part-whole relation with the previously mentioned indirect antecedent (*anaphoric uniqueness*) or through general knowledge that a given situation contains only one unique referent (*non-anaphoric uniqueness*).

(4.24) Different meanings of indefiniteness (Tuniyan, 2018, p. 94)

a. Non-familiarity: An NP is non-familiar if it refers to a new referent for the speaker and the hearer or to a referent that is known to the speaker but unknown to the hearer (*non-anaphoric non-familiarity*)

b. Non-uniqueness: An NP is non-unique if it refers to a non-unique referent through a non-unique member-set/part-whole relation with the previously mentioned direct/indirect antecedent (*anaphoric non-uniqueness*)

Based on these characterisations, Tuniyan postulated that the English definite article encodes [+familiar, ±anaphoric] or [+unique, ±anaphoric] whereas the indefinite article encodes [−familiar, −anaphoric] or [−unique, +anaphoric].[30] Furthermore, assuming that both Russian and Mandarin's demonstratives can express familiarity (i.e., [+familiar, ±anaphoric]) like the English definite article, Tuniyan predicted that L2 English learners with L1 Mandarin and L1 Russian would initially map the feature of the L1 demonstratives onto the L2 definite article. Specifically, those L2 learners were expected to particularly associate anaphoricity ([+anaphoric]) with the English definite article on the grounds that the L1 demonstratives are typically used in anaphoric contexts. Consequently, L2 learners from article-less language backgrounds are predicted to overuse the definite article in [−unique, +anaphoric] contexts, where the indefinite article is appropriate on the one hand, and underuse it in [+unique, −anaphoric] (i.e., obligatory) contexts, on the other. The researcher tested these predictions, using an AJT and a written sentence production task (see Chapter 2 for more information). Six conditions were tested as exemplified in (4.25–4.26) for the AJT (three each for the definite and indefinite conditions). In [+familiar] or [+unique] contexts, *the* is felicitous and *a* is not, irrespective of anaphoricity (4.25), and the opposite is true in [−familiar] or [−unique] contexts (4.26).

---

30 Note that Tuniyan uses the term anaphoricity differently from Cho (2007) and Schwartz (2009). In Cho and Schwartz, anaphoricity means familiarity, which is the type of definiteness established through anaphoric relations whether directly or indirectly (bridging). In these studies, bridging definiteness expressed situationally, for example, through a part-whole relation (e.g., *car*-*engine*) is not considered anaphoric, but non-anaphoric (unique) definite. This issue is discussed in detail later in section 4.4.2.

(4.25) Definite conditions (adapted from Tuniyan, 2018, p. 151)

a. Previous mention: [+familiar, +anaphoric]

Mary often goes shopping, and last Friday she went to a new shopping mall. She bought a bag there, and she was very happy. She used **the/#a** bag straight away.

b. Unique bridging: [+unique, +anaphoric]

Michael likes going out, so he often goes to parties. Last Saturday he went to a wedding, and he had fun there. He even danced with **the/#a** bride.

c. Out-of-the-blue definite: [+unique, −anaphoric]

Patrick went camping last summer, but one night he could not fall asleep. He got up, and he did not know what to do. So he watched **the/#a** sky for a while.

(4.26) Indefinite conditions (adapted from Tuniyan, 2018, p. 153)

a. Partitive: [−unique, +anaphoric]

Betty decided to get a kitten, so she went to a pet shop. The pet shop had five kittens, and she played with them for a while. Then she chose **a/#the** kitten.

b. Non-unique bridging: [−unique, +anaphoric]

Alex is a photographer, and last Saturday he worked at a big wedding party. It was a long day, and he got bored being by himself. So he talked to **a/#the** guest for a while.

c. Out-of-the blue indefinite: [−familiar, −anaphoric]

Aaron is a policeman, and last night he was at work. He was tired, and he fell asleep. When he woke up, he was surprised. He saw **a/#the** mouse in his office.

That is, learners were predicted to infelicitously accept/use *the* more in the [−unique, +anaphoric] contexts (4.26a–b) than the [−familiar, −anaphoric] contexts (4.26c), and reject/underuse it more frequently in the [+unique, −anaphoric] contexts (4.25c) compared to the [+familiar, +anaphoric] and [+unique, +anaphoric] contexts (4.25a–b). These predictions were partially supported across the tasks. As predicted, the L2 learners were generally target-like in the [+familiar, +anaphoric] (i.e., previous mention) and [−familiar, −anaphoric] (i.e., out-of-the-blue indefinite) contexts, except for the beginner L1 Russian group, who were non-target-like in all the contexts. However, no groups showed less target-like acceptance/use of *the* in the out-of-the-blue definite contexts ([+unique, −anaphoric]) compared to the anaphoric familiar/unique contexts, against the predictions. Furthermore, what is potentially a L1-related difference was observed. On the one hand, Russian-speaking learners often incorrectly used/accepted *the* in the anaphoric indefinite contexts (both partitive and non-unique bridging) significantly more than the non-anaphoric indefinite contexts. On the other hand, the L1-Mandarin groups also showed a similar tendency but only in non-unique bridging contexts. As to why the L1 Mandarin learners were affected by anaphoricity in non-unique bridging but not in partitive indefinite contexts, Tuniyan attributed this difference to an observation about Mandarin from her small-scale experimental study that partitive indefinite nouns ([−unique, +anaphoric]) are often overtly marked with an unstressed numeral + classifier (CL), *yi* CL rather than being left bare, whereas there is no such preference in non-unique bridging contexts. That is, this tendency towards overt realisation might have a facilitative effect in partitive indefinite contexts, in line with the cline of difficulty (as detailed in Chapter 2).[31]

Finally, Zhang (2020), building on Ionin et al. (2012), investigated whether L1

[31] Although it is possible that the L1 overt realisation had a facilitative effect, this analysis seems questionable because according to Tuniyan's experimental data from native Russian speakers, in Russian, overt marking with the numeral *odin* is favoured over covert marking with bare nouns in both partitive indefinites and non-unique bridging anaphoric: that is, it remains to be explained why Russian-speaking learners did not benefit from the L1 overt marking preference as Mandarin-speaking counterparts might have.

Mandarin speakers could acquire subtle contrasts between the English definite article, *the* and the demonstrative, *that* in terms of semantics-pragmatics mappings or conditions for discourse-based expression of definiteness. As mentioned earlier, Mandarin does not have an article system but has a demonstrative, *na* 'that', which shares the function of signifying definiteness as the English definite article. Zhang administered the same forced elicitation task as used in Ionin et al. to advanced and near-native English learners with L1 Mandarin in order to examine whether those L2 learners could acquire native-like discourse-sensitive preference of one determiner over the other in English. The researcher focused on the contexts where the definite article is preferred over the demonstrative (i.e., the referent is unique in the whole discourse) (4.17b), repeated as (4.27a), ('Whole' scenario, henceforth), and vice versa (i.e., the referent is salient in the immediate discourse) ('Salient' scenario, henceforth) (4.17c), repeated as (4.27b). Additionally, a different group of native Mandarin speakers were tested with a Mandarin equivalent of the elicitation task to confirm how similarly and differently to the English determiners the L1 demonstrative is distributed. The relevant contexts were compared, using *na* + NP vs. bare NP contrast in place of *the* + NP vs. *that* + NP contrast in the English version.

(4.27) a. 'Whole' scenario: *the* preferred over *that*

Vicky was getting ready for a long train trip, and she wanted something to read on her trip, so she went to the library, and got out a book and a new magazine, and packed them in her bag. The next day, Vicky got on the train. She found her seat and sat down. Then, she read ____ book. It was really interesting.

b. 'Salient' scenario: *that* preferred over *the*

Richard went to a bookstore and bought two books to read. One of the books turned out to be long and boring. The other book had a really exciting storyline. So, Richard finished ____ book. He read it in just one night.

(Ionin et al., 2012, p. 80)

The results showed that the English native controls for the English version of the task, as predicted, almost unanimously preferred *the* in the 'Whole' scenario and *that* in the 'Salient' scenario (≥98%), indeterminately accepting the disfavoured determiner (44–46%). As to the L2 English learners, the advanced group did not show a target preference in either scenario whereas the native-like group showed a target preference of *that* in the 'Salient ' scenario but not a target preference of *the* in the 'Whole' scenario. On the other hand, the results of the native Mandarin controls suggested that both *na* + NP and bare NP were acceptable but *na* was preferred (91.2% vs. 70.0%) in the 'Whole' scenario. However, in the 'Salient' scenario, the bare NP option was rejected (accepted only 2.2%) and the acceptance rate of the demonstrative option was below chance level (36.4%). Zhang explains these results as follows. Firstly, judging from the native English and Mandarin controls' data, the demonstrative *na* can be considered a Mandarin equivalent of the English definite article *the*, since both were highly accepted in the 'Whole' scenario but only indeterminately accepted in the 'Salient' scenario. Furthermore, no Mandarin determiner seems to exist that corresponds to English *that*. However, the partially overlapping distribution of *na* and *that* as well as *the* possibly led Mandarin-speaking learners to map the properties of *na* onto *the* and *that,* namely conditions for expression of definiteness that are discourse-sensitive (semantics-pragmatics mappings), along the lines of the FRH. This means that when either of the licencing conditions for *the* or *that* is satisfied, L2 learners are expected to allow both *the* and *that*, in other words, incorrectly accept *that* in the 'Whole' scenario and *the* in 'Salient' scenario, as observed with the advanced learners. However, the native-like group's partially target-like preference suggests that with increased proficiency and input, they may be able to acquire the target semantics-pragmatics mappings for definiteness marking. Particularly, given the target-like asymmetry in acceptance of *the* between the 'Whole' and 'Salient' scenarios, they seem to have acquired the semantics-pragmatics mapping of *the* first, due to a positive L1 influence: the demonstrative *na* is constrained in a similar way to *the*. In contrast, due to the lack of a lexical item with the same semantics-pragmatics mapping as *that*, they will continue to infelicitously accept *that* in the 'Whole' scenario as a negative L1 effect. Zhang concluded that "convergence at the semantics-pragmatics interface is not impossible for L2 learners, but may be constrained by asymmetries in the L1–L2 realisation of semantics-pragmatics mappings." That is, L1 transfer in semantics-pragmatics mappings may facilitate or hinder the L2 acquisition of corresponding properties.[32]

In sum, L2 research into definiteness properties has started shifting to the FRH framework. Some studies (Cho, 2017; Feng, 2019) found relatively more target-like use of the definite article in a given definite context (i.e., anaphoric) than another (i.e., non-anaphoric), which cannot seem to be easily accounted for as the effect of semantic universals. There is also an attempt to reinterpret some L2 article substitution errors, namely overuse of the definite article in indefinite contexts, which has been claimed to be the effect of the semantic universal, presuppositionality within the FRH as an effect of L1 demonstratives (Tuniyan, 2018). Furthermore, it has been suggested that the lexical transfer from the L1 demonstrative onto the L2 definite article, proposed in earlier studies (e.g., Ionin et al., 2012) is compatible with the FRH: transfer of semantics-pragmatics mappings between the relevant L1 and L2 lexical items (Zhang, 2020).

[32] Tuniyan's (2018) observation mentioned in Chapter 2 that her L2 learners were more target-like in contexts where the demonstrative and the definite article behave similarly than where they behave differently also echoes Zhang's findings.

4.3 Definiteness in article-less L2s: influence of L1 article semantics on L2 demonstratives

Despite the vast amount of research into the L2 acquisition of definiteness properties in article languages (predominantly English), there has been little exploration of article-less L2s. Notable exceptions are Cho and Slabakova's (2014) investigation of L2 Russian (reviewed in Chapter 2), and Crosthwaite et al.'s (2018) investigation of L2 Mandarin. This section reviews Crosthwaite et al. (2018), who investigated definite discourse reference in L2 Mandarin by L1 speakers of English, Japanese, and Korean.

Crosthwaite et al.(2018) built on Crosthwaite (2014), who investigated use of bridging reference in unique definite contexts in L2 English of L1 Korean and Mandarin speakers, using a picture sequence narrative task. Crosthwaite's experimental data from native speakers of English, Mandarin, and Korean showed that these three languages adopt different syntactic strategies to mark uniqueness in bridging reference.[33] English uses the definite article and the indefinite article to mark uniqueness and non-uniqueness in bridging contexts, respectively. In Mandarin, non-uniqueness is encoded through a combination of numeral + classifier before the noun, whereas uniqueness is consistently realised covertly in the form of bare nouns. In Korean, nouns are, in principle, bare whether the referent is unique or non-unique. L2 learners' data from the English task suggested that Mandarin-speaking learners acquired unique bridging reference in L2 English at lower proficiency levels than Korean-speaking counterparts. Crosthwaite attributed this to L1 difference, namely the contrast between syntactic distinction of uniqueness in Mandarin and the lack thereof in Korean. That is, grammaticalisation of

[33] Crosthwaite (2014) uses the notion of inferability rather than uniqueness. That is, in unique bridging contexts, the definite article indicates that a unique referent is inferable whereas the indefinite article the opposite. However, here I use the term, uniqueness to replace Crosthwaite's inferability for convenience.

uniqueness in the L1 may have facilitated Mandarin-speaking learners in acquiring the grammaticalisation of the corresponding property in their L2 English. Crosthwaite et al. (2018) investigated the L2 acquisition of the same property as Crosthwaite (2014) but in the opposite direction, targeting L2 Mandarin learners from article and article-less L1 backgrounds, namely English and Korean/Japanese. The researchers examined how those learners mark definiteness in their L2 Mandarin by using a revised version of the oral picture sequence narrative task used in Crosthwaite (2014). In Japanese and Korean, unique bridging relations are expressed covertly through bare nouns as in Mandarin whereas they are encoded overtly with the definite article in English (as detailed in Chapter 2).

The results suggested that English-speaking learners were more likely to infelicitously use demonstrative + NPs more than native speakers of Japanese and Korean in contexts where native speakers of Mandarin rarely used demonstrative + NPs but mostly used bare NPs instead, as exemplified in (4.28) (*jiaoshi* 'classroom' and *laoshi* 'teacher' are considered unique in the presented setting and mentioned for the first time).

(4.28) Introductions of definite reference (in the case of a 'school' setting)

Tamen	liahoulai	jiu	huıdao	**jiaoshi**	li	yiqı	ting	**laoshi**
They	both later	then	return	**classroom**	inside	together	hear	**teacher**

'They both later returned to the classroom to hear the teacher together.'

The researchers accounted for these results as crosslinguistic effects. That is, the similarity between the L1 and the L2 may have triggered positive influence for the Korean- and Japanese-speaking learners, on the one hand, and the difference between the

L1 and the L2 led to negative transfer for English, on the other.[34] Specifically, they proposed that L1-English learners of Mandarin used the L2 demonstrative as a de facto L1 definite article.

Therefore, Crosthwaite et al. (2018)'s findings suggest that crosslinguistic transfer from the L1 definite article onto the L2 demonstrative is possible. Although their study was not framed within the FRH, this transfer in the opposite direction from what has been discussed in the previous section could also naturally be explained as L1 transfer from the perspective of the FRH.

4.4 Discussion

Now I will discuss some outstanding questions regarding the following two topics that have particularly significant implications for the present study: (i) L2 learnability of bridging reference and L1 transfer, and (ii) semantic categorisation of different types of definiteness.

4.4.1 L2 learnability of bridging reference and L1 transfer

As mentioned in Section 4.2.3, it has been argued that bridging may be an independent source of difficulty in the L2 computation of definiteness (e.g., Cho, 2017; Feng, 2019). Since the present study also involves the L2 acquisition of bridging definiteness, namely bridging with the Japanese demonstrative, *sono*, the following two questions should be addressed before acquisition tasks and predictions can be formulated as to the relative

[34] However, Crosthwaite et al. (2018) showed that Japanese and Korean learners may not be completely target-like in non-unique bridging contexts. They reported that those learners often marked non-unique bridging relations with the form [numeral + classifier + noun], the target form used for (non-unique) non-bridging relations. Since Japanese and Korean use bare nouns in such contexts, this cannot be attributed to the L1s.

difficulty with which to acquire the relevant property.

(4.29) Question 1: Is bridging generally difficult for L2 learners to acquire?

(4.30) Question 2: Are there any crosslinguistic effects in acquiring bridging definiteness?

Within the main framework of the present study, namely the FRH, L1 transfer is expected in terms of Question 2 but no particular challenge is anticipated in terms of Question 1. However, if the answer to Question 1 is affirmative, then the challenging nature of bridging could be a confounding factor in examining crosslinguistic effects subsumed under the FRH (or any other full transfer account, such as Sprouse, 2006).

Starting with the question of whether bridging is generally difficult for L2 learners, it seems to depend on the context. Recall that although the advanced learners in Cho (2017) showed non-native-like judgement in bridging definite contexts in general (anaphoric or non-anaphoric), Feng (2019) found that her learners had difficulty only with non-anaphoric bridging. These results suggest that although bridging may be difficult, because of, for example, the trouble of establishing the implied relation between the bridging referent and its anchor (Cho, 2017), not all bridging contexts are equally challenging (possibly due to L1 positive transfer, as discussed next). Specifically, whereas non-anaphoric bridging may pose a persistent challenge due to its extra computational cost imposed at the semantics-pragmatics interface along the lines of Feng, anaphoric bridging may not necessarily do so. This is compatible with Cho's observation that even the intermediate learners showed native-like judgements in anaphoric bridging contexts as they did in direct anaphoric contexts.[35] Meanwhile, some studies suggest that even

[35] Cho (2017) argues that this is a case of positive transfer of the features of the L1 demonstrative [+definite, +anaphoric] (section 4.2.3). However, the point here is that the potential challenge posed

non-anaphoric bridging contexts are not necessarily problematic, either. For example, consider the results of Ko et al. (2010) and Tuniyan (2018). Although these studies did not explicitly focus on bridging contexts, they did test L2 article choice in non-anaphoric (unique) definite bridging contexts (i.e., (4.14) and (4.25b)). The participants in these studies did not particularly show a sign of difficulty in those contexts. In sum, bridging definiteness seems potentially difficult to acquire for L2 learners, particularly in non-anaphoric contexts; however, some studies have reported generally target-like performance in even non-anaphoric contexts. Therefore, the difficulty of bridging contexts might covary with other factors such as task effects.

Turing to Question 2, there are a number of studies that suggest L1 transfer in L2 bridging contexts. For example, the intermediate groups in Cho (2017) and Feng (2019) showed target-like performance in contexts where the L1 demonstrative can be used for bridging, namely anaphoric definite bridging. These can be considered positive L1 influence. Furthermore, Crosthwaite (2014) reported another interesting difference between L2 English learners from different article-less L1 backgrounds regarding bridging contexts. Specifically, the L1 grammaticalisation of bridging uniqueness (i.e., L1 Mandarin) may help the acquisition of the same property in the L2 whereas the lack thereof in the L1 may result in a delayed acquisition (i.e., L1 Korean). L1 influence has also been suggested in the L2 acquisition of languages that do not have articles or do not obligatorily mark definiteness. Specifically, Crosthwaite et al. (2018) documented what appears to be L1 influence in the non-native Chinese of Korean, Japanese, and English speakers regarding unique (non-anaphoric) bridging. That is, correspondences between the L1 and L2 in terms of the relevant bridging contexts (i.e., L1 Korean and Japanese) resulted in facilitative effects, whereas mismatches thereof may lead to persistent

by anaphoric definite bridging seems to get overridden by such a facilitative effect relatively easily compared to the difficulty with non-anaphoric definite bridging.

difficulties (i.e., L1 English).

4.4.2 Semantics of different definiteness

Now let me discuss how different definite contexts should be semantically categorised by comparing approaches by Ko et al. (2010), Cho (2017), and Tuniyan (2018). In what follows, I will first compare Ko et al. (2010) with Tuniyan (2018) and then Cho (2017) with Tuniyan (2018). The first comparison mainly concerns overuse of the definite article in indefinite contexts, and the second the distinction between different bridging contexts.

The main difference between Ko et al. (2010) and Tuniyan (2018) lies in what is assumed to cause misuse of the L2 definite article in indefinite contexts. Ko et al. argue that it is because L2 learners associate presuppositionality with the definite article whereas Tuniyan contends that it is anaphoricity rather than presuppositionality. They make similar predictions in many contexts. For example, they both predict misuse of the definite article in indefinite contexts where a set that includes the referent denoted by the target noun has been already introduced ([−definite, +partitive, (+presuppositional)] in Ko et al. (4.10a); [−unique, +anaphoric] in Tuniyan (4.26a–b)). However, they offer different predictions in terms of whether there is a contrast between unique bridging and out-of-the-blue definite contexts ([+definite, +partitive, (+presuppositional)] (4.14) vs. [+definite, −partitive, (+presuppositional)] (4.15) in Ko et al.; [+unique, +anaphoric] (4.25b) vs. [+unique, −anaphoric] (4.25c) in Tuniyan). That is, on Ko et al.'s presuppositionality account, target-like use of *the* is predicted in both contexts equally (hence no contrast) whereas on Tuniyan's anaphoricity account, target-like use is expected more in unique bridging contexts.[36] Therefore, some form of difficulty with out-of-the-blue definite contexts is expected on Tuniyan's account but not Ko's. However,

[36] However, unique bridging contexts may be challenging for a different reason as discussed in the previous section.

this prediction has not been borne out. In fact, the unpredicted target use of *the* by lower proficient learners in the out-of-the-blue definite ([+unique, −anaphoric]) contexts found in Tuniyan is more compatible with Ko et al.'s account. In sum, since no definitive evidence has been provided to support Tuniyan's anaphoricity account against Ko's presuppositionality account, it can be said that the latter is, though tentatively, empirically more plausible.

Meanwhile, Cho (2017) and Tuniyan (2018) categorise bridging definite contexts in distinct ways, using different characterisations of anaphoricity. Which approach is more plausible on theoretical and empirical grounds? Firstly, let me clarify how different the concept of anaphoricity is between Cho and Tuniyan. Cho closely follows Schwarz's (2009) division of definiteness by assigning distinct feature sets to anaphoric (i.e., familiar definite) and non-anaphoric (i.e., unique definite) bridging. In contrast, Tuniyan seems to treat both types of bridging as anaphoric, since she treats unique (non-anaphoric in Cho's terms) definite bridging as [+unique, +anaphoric] ((4.22) and (4.25b)). Although Tuniyan did not examine what corresponds to anaphoric definite bridging in Cho's terms, this type of bridging should be expressed as [+familiar, +anaphoric] in Tuniyan's approach. That is, the main difference concerns the categorisation of unique bridging: it is [−anaphoric] in Cho but [+anaphoric] in Tuniyan. Another difference between Cho (2017) and Tuniyan (2018) lies in what features are assumed to transfer from the L1 demonstrative onto the L2 definite article. Cho proposes that L2 learners associate the whole feature set of the L1 demonstrative, [+definite, +anaphoric] with the L2 definite article. However, Tuniyan argues that they transfer only [+anaphoric] on the grounds that the L1 demonstratives are typically used in [+anaphoric] contexts.[37] Therefore, Cho predicts more target-like performance in familiar definite

[37] Tuniyan (2018) also assumes that demonstratives represent the feature set of [+familiar, ±anaphoric], unlike Cho's (2017) set of [+definite, +anaphoric]. Tuniyan presents (i) as an illustration

bridging contexts than in unique definite bridging contexts whereas Tuniyan does not expect such a contrast.

One theoretical problem with Tuniyan's (2018) proposal is that it does not seem straightforward enough why learners map only [+anaphoric] from the L1 demonstrative onto the L2 definite article. Put differently, it seems unclear why [+familiar (=+definite)] is not subject to transfer. On the other hand, Cho analyses the L1 demonstrative feature set as [+definite, +anaphoric] and assumes the features to be transferred all together. Although Tuniyan argues that L2 learners tend to associate the L2 definite article with [+anaphoric] because the L1 demonstrative is most typically used in anaphoric contexts, this does not explain why [+familiar (+definite)] does not transfer together with [+anaphoric]. Since demonstratives invariably express [+familiar (+definite)], it seems more plausible to assume that this feature also transfers, in line with Cho (2017). Furthermore, Tuniyan's feature analysis fails to capture the crosslinguistic division between the two types of bridging, namely non-anaphoric and anaphoric bridging (in Cho's terms), which are attested in a number of languages as shown in Chapter 2. Additionally, Cho's account has gained more empirical support than Tuniyan's. That is, the results of studies testing both types of bridging such as Cho (2017) and Feng (2019) show that learners seem more target-like in familiar bridging contexts than unique

---

of use of demonstratives in [+familiar, −anaphoric] contexts, where the referent is salient through the presence of the perceptually most salient antecedent (as in the definition of familiarity by Tuniyan in (4.23)).

(i) Visible situation with more than one referent, one of which is more perceptually salient
That/the cat is hungry.
(Tuniyan, 2018, p. 105)

This appears to be a case of deictic/exophoric use, which typically has to be accompanied by a pointing gesture of some sort (Chapter 3). Since Tuniyan did not examine this context experimentally, and the present study only concerns anaphoric use of demonstratives, following Cho and Schwartz (2009), I assume [+familiar (+definite), +anaphoric] as the target feature. It is beyond the scope of the present thesis whether demonstratives should be analysed as [+definite, ±anaphoric] or [+definite, +anaphoric].

bridging contexts (section 4.2.3). This supports Cho's account rather than Tuniyan's. Therefore, both theoretically and empirically, Cho's proposal seems to be a better choice than Tuniyan's.[38]

In conclusion, the most reasonable way to go about approaching L2 definiteness seems to be a combination of Ko et al. (2010) and Cho (2017). That is, whereas misuse of L2 articles in indefinite contexts is due to the effect of semantic universals, the contrast between definite contexts in terms of use of the L2 definite article is attributed to the transfer of L1 demonstratives, specifically [+definite, +anaphoric] in Cho's terms.

4.5 Implications for the present study

To conclude, let me discuss some key implications that the previous findings and the discussions above have for the present study.

Firstly, the previous findings suggest that the FRH is an appropriate framework for exploring the L2 acquisition of definiteness marking by demonstratives in article-less languages. Specifically, it seems possible that article L1 learners of an article-less L2 transfer the features of the L1 definite article onto L2 demonstratives due to their semantic and functional similarities, as predicted by the FRH and suggested in Crosthwaite et al. (2018). Therefore, the FRH is a promising approach to the L2 acquisition of definiteness-related properties of the Japanese demonstrative, *sono*. In particular, L1 transfer (whether positive or negative) may occur in acquiring definite marking by the demonstrative, including its bridging definiteness properties.

Secondly, bridging may be generally challenging to L2 learners, and this can

[38] Like Tuniyan (2018), Ko et al.'s (2010) presuppositionality account does not predict an asymmetry between familiar/anaphoric vs. unique/non-anaphoric bridging contexts in terms of target use of the definite article because both are definite thus presuppositional by definition.

be an obstacle for research within the FRH including the present study. If, as Cho (2017) argues, bridging cause difficulty for L2 learners in general, then learners may have trouble in bridging contexts whether anaphoric or non-anaphoric even with corresponding properties in the L1 (e.g., L1 Korean L2 Japanese). Crucially, it is possible that facilitative L1 transfer predicted by the FRH is overridden by the pragmatics-based challenge posed by bridging, resulting in non-target-like performance regardless of the L1. Anaphoric definiteness marking with the demonstrative, *sono,* a target property in the present study, also allows for bridging. The main purpose of the present study is not to investigate L2 acquirability of bridging, but rather to examine whether L2 learners develop their knowledge of the demonstrative in the way predicted within the FRH (i.e., whether they are constrained by the L1 transfer). Therefore, in order to reduce the risk of obscuring potential L1 differences, the present study does not examine non-anaphoric (unique) bridging contexts, which seem particularly difficult for L2 learners (e.g., Cho, 2017; Feng, 2019).

Another thing to consider in testing bridging contexts in article-less L2s is the choice of methodology. Recall that what seems to be positive L1 transfer has been found from the production data presented by Crosthwaite et al. (2018), a study focusing on bridging in an article-less L2 (Mandarin). That is, Japanese- and Korean-speaking learners showed target-like production of bare nouns in Mandarin unique definite bridging contexts, as opposed to English-speaking counterparts, who did not. However, this kind of target-like performance should be interpreted with caution. Although the relevant data suggests that learners know unique bridging can be expressed with bare nouns, it does not constitute evidence that they also know the form, [demonstrative + noun] is infelicitous. For example, they may simply have preferred bare nouns in general, because bare nouns are the default form in their L1s. In order to rule out such a possibility, knowledge about infelicity should also be tested, and to that end, judgement tasks will be more informative. Therefore, the present study uses an acceptability judgement task as a primary measure for tapping into the knowledge about contexts in which the Japanese demonstrative is (in)felicitous.

Finally, as to the characterisation of different definite contexts in terms of features, Cho (2017)'s semantic analysis seems empirically and theoretically sounder than Tuniyan's (2018). Therefore, following Cho, the present thesis adopts the feature specifications given in (4.31) for different determiners, and (4.32) for bridging contexts:

(4.31) Demonstratives: [+definite, +anaphoric]
Definite articles: [+definite, ±anaphoric]

(4.32) Familiar bridging: [+definite, +anaphoric]
Unique bridging: [+definite, −anaphoric]

What has been discussed in this section feeds into the design of the present study (to be described in detail in Chapter 6) and will be revisited in light of the experimental results (in Chapter 9).

Chapter 5: The present study: acquisition tasks and predictions

5.1 Introduction

In this chapter, acquisition tasks and predictions for relative ease/difficulty are described for each target Japanese property and for L2 learners from each L1 background (i.e., English and Korean). The Japanese properties under investigation are reiterated in (5.1) and (5.3) with corresponding or related Korean and English properties in (5.2) and (5.4). Note that these are the generally agreed properties of *sono* and the Japanese NQs discussed in Chapter 3.

(5.1) Target properties of the Japanese demonstrative *sono*

- *Sono* must be anaphoric thus cannot be used to mark non-anaphoric (i.e., unique) definiteness.
- *Sono* can be used in indirectly anaphoric (i.e., bridging) contexts, as well as directly anaphoric contexts.

(5.2) English and Korean properties relevant to the Japanese properties in (5.1)

- The Korean demonstrative *ku* has the same properties as Japanese *sono* in (5.1).
- The English demonstrative *that* can be used in directly anaphoric contexts like *sono* and *ku*, but cannot be used for bridging anaphoric contexts.
- The English definite article *the* can be used in definite contexts whether the referent is anaphoric (including bridging) or non-anaphoric.

(5.3) Target properties of Japanese NQ constructions

- The Japanese floating NQ construction must be indefinite, but the post-nominal NQ construction can be either indefinite or definite.

(5.4) English and Korean properties relevant to the Japanese properties in (5.3)

- The Korean floating and post-nominal NQ constructions have the same properties as their Japanese counterparts in (5.3).
- English does not allow floating or post-nominal NQ construction (, only allowing the pre-nominal NQ construction).

Assuming these properties as given (the Japanese properties in (5.1) and (5.3) have indeed been attested in experimental settings, as shown in the next chapter), I will describe acquisition tasks and predict relative ease/difficulty of those tasks, on the basis of the Feature Reassembly Hypothesis (FRH) (Lardiere, 2008, 2009) and the cline of difficulty in feature acquisition (Cho & Slabakova, 2014; Slabakova, 2009). Although more nuanced predictions will be provided later in Chapter 6 in light of the findings of experimental studies with native Japanese speakers regarding the disputed properties (discussed in Chapter 3), the predictions will remain essentially similar to the way they are presented in this chapter.

In what follows, acquisition tasks and their relative ease/difficulty in the targeted learner groups (L1-Korean and L1-English) are first outlined focussing on each phenomenon. Then, after a brief discussion of some confounding factors, the chapter concludes with overall predictions for the research questions presented in Chapter 1.

5.2 Overt definiteness marking with the demonstrative *sono*

5.2.1 Acquisition tasks

Starting with the properties of *sono*, as stated in Chapter 4, I assume with Cho (2017) and Schwartz (2009) that anaphoric and non-anaphoric (unique) definiteness encode the feature sets of [+definite, +anaphoric] and [+definite, −anaphoric], respectively. Therefore, feature specifications of the Japanese, Korean, and English lexical items relevant in this thesis (i.e., *sono*, *ku*, *the*, and *that*) can be summarised as in Table 5.1. Note that the availability for bridging use is indicated as the [bridging] feature: [−bridging] means that bridging use is not possible whereas [±bridging] means that bridging use is optionally available (i.e., possible in bridging contexts).[39]

Table 5.1 Crosslinguistic differences between English determiners and Japanese and Korean demonstratives

Lexical item		Feature specification
Japanese	*sono*	[+definite, +anaphoric, ±bridging]
Korean	*ku*	[+definite, +anaphoric, ±bridging]
English	*the*	[+definite, ±anaphoric, ±bridging]
	that	[+definite, +anaphoric, −bridging]

According to the FRH, L2 learners will map the features of the L1 lexical items onto the L2 counterparts on the basis of perceived semantic and functional similarities. If there is a mismatch in feature specification between the relevant L1 and L2 items, they will have to reconfigure the initially mapped feature set into the target by means of available input. Thus, the feature overlaps between the lexical items in Table 5.1 may lead to L1-to-L2

[39] The notation of “[±bridging]” is only for convenience, namely for the purpose of describing distinct properties of the Japanese/Korean demonstratives and the English determiners. In other words, there is no intention of proposing that such a pragmatic feature exists.

mappings as elucidated below. (Since the cline of difficulty makes the same assumption as the FRH in terms of L1-to-L2 feature transfer, the description of acquisition tasks based on the FRH below also applies to the cline of difficulty.)

Korean-speaking learners are predicted to map the features of *ku* onto *sono*. Once this initial mapping task is completed, no subsequent change to the L1-based features is required (i.e., feature reassembly is unnecessary), since *ku* and *sono* seem to be syntactically and semantically alike (as detailed in Chapter 3).

As to English-speaking learners, however, the situation seems more complicated, given the similarities between *sono* and *the* on the one hand, and between *sono* and *that*, on the other (as shown in Chapter 3). Specifically, there are at least three possibilities as follows:

(5.5) Possible scenarios for L1-English learners' acquisition of *sono*

- Scenario 1: The features of *the* are mapped onto *sono*.
- Scenario 2: The features of *that* are mapped onto *sono*.
- Scenario 3: The features of *the* and *that* are both mapped onto *sono*.

Scenario 1 is where L1-English learners associate the features of *the* with *sono* presumably due to the similarities between these two lexical items in anaphoric definite contexts (i.e., both can be used anaphorically whether directly or indirectly anaphoric). The features of *the* ([+definite, ±anaphoric , ±bridging]) may be chosen over the features of *that*, ([+definite, +anaphoric, −bridging]) because the positive input relevant to *sono* ([+definite, +anaphoric, ±bridging]) may prevent the learner from associating *that* with *sono* due to the feature clash in terms of [bridging] (i.e., [−bridge] of *that* clashes with [±bridge] of *sono*); in contrast, [±anaphoric] of *the* can accommodate [+anaphoric] of *sono*, hence no clash. Furthermore, although *sono*, like *that*, is [+anaphoric], the absence of *sono* in [−anaphoric] contexts would not provide strong evidence for learners to assume that *sono* is restricted to [+anaphoric]. An empirical motivation for this mapping is Crosthwaite et al.'s (2018) observation that their English-speaking learners of L2 Mandarin appeared to (infelicitously) equate the L2 Mandarin demonstrative (*nei*) with the English definite article. Additionally, Cho (2017) suggests transfer of the features of the L1 (Korean) demonstrative onto the L2 (English) article (see Chapter 4). Despite being in the opposite learning direction (i.e., article-less L1 to article L2 rather than article L1 to article-less L2), this can be taken as supporting evidence for crosslinguistic mappings between demonstratives and definite articles. In Scenario 1, the relevant reassembly task for English-speaking learners is to constrain the [anaphoric] feature, from [±anaphoric] to [−anaphoric].

In Scenario 2, English-speaking learners initially map the features of *that* onto *sono*, motivated by the [+anaphoric] feature shared by them. L2 instruction may also play a role in biasing learners towards this mapping. Although anaphoric use of *sono* (especially in bridging contexts) is not usually taught in Japanese language education, its exophoric or deictic use is typically introduced at a relatively early stage (confirmed by several Japanese language teachers and an inspection of widely used Japanese textbooks).[40] Moreover, in instruction or textbooks targeting English-speaking learners, deictic *sono* is commonly treated as a translation equivalent of English *that*.[41] Such pedagogical practices may lead learners to map onto *sono* the features of *that* ([+definite, +anaphoric, −bridging]) rather than of *the*. The reassembly task in this case would be to

---

[40] The consulted textbooks were the *Genki* series (Banno, Ohno, Sakane, Shinagawa, & Takashiki, 1999) and the *Minna-no-Nihongo* 'Japanese for All' series (3A Network, 1998).

[41] For example, one of the most widely used textbooks for English-speaking beginning to intermediate learners of Japanese, *Genki* series (Banno et al., 1999) introduces deictic *sono* in its second lesson of the first volume (p. 32), where it is simply translated as *that*.

extend the [bridging] feature, from [−bridging] to [±bridging].

Finally, in Scenario 3, learners map the features of both *the* and *that* onto *sono* presumably due to the perceptually equivalent degrees of semantic and functional similarities of *the* and *that* with *sono*. However, there are at least two theoretically conceivable subcases of Scenario 3 as follows:

(5.6) Subcases of Scenario 3:

- Scenario 3A: Learners create a single lexical entry for *sono* and map the features of both *the* and *that* onto it.
- Scenario 3B: Learners create two lexical entries for *sono*, mapping the features of *the* onto one and the features of *that* onto the other.

On the behavioural level, in either case, learners are expected to initially allow *sono* where either *the* or *that* is possible (i.e., in definite contexts including bridging, whether it is anaphoric or non-anaphoric). Empirical support for such behaviour comes from, for example, Ionin et al.'s (2012) finding that Korean-speaking English learners apparently tend to map the meaning and function of the L1 demonstrative (*ku*) onto both *the* and *that* (see Chapter 4). On the conceptual level, however, the two subcases can be distinguished as illustrated below.

In Scenario 3A, initially mapped onto the single lexical entry would be the union of the feature sets of *the* and *that*, namely [+definite, ±anaphoric, ±bridging] (the same as Scenario 1). The relevant feature reassembly task would be then the restriction of the feature [±anaphoric] to [+anaphoric] (again, the same as Scenario 1). In Scenario 3B, on the other hand, two lexical entries are created, one with the features [+definite, ±anaphoric, ±bridging] (i.e., *the*) and one with the features [+definite, +anaphoric,

−bridging] (i.e., *that*). Scenario 3B could be further divided into two paths after this initial mapping stage. In order to complete the acquisition of *sono*, learners are, in theory, required to select one of the two lexical entries at some point and adjust its feature composition to the target ([+definite, +anaphoric, ±bridging]). The relevant feature reassembly then varies depending on which lexical entry is selected. If learners go with the entry with the features of *the* ([+definite, ±anaphoric, ±bridging]), the reassembly task would be to restrict the anaphoricity feature ([±anaphoric] => [+anaphoric]) (the same as Scenarios 1 and 3A). On the other hand, when the entry with the features of *that* ([+definite, +anaphoric, −bridging]) is chosen, the learners' task would be to relax the bridging feature ([−bridging] => [±bridging]) (the same as Scenario 2).

In sum, whereas Korean-speaking learners are expected to map the features of *ku* onto *the* with no reassembly task to be done, there are several possibilities for English-speaking learners in terms of what features they initially map onto *sono*; and the relevant feature reassembly task differs accordingly. With each scenario empirically and theoretically motivated in different ways, it seems ultimately an empirical question which will be true. Next, let us consider the relative ease/difficulty of each acquisition task above.

5.2.2 Relative ease/difficulty and acquisition order

Table 5.2 summarises the possible acquisition scenarios for each L1 group discussed in the previous section.

Table 5.2 Summary of acquisition scenarios for each L1 group

L1-Korean		<u>Initial mapping</u> L1 *ku* = L2 *sono* [+definite, +anaphoric, ±bridging]	<u>Feature reassembly</u> not necessary
L1-English	Scenario 1	<u>Initial mapping</u> L1 *the* = L2 *sono* [+definite, ±anaphoric, ±bridging]	<u>Feature reassembly</u> [±anaphoric] => [−anaphoric]
	Scenario 2	<u>Initial mapping</u> L1 *that* = L2 *sono* [+definite, +anaphoric, −bridging]	<u>Feature reassembly</u> [−bridging] => [±bridging]
	Scenario 3A	<u>Initial mapping</u> L1 *the/that* = L2 *sono* [+definite, ±anaphoric, ±bridging]	<u>Feature reassembly</u> [±anaphoric] => [−anaphoric]

	Scenario 3B	<u>Initial mapping</u>
		(i) L1 *the* = L2 *sono*
		[+definite, ±anaphoric, ±bridging]
		
		(ii) L1 *that* = L2 *sono*
		[+definite, +anaphoric, −bridging]
			<u>Feature reassembly</u>
			If (i) is selected,
			[±anaphoric]
			=> [−anaphoric]
			
			If (ii) is selected,
			[−bridging]
			=> [±bridging]

The key point in predicting relative ease/difficulty is that no feature reassembly is required for L1-Korean learners whereas some feature reassembly is necessary for L1-English learners in any of the presented scenarios. On accounts that postulate L1 transfer of lexical features including the FRH, the relevant properties of *sono* are predicted to be more difficult for English-speaking learners than Korean-speaking learners. Thus, Korean-speaking learners will acquire the properties of *sono* earlier, or more easily, than English speakers. Crucially, this prediction remains essentially the same, irrespective of which scenario holds true with English-speaking learners. However, the relative ease/difficulty that L1 English learners will experience may depend on the acquisition task. More precisely, which features to select for the initial mapping potentially influences the achievability of the subsequent feature reassembly task. For example, it has been argued that to constrain a superset grammar to its subset may be impossible because some form of negative evidence is necessary (e.g., Inagaki, 2001; Trahey & White, 1993). This means that the task of constraining the feature [±anaphoric] to [+anaphoric] (Scenarios 1, 3A, and 3B(i)) may be more challenging than that of relaxing the feature [−bridging] to [±bridging] (Scenarios 2 and 3B(ii)). In relaxing a feature set, no serious learnability problem is expected, since such a task can be motivated by means of positive evidence (i.e., instances of *sono* in [+bridging] contexts) only. However, in constraining a feature distribution, acquisition may be delayed if negative evidence necessary for the feature reassembly (i.e., the information that use of *sono* is infelicitous in non-anaphoric definite contexts) is scarce. Indeed, this seems to be the case: the examination of a selection of widely used Japanese textbooks along with informal consultation with several Japanese language teachers confirms that the incompatibility of *sono* with non-anaphoric (unique) definiteness are not explicitly taught. Thus, it can be predicted that English-speaking learners may not be able to complete the feature reassembly task, although some (typically advanced) learners might be, as observed in some previous studies investigating L2 learners under similar learnability problems of retreating from an L1-based representation to a more constrained target (e.g., Gabriele, 2009; Marsden, Whong, & Gil, 2018; Slabakova, 2006).

In sum, Korean-speaking learners are predicted to acquire the relevant properties of *sono* generally earlier, or more easily, than English-speaking counterparts.

As to L1-English learners, however, the learnability of the properties may vary depending on which acquisition task they are expected to tackle. When the feature reassembly task involves constraining a feature set (Scenarios 1, 3A, and 3B(i)), it is predicted to be more challenging than the task of extending the distribution of features (Scenarios 2 and 3B(ii)). (What task is tackled by English-speaking learners will be discussed in light of experimental findings in Chapter 9.)

5.3 Covert definiteness marking with numeral quantifier constructions

5.3.1 Acquisition tasks

As a prerequisite to the acquisition to the definiteness constraint, learners must know the essential properties of Japanese NQs: namely, that Japanese numerals need to combine appropriate classifiers; and that NQs can occur in multiple positions including post-nominal and floating positions. Taking the Minimalist view that the main task of language acquisition is to assemble relevant features on lexical items in the target language, it can be understood that learners must acquire a set of features that allow post-nominal and floating NQ constructions as well as their semantic properties (specifics of these features are not discussed here, for simplicity, because predictions can be made based on the crosslinguistic differences).[42]

Now let us consider what learners from each L1 background need to acquire about the Japanese NQs. Assuming that Japanese and Korean NQs encode the same set of features (hence the corresponding properties), the FRH predicts that Korean-speaking learners of Japanese will map those features from Korean NQs onto the Japanese

42 Details of the features will be relevant to the discussion of the main experimental results in Chapter 9.

counterparts once they detect Japanese NQs. The parallelism between Japanese and Korean in the syntax-semantics of NQs means that no feature reassembly is necessary for Korean-speaking learners with regard to the definiteness constraint on floating NQs. In contrast, English-speaking learners must learn the relevant features without benefit of L1 knowledge. One possible prediction based on the FRH is that they will map the properties of English NQs onto Japanese NQs. In this case, the reassembly task should involve acquiring the following three mismatches between Japanese and English NQs.

(5.7) a. Japanese numerals must combine with appropriate classifiers in contrast to English numerals, which do not.

b. Japanese NQs can appear in post-nominal and floating positions in contrast to English NQs, which can appear only in pre-nominal position.

c. Japanese NQs must be indefinite in floating position in contrast to English counterparts which can be definite or indefinite.

Whereas the properties of Japanese NQs in (5.7a) and (5.7b) are, in principle, acquirable based on positive evidence, the acquisition of (5.7c) may cause a learnability problem for English-speaking learners. That is, acquisition of the lack of definite interpretation for floating NQs is not motivated by information available to learners either through input or the L1. It is possible that L2 learners could make use of L2 instruction as negative evidence. However, the definiteness constraint on floating NQs does not seem to be covered in L2 Japanese language instruction (according to the consultation with the Japanese language instructors and the textbook inspection mentioned above). Therefore, the acquisition of the constraint seems to constitute a poverty-of-the-stimulus (POS) problem for English-speaking speakers. This learnability problem is discussed further in the next section.

5.3.2 Relative ease/difficulty and acquisition order

The relative ease/difficulty of the L2 acquisition of the covert definiteness property of Japanese NQs varies depending on whether the FRH or the cline of difficulty approach is adopted, as shown below. In what follows, predictions will be provided first based on the FRH followed by the cline of difficulty.

To repeat, the key factor for the FRH is the necessity of feature reassembly. Generally, the relevant property will be harder to acquire when feature reassembly is required than when it is not. L1-Korean learners are then predicted to acquire the properties of the Japanese NQs faster, or more easily, than L1-English learners as a result of L1 positive transfer. On the other hand, whether English-speaking learners can acquire the definiteness constraint on Japanese floating NQs seems an empirical question, particularly given the POS problem for L1-English learners discussed above. However, given the previous findings suggesting that L2 learners can overcome such a learnability problem by virtue of the domain-specific syntax-semantics computation mechanisms (i.e., UG) (cited in Chapter 2), the definiteness constraint may be acquirable despite the POS. The most directly relevant empirical support for this possibility comes from Okuma (2019). Okuma provides evidence that L1-English learners can acquire what seems to be another POS property of Japanese NQ constructions, namely the semantic constraint that floating NQs cannot have a collective reading as opposed to post-nominal NQs, which can (see Chapter 2).

Now looking at the situation from the standpoint of the cline of difficulty, there is, along with feature reassembly necessity, another important factor, namely overt vs. covert feature realisation (i.e., whether the relevant feature is expressed overtly with a particular morpheme, or covertly through context or word order). In this view, the definiteness constraint on floating NQs vs. the absence thereof with post-nominal NQs can be regarded as a covert property because it is expressed through word order (as discussed in Chapter 2). The predictions of the cline of difficulty for six types of acquisition task are repeated in Figure 5.1.

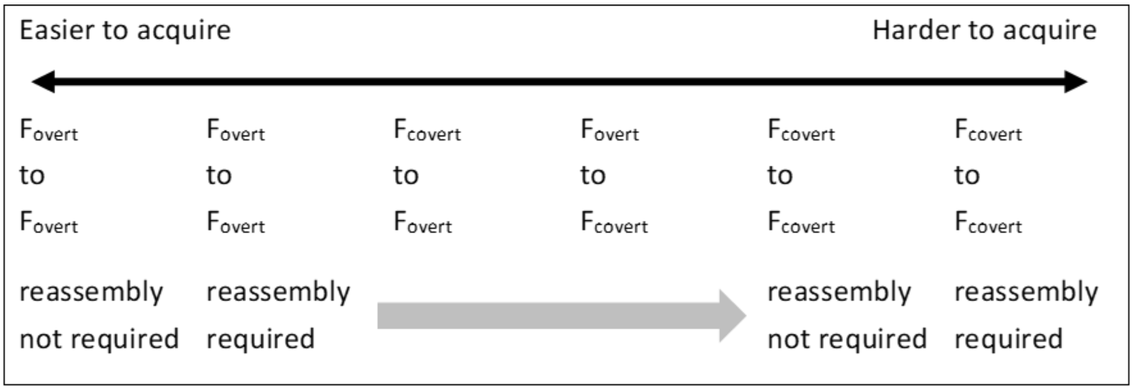


Figure 5.1 Cline of difficulty in feature acquisition (adapted from Cho & Slabakova 2014, p. 166)

As discussed in Chapter 2, the cline of difficulty suggests the following.

(5.8) a. Covert feature realisation is predicted to be more challenging to acquire than overt realisation.

b. In acquiring the covert realisation of a feature, an L1 overt realisation of that feature benefits learners more than an L1 corresponding covert realisation.

(5.8a) means that regardless of the L1, the covert definiteness property of the Japanese NQs is predicted to be more difficult to acquire than the overt property of *sono* (I will come back to this later in Section 5.5). More importantly here, (5.8b) leads to the prediction that the English-speaking learners may acquire the covert property of the NQs with less difficulty than Korean-speaking learners as a result of facilitation from the L1 functional morphology for definiteness (the English articles) (i.e., the task for the L1-

English group corresponds to the fourth point on the cline, whereas for the L1-Korean group to the fifth point). Note that this does not necessarily mean that the corresponding L1 covert property (i.e., the definiteness constraint on Korean NQs) will not have a facilitative effect. But it rather suggests that the degree of facilitation may be weaker than that of the L1 overt morphology (i.e., the English articles).

Before presenting predictions to address the research questions, let us consider some potential confounding factors that might affect the learnability of the target properties.

5.4 Potential confounding factors

Firstly, given the Interface Hypothesis (Sorace, 2011; Sorace & Filiaci, 2006), L2 learners may have persistent difficulty acquiring definiteness properties, which require integration of relevant syntax-semantics knowledge (i.e., internal components of the grammar) with discoursal/pragmatic information (i.e., external components). This may be the case regardless of whether feature reassembly is necessary or whether features are overt or covert.

Another factor, which is specifically relevant to the acquisition of *sono*, is the potentially challenging nature of bridging (i.e., contexts where anaphoric relations are expressed indirectly). If bridging is generally difficult for L2 learners for pragmatic reasons such as presupposition accommodation, as discussed in Chapter 4, in relation to Cho (2017), then *sono* is expected to be more difficult to acquire in indirectly anaphoric (i.e., bridging) than directly anaphoric contexts. This means that even if learners have correctly assembled the target feature set of [+definite, +anaphoric, ±bridging] onto *sono*, they may behave in a less target-like manner in bridging anaphoric than directly anaphoric contexts.

These factors are not taken into account in outlining overall predictions in the next section for simplicity but will be considered later in the discussion of the main study results (Chapter 9).

5.5 Conclusion: overall predictions

Finally, let me conclude the chapter by presenting overall predictions for the research questions, which were introduced in Chapter 1 and repeated as (5.9–5.11), based on the discussions above.

(5.9) Research question 1:

Is a covert feature expression more difficult to acquire than an overt feature expression?

(5.10) Research question 2:

Does the necessity of feature reassembly make the acquisition task more difficult?

(5.11) Research question 3:

In which situation is the acquisition of a covert feature expression less difficult, (i) when the L1 has a functional morpheme that realises the feature overtly or (ii) when the L1 has a corresponding covert expression?

Starting with Research question 1, the cline of difficulty maintains that the answer is yes (the FRH does not offer a testable prediction). That is, the covert definiteness property of Japanese NQs is predicted to be more difficult than the overt property of *sono*.

Turning to Research question 2, the FRH and the cline of difficulty both predict that the necessity of feature reassembly makes the task more difficult. This means that the property of *sono* is predicted to be more difficult to acquire for English-speaking learners than Korean-speaking counterparts.

As to Research question 3, the FRH and the cline of difficulty make distinct predictions. In terms of the FRH, what is crucial is whether the L1 has a corresponding property or not. This means that learners are expected to acquire a covert property with less difficulty when a corresponding covert L1 property exists than when not. Therefore, Korean learners are predicted to acquire the definiteness property of Japanese NQs faster, or more easily, than English learners: having a functional morpheme that overtly realises the relevant property will not be an advantage for L1-English learners over L1-Korean counterparts. In contrast, the cline of difficulty predicts that L1-English learners will acquire the Japanese covert definiteness property faster, or more easily, than L1-Korean learners. From the perspective of the cline of difficulty, the challenge pertaining to L2 covert property acquisition would be better alleviated with an L1 overt realisation of the relevant property through functional morphology (i.e., the English articles) than with an L1 corresponding covert property (i.e., the definite constraint on Korean NQs).

The predictions for each research question are summarised in Table 5.3. (5.12) and (5.13) present the predictions in ways more specific to the present thesis ("$x > y$" means x is acquired faster, or more easily, than y). These predictions are evaluated in the experimental results to be reported in later chapters.

Table 5.3 Summary of predictions for each research question (RQ)

Approach	RQ1	RQ2	RQ3
FRH	N/A	Yes	(ii)
Cline of difficulty	Yes	Yes	(i)

(5.12) FRH approach

- L1-Korean: *sono* > L1-English: *sono*
- L1-Korean: NQs > L1-English: NQs

(5.13) Cline of difficulty approach

L1-Korean: *sono* > L1-English: *sono* > L1-English: NQs > L1-Korean: NQs

Chapter 6: Experiment design and preliminary studies with native Japanese speakers

6.1 Introduction

This chapter reports on a series of experimental studies with native Japanese speakers conducted to check that those properties of the target phenomena that are agreed in the literature are measurable, and to investigate which of the competing accounts best describes the disputed properties (discussed in Chapter 3 and repeated below).

(6.1) Properties of the demonstrative *sono*

Agreed:

- The use of *sono* is preferred in direct anaphoric contexts over bare forms.
- The use of *sono* is infelicitous in non-anaphoric (i.e., unique) definite contexts.
- The use of *sono* is felicitous in bridging (anaphoric) definite contexts.

Disputed:

- Whether the use of *sono* is preferred over bare forms in bridging contexts.

(6.2) Properties of NQs

Agreed

- Floating NQs are acceptable only in [−definite] contexts.
- Post-nominal NQs are acceptable in [+definite] contexts.

Disputed:

- Whether floating NQs are acceptable in [+specific, −definite] contexts.

- Whether post-nominal NQs are acceptable in [−specific, −definite] contexts.

Native Japanese data were collected by means of two acceptability judgement tasks (AJTs) and two self-paced reading tasks (SPRTs) in total. Another important objective of these studies was to carefully develop testing materials for the main studies that involve L2 learners (reported in subsequent chapters), by fine-tuning the experimental design and test instruments in light of experimental data. Chronologically, the pilot studies were conducted in the order shown in (6.3) but the AJTs will be reported before the SPRTs below.

(6.3) AJT ver. 1 => SPRT ver. 1 => SPRT ver. 2 = AJT ver. 2

(“ *x* = *y*” means *x* and *y* were conducted concurrently.)

It will be shown that all agreed properties are experimentally observable offline, by means of untimed AJTs, but not online, by means of SPRTs. As to the disputed properties, the following conclusions will be drawn based on AJT results.

(6.4) Bare NPs are as acceptable as *sono* NPs in bridging anaphoric contexts.

(6.5) Floating NQs are acceptable in [+specific, −definite] contexts whereas post-nominal NQs are acceptable in [−specific, −definite] contexts.

In what follows, after rationales for the choice of method are provided, the series of experimental studies is presented. Pilot AJT ver. 1 is reported first, focusing mainly on the results concerning the disputed properties of NQs. Then, Pilot AJT ver.

2 will be presented with a view to providing further evidence of the definiteness constraint of NQs and the agreed and disputed properties of *sono*. As to the SPRTs, after stating some key assumptions in the self-paced reading paradigm and describing the postulated processing mechanisms for the target linguistic properties, I will outline the experiment designs of the two studies together (Pilot SPRTs ver. 1 and ver. 2), followed by their key findings. Finally, implications of the overall findings will be discussed in relation to the acquisition tasks and predictions for the research questions offered in the previous chapter.

All pilot experiments reported in the present thesis were approved by the Language & Linguistic Science ethics committee at the University of York and the procedure always included gaining informed consent from participants before participation. Furthermore, all statistical analyses of experimental data reported in the thesis were conducted in R environment (R Core Team, 2018).

6.2 Rationales for combining online and offline methods

The present thesis employs two different methods, namely an untimed AJT and an SPRT. Explicit judgement tasks such as AJTs have been widely employed in L2 acquisition research for decades, for its advantage of explicitly testing what is unacceptable as well as what is acceptable: this would not be achievable with other measures such as production or comprehension (e.g., Cowart, 1997; Schütze, 1996; Spinner & Gass, 2019). In recent years, it has become more commonplace that explicit offline tasks are combined with online (i.e., real-time processing) methods such as the self-paced reading and eye-tracking paradigms. Generally, online tasks are assumed to tap into more implicit and automated linguistic knowledge than offline methods such as grammaticality/acceptability judgement tasks, which arguably require more explicit metalinguistic reasoning on the part of participants (e.g., Ellis, 2005, 2006;

Jiang, 2012; Orfitelli & Polinsky, 2017). Among the most obvious and primary motives for combining offline and online tasks are triangulation, and maximisation of the chance of obtaining evidence for relevant underlying linguistic knowledge. Indeed, a number of studies using both kinds of method have reported that target phenomena were observed with only one type of method but not in the other. For example, it has been documented that L2 learners tend to be less target-like in explicit tasks such as grammaticality/acceptability judgement than in implicit tasks such as self-paced reading. Some attribute this to extra cognitive demand of metalinguistic reasoning imposed on L2 learners in performing explicit judgment tasks (e.g., Hopp, 2009; Ionin, Choi, & Liu, 2019; Orfitelli & Polinsky, 2017). On the other hand, others showed more target-like performance in offline tasks compared to online tasks (e.g., Hopp, 2010; Roberts, Gullberg, & Indefrey, 2008; Roberts & Liszka, 2013). One account compatible with such observations is the Morphological Congruence Hypothesis (e.g., Jiang, Novokshanova, Masuda, & Wang, 2011), which maintains that online sensitivity to L2 specific grammatical morphemes cannot be acquired. However, given the conflicting findings in the literature, task effects may interact with the types of linguistic property under investigation: implicit knowledge may be easier to measure than explicit knowledge with some structures but the opposite may be true with others (e.g., Ellis, 2006; Zufferey, Mak, Degand, & Sander, 2015).

Regarding definiteness linguistic properties, whereas explicit judgement tasks have been widely used in the research into L2 English articles (as shown in Chapter 4), the implementation of online methods has been relatively limited. Kim and Lakshmanan (2009) tested the effect of specificity (e.g., Ionin et al., 2004) on L2 English article choice by L1-Korean learners, using an SPRT as well an AJT. They reported converging evidence for the relevant semantic effect from the offline and online tasks. Trenkic, Mirkovic, and Altmann (2014) tested whether Mandarin-speaking L2-English learners could effectively use article information for reference resolution. The results of their eye-tracking experiment suggest that the L2 learners can utilise articles for real time processing. More recently, some studies (Cho, 2020; Ionin, Choi, & Liu, 2019, 2020) used both untimed judgement tasks and self-paced reading tasks to investigate whether L2 learners from article-less L1 backgrounds (i.e., Korean and Mandarin) are sensitive to article errors such as omissions in obligatory contexts and inappropriate uses (e.g., infelicitous use of *the* in indefinite contexts). These studies unanimously show that L2 learners are more sensitive to the relevant properties in online tasks (i.e., self-paced reading) than in offline tasks (i.e., grammaticality/acceptability judgement). These results suggest that L2 knowledge of definiteness properties may be easier to observe by means of online rather than offline tasks. However, to my knowledge, there has been no real-time processing research into definiteness properties in article-less L2s or languages such as Japanese where definiteness is not distinguished by functional morphology. That is, it remains empirically open whether the observed online vs. offline contrast holds true with the target Japanese definiteness properties of this thesis, or even whether these properties are experimentally measurable (either online or offline). Therefore, testing whether they can be observed with native Japanese speakers is a crucial first step for the present thesis.

6.3 Pilot AJT ver. 1

This first pilot AJT was conducted to investigate whether the properties of the demonstrative *sono* and NQs discussed in the theoretical literature (the agreed and disputed properties of these linguistic phenomena shown in (6.1−6.2)) are experimentally observable. However, since the most important objective was to gain insights into what are the best descriptions of the disputed properties of NQs, I will report the results, focusing on NQ items. (For conciseness, the results for the *sono* items will not be presented, but they are essentially similar to those in Pilot AJT ver. 2, reported later in Section 6.4).

6.3.1 Method

6.3.1.1 Participants

Twenty Japanese native speakers participated in the study. All participants were adult Japanese speakers (18 years old or older). Four participants were postgraduate students studying at a university in the UK while the rest were Japanese speakers living in Japan. No demographic information such as age and gender was collected because there was no particular theoretical interest in doing so.

6.3.1.2 Test instrument

The AJT was administered, using Qualtrics (version: Oct. 2019). The validity of web-based AJTs has been attested by recent studies in terms of their strong correlation with offline informal judgement ratings (e.g., Juezk, 2016; Sprouse, Schütze, & Almeida, 2013) and laboratory-based judgement ratings (Sprouse, 2011). The task was to read a series of short passages that established context, each of which was followed by a target sentence to be rated in terms of its naturalness as a continuation of the preceding text, on a 7-point scale (0 = "completely odd", 6 = "completely natural").

The AJT tested the native Japanese speakers' knowledge of the relevant properties of the demonstrative *sono* and of NQ constructions (6.1–6.2) in six types of context (three contexts concerning *sono*, anaphoric, non-anaphoric, and bridging will not be reported, as mentioned above). Specifically, there are three relevant contexts for NQs, namely [+definite, +specific], [−definite, +specific], and [−definite, −specific] contexts (for the definitions of definiteness and specificity, see Chapter 3), as exemplified in (6.6–6.8) (the context is given in English in these examples for convenience but was presented in Japanese in the actual experiment). The choice of construction (either post-nominal, where the NQ precedes the case marker, or floating, where the NQ follows the case marker) was predicted to affect acceptability of the sentence in different ways. In [+definite, +specific] contexts (6.6), whereas the post-nominal construction will be acceptable according to all of the accounts reviewed in Chapter 3, the floating construction is predicted to be unacceptable. However, predictions vary in the two indefinite contexts. In [−definite, +specific] contexts (6.7), some accounts (Prediction C in Chapter 3: Kobayashi & Yoshimoto 2001; Nakanishi, 2007; Shin, 2017, among others) predict both NQ constructions to be felicitous whereas according to others (Prediction A: Furuya, 2012; Watanabe, 2006; and Prediction B: Downing, 1996; Huang & Ochi, 2014), only the post-nominal construction will be acceptable. In [−definite, −specific] contexts (6.8), some argue that both constructions are allowed (Predictions A and C) whereas others expect post-nominal NQs to be degraded in acceptability (Prediction B).

(6.6) NQs: [+definite, +specific] contexts (post-nominal vs. floating)

PREAMBLE: *Taroo has two little sisters.*

Taroo-wa sensyuu sono {imooto huta-ri-o vs. imooto-o huta-ri}
Taroo-TOP last.week SONO sister 2-CL-ACC vs. sister-ACC 2-CL

yuuenti-ni tureteikimasita
amusement-to took

‘Taroo took the two sisters to an amusement park last week.’

(6.7) NQs: [−definite, +specific] contexts (post-nominal vs. floating)

PREAMBLE: *Hanako got two PCs and one printer from a friend of hers.*

Hanako-wa toriaezu {pasokon iti-dai-o vs. pasokon-o iti-dai}

Hanako-TOP for.now PC 1-CL-ACC vs. PC-ACC 1-CL

sigoto-yoo-ni tukau-koto-ni simasita.

work-use-for use-ing-DAT decided

'Hanako decided to use one (of the) PC(s) for work for the time being.'

(6.8) NQs: [−definite, −specific] contexts (post-nominal vs. floating)

PREAMBLE: *Taroo does online shopping almost every day.*

Kinoo-wa {hon san-satu-o vs. hon-o san-satu} kaimasita.

yesterday-TOP book 3-CL-ACC vs. book-ACC 3-CL bought

'He bought three books yesterday.'

Additionally, the following three manipulations were made for the NQ items. Firstly, the classifiers used in the test items were chosen among those introduced in basic-level Japanese language textbooks to make the AJT learner-friendly.[43] Secondly, only predicates denoting events, in other words, S-level predicates were used in consideration of the incompatibility of floating NQs with I-level predicates (as discussed in Chapter 3). Finally, the target nouns associated with NQs were placed in object position. This was to avoid adding another variable of position of the target NP,

---

[43] The textbooks consulted includes the *Genki* series (Banno, Ohno, Sakane, Shinagawa, & Takashiki, 1999) and the *Minna-no-Nihongo* series (3A Network, 1998). The chosen classifiers (CLs) are *nin* (CL for human), *satu* (CL for books), *hon* (CL for long and round objects), *hiki* (CL for small animals), *dai* (CL for vehicles and big equipment), *ma*i (CL for thin, flat objects) and *ko* (CL for small objects).

particularly, information bias of the topic marker *wa*, which often marks external arguments in matrix clauses in Japanese.[44]

There were 12 test items for each of the six context types. The set of test items in each context type had two different versions of target sentence which minimally differed from each other regarding the form of NP. In the NQ items, the target NP contained either a post-nominal NQ or a floating NQ (as can been in the examples above). The test items were divided into four lists with each containing 36 critical items (6 items × 6 contexts). Each item was presented in only one type of NQ (e.g., post-nominal or floating) for each participant. Moreover, 18 unacceptable fillers (3 items × 6 types) were added to each item list to balance out the numbers of acceptable and unacceptable items (the detail of the fillers is not provided because it is not essential for the purpose of this section).[45] In order to prevent order effects, the stimuli were pseudo-randomised for each participant.

6.3.1.3 Procedure

The participants accessed the AJT through a web link provided by the researcher. The task was designed to be completed individually by participants either on personal computers or on mobile devices. After agreeing to participate in the study, they were asked to read a brief explanation of the task and familiarize themselves with the task format through practice items. During the experiment, it was impossible to go back and change answers to the previous items. There was no time limit for the task. The participants completed the task within 25 minutes on average and received online

[44] Topic-marked NPs tend to be given a [+definite, +specific] interpretation regardless of the choice of NQ construction, because of the correlation between topic/old information and definiteness (as discussed in Chapter 3). In contrast, object NPs in the canonical position are not subject to such bias.
[45] Note that *sono* items served as fillers for the NQ items and vice versa, too. Therefore, the filler to critical item ratio is practically 36:18 per target property.

shopping vouchers for their participation.

6.3.2 Results

The mean ratings of the two types of NQ construction are provided in Table 6.1 and Figure 6.1. In the [−definite] context, both construction types received mean ratings of at least 4.7 on the 6-point scale with no meaningful differences whereas in the [+definite] context, the post-nominal construction was preferred to the floating counterpart with a more than 2-point difference (Table 6.1).

Table 6.1 Acceptability ratings for NQ construction items: Pilot AJT ver. 1 (SD in parentheses)

	[+def, +spec]	[−def, −spec]	[−def, −spec]
Post-nominal	5.05 (1.47)	4.73 (1.51)	4.97 (1.22)
Floating	2.78 (2.03)	4.70 (1.39)	5.20 (1.20)

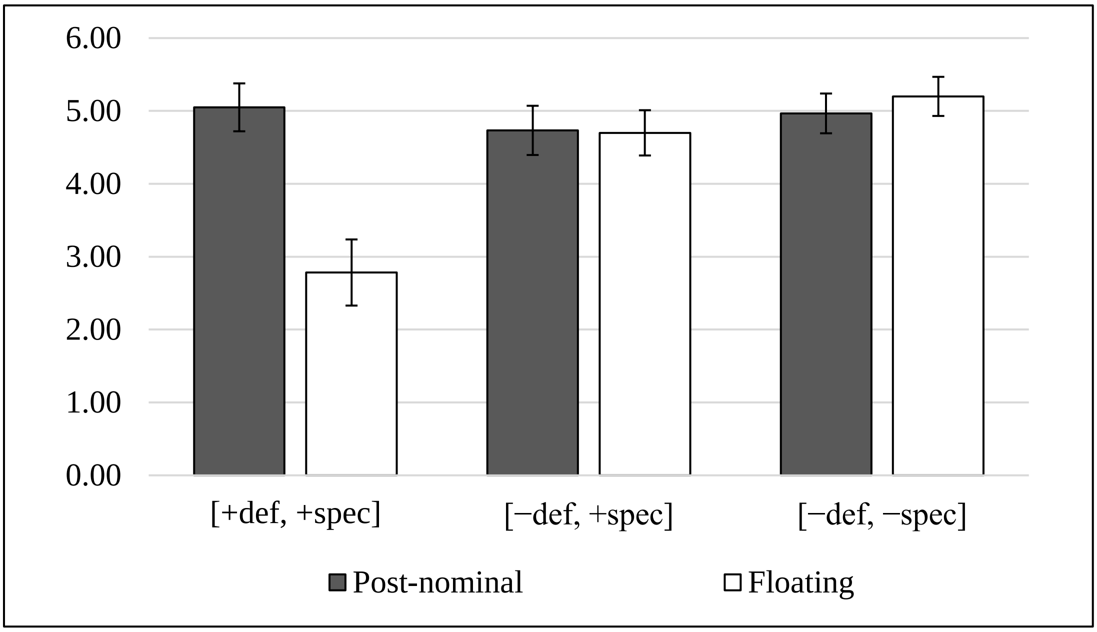


Figure 6.1 Mean ratings for NQ construction items (error bars = SE)

For further analysis, the raw ratings were converted into z-scores and tested with linear mixed effects modelling using the *lme4* package (Bates, Mächler, Bolker,

& Walker, 2015) to evaluate how NQ type affected acceptability ratings in each context.[46] Models were fit with NQ type post-nominal vs. floating, coded as −0.5 vs. 0.5) as a fixed factor and with random intercepts for participants and items as well as by-participants and by-items slopes for NQ type (i.e., maximal random structure (Barr, Levy, Scheepers, & Tily, 2013). Separate models were constructed for each context because the target NP and sentence structures were controlled for only within each context but not across contexts. In order to examine whether NQ type had a significant impact on the ratings in each context, the models were compared to the respective null models (i.e., the maximal critical models minus the fixed effect under investigation, in this case, the effect of NQ type), following Winter (2020). The null model was compared to the critical model, using the *anova* function (i.e., the likelihood-ratio testing). Table 6.2 presents the outputs for the relevant critical models and the model comparisons. The model comparisons revealed that NQ type had a significant effect in the [+def(inite), +spec(ific)] context. However, non-significant results were obtained in the [−def, +spec] and [−def, −spec] contexts. These results suggest that both NQ constructions are felicitous in indefinite contexts regardless of specificity but only the post-nominal construction is compatible with definite contexts, which confirmed the agreed properties. In terms of the disputed properties, the results favour Prediction C (e.g., Kobayashi & Yoshimoto, 2001; Nakanishi, 2007; Shin, 2017).

[46] Z-transformation is known to mitigate potential individual variations, scale bias, and thus wield a higher statistical power compared to raw scores by expressing each participant's responses via a standardized unit (Schütze, 2013).

Table 6.2 Fixed effect estimates of critical linear mixed-effects models and likelihood-ratio testing results of critical vs. null model comparisons for NQ items: Pilot AJT ver. 1

Fixed Effects	β	*SE*	*t*	χ^2	*p*
[+def, +spec]					
(intercept)	−0.313	0.094	−3.313		
quantifier type	0.920	0.138	6.652	23.684	< .001***
[−def, +spec]					
(intercept)	−0.618	0.065	−9.502		
quantifier type	0.026	0.110	0.233	0.054	.816
[−def, −spec]					
(intercept)	−0.782	0.047	−16.720		
quantifier type	−0.097	0.095	−1.020	0.986	.321

Note. Formula: z-rating ~ quantifier type + (quantifier type | participant) + (quantifier type | item). Coding: quantifier type: post-nominal = −0.5 vs. floating = 0.5. *** p < .001. P values were calculated for χ^2 values obtained through comparisons of the full vs. null model in terms of the fix effect of *quantifier type* by means of the likelihood-ratio testing.

6.3.3 Discussion

The results confirmed the agreed properties the NQ constructions repeated below.

(6.9) Agreed properties of NQ constructions

- Floating NQs are allowed only in [−definite] contexts.
- Post-nominal NQs are compatible with [+definite] contexts.

As to their disputed properties (repeated below in (6.10)), the results revealed the following.

(6.10) Disputed properties of NQ constructions

- Whether floating NQs are acceptable in [+specific, −definite] contexts.
- Whether post-nominal NQs are acceptable in [−specific, −definite] contexts.

The evidence suggested that the floating construction was constrained to be [−definite] rather than [−specific, −definite]; whereas the post-nominal construction was highly acceptable in [−specific, −definite] contexts as well as in the other tested contexts. This acceptability paradigm suggests that the constraint on the floating NQs concerns definiteness rather than specificity, which matches Prediction C (the floating-NQs-as-adverbials account).

However, some potential complications were identified that could affect the interpretation of the results as follows. The first concerns the use of *sono* in the [+definite, +specific] condition. *Sono* was included to explicitly mark the target NP in this condition as definite. However, it was pointed out by some native Japanese-speaking linguists (T. Hokari, personal communication, April 27, 2018; T. Nakamura, personal communication, Dec 7, 2017) that the relative unacceptability of the floating construction (e.g., *sono hon-o san-satu*) in the definite context in contrast to post-nominal NQs (e.g., *sono hon san-satu-o 'the three books'*) in this AJT may be attributed to the former's incompatibility with *sono.* Specifically, they suggested that the acceptability of floating NQs would improve if *sono* is replaced with its explicitly plural version, *sorerano*. This is puzzling given the fact that *sono* is widely used not only with definite singulars but also plurals. Indeed, there seems no clear difference in acceptability between *sono* and *sorerano* in the post-nominal construction (6.11a); however, the relevant contrast seems to arise in the floating construction (6.11b) (according to the intuition of the author and a few native-Japanese informants).

(6.11) Comparisons of *sono* and *sorera*

PREAMBLE: *Hanako bought three books yesterday.*

a. Post-nominal

{*Sono* vs. *sorerano*} hon *san-satu-o* moo yomi-oeta.

SONO vs. SONO.PL book 3-CL-ACC already read-finished

b. Floating

{#*Sono vs.* ?*sorerano*} *hon-o* *san-satu* moo yomi-oeta.

SONO vs. SONO.PL book-ACC 3-CL already read-finished

(Judged under the interpretation, 'she has already finished reading the three books.')

The floating NQ number incompatibility with *sono* in (6.11b) might be due to a pragmatic factor. That is, when it modifies a countable bare NP, a singular reading seems predominant (Nemoto, 2005) presumably because Japanese has a way to solve number ambiguity by using the explicitly plural *sorerano.* Therefore, one would expect that a *sono* NP should be singular because if the NP is meant to be plural, the speaker is likely to use the unambiguous *sorerano* along the lines of "the maxim of manner" (Grice, 1975). This may partially account for the lower rating of floating NQs. If this asymmetry between *sono* and *sorerano* is true, the genuine contrast between post-nominal and floating constructions observed in this experiment might be more subtle. To avoid this potential complication, it was decided to remove *sono* in the definite context for subsequent experiments.

Additionally, some minor revisions were made to the test battery for subsequent studies as follows. Firstly, sentence structures used for the target sentences were more strictly controlled (e.g., since some sentences started with an explicit subject but others included a null counterpart in Pilot AJT ver. 1, revisions were made for all sentences to include an explicit subject). Secondly, the numbers of animate, inanimate, human, and non-human nouns were balanced out. Finally, since it is already clear at this point that neither type of NQ construction is constrained in terms of specificity, it was decided that the [−definite, +specific] context was not to be included in subsequent experiments for simplicity. Finally, the number of test items for each context was reduced from twelve to eight with an intention of reducing task load. Finally, the number of classifiers used was reduced from 7 to 6.[47]

6.4 Pilot AJT ver. 2

The main objectives of Pilot AJT ver. 2 were to confirm the validity of the test sentences used in Pilot SPRT ver. 2 (to be presented later) by means of an offline behavioural measure; and to ascertain that the target linguistic properties of *sono* and the NQ constructions observed in Pilot AJT ver. 1 could be replicated. I focus on the second objective for the purpose of this section.

6.4.1 Method

6.4.1.1 Participants

Twenty native Japanese speakers participated in this experiment (they had not participated in any previous experiments or SPRT ver. 2). Most participants were resident in Japan while four of them had lived in the UK for more than two years studying for undergraduate or postgraduate degrees at the time of testing.

[47] Specifically, the classifier, *satu* (CL for books) was removed (because the number of nouns this classifier can select is more limited compared to the others), and the classifier *ko* (CL for small objects) was replaced with a more general classifier, *tu*.

6.4.1.2 Test instrument

The experimental design of Pilot AJT ver. 2 is essentially similar to that of Pilot AJT ver. 1. However, it differed in the following two ways (along with the revisions described in Section 6.3.3).

First, the target sentences were lengthened with extra adverbials or adjuncts. This modification was made primarily for addressing a potential methodological problem regarding SPRT Pilot ver. 1, so it will be explained in that context, in Section 6.5. (6.12) and (6.13) exemplify the revised NQ items (extra adverbials are in squares). Note that an extra adverbial is placed between the case marker and the NQ for the floating construction and after the case marker for the post-nominal construction. However, the prediction for native judgement remained the same: floating NQs would be rated significantly lower than post-nominal NQs in the [+definite] context (6.12), but the both types of NQ would be rated equally high in the [−definite] context (6.13).

(6.12) NQs: [+definite] contexts (post-nominal vs. floating)

PREAMBLE: *Taroo is a good tennis player. He played games with his friends, Takasi, Hirosi, and Goroo yesterday.*

Taroo-wa itumo-no-yoni
Taroo-TOP like.always

{tomodati san-nin-o $\boxed{\text{kantanni}}$ vs. tomodati-o $\boxed{\text{kantanni}}$ san-nin} sugu
friend 3-CL-ACC easily vs. friend-ACC easily 3-CL quickly

makasite-simai-masita
beat-finish-POL.PAST

'Taroo, as always, beat the three friends quickly and easily.'

(6.13) NQs: [−definite] contexts (post-nominal vs. floating)

PREAMBLE: *Taroo works as a vet at an animal hospital. Although he mostly treats dogs, ...*

Taroo-wa senzitu

Taroo-TOP the.other.day

{ neko san-biki-o mezurasiku vs. neko-o mezurasiku san-biki } tuzukete

{ cat 3-CL-ACC unusually vs. cat-ACC unusually 3-CL } in.a.row

tiryoo-si-masita.

treatment-do-POL.PST

'Unusually for him, Taroo treated three cats in a row the other day.'

A similar modification was made to *sono* items for coherence, as shown in (6.14)–(6.16) for each context type (anaphoric, non-anaphoric, and bridging). Native Japanese speakers were predicted to rate *sono* NPs significantly lower than bare NPs in the non-anaphoric context (6.15), but significantly higher than or as high as bare NPs in the anaphoric and bridging contexts (6.14 and 6.16), given the characteristics of *sono* proposed in the literature (6.1).

(6.14) *Sono*: anaphoric contexts (*sono* NP vs. bare NP)

PREAMBLE: *There is a very popular restaurant which serves great food for reasonable prices near Taroo's house. ...*

Taroo-wa mukasi-kara

Taroo-TOP past-from

{ *sono* resutoran-o vs. resutoran-o } dare-yori-mo

{ SONO restaurant-ACC vs. restaurant-ACC } more.than.anyone

often use-do-ASP.POL.PST

yoku riyoo-site-imasu

'Taroo has eaten at the restaurant more often than anyone since long ago.'

(6.15) *Sono*: non-anaphoric contexts (*sono* NP vs. bare NP)

PREAMBLE: *Taroo likes taking a night walk. When the weather is nice and the sky is clear of clouds, …*

Taroo-wa mukasi-kara { *sono* tuki-o vs. tuki-o } yukkuri

Taroo-TOP past-from { SONO moon-ACC vs. moon-ACC } slowly

tanosinde mite-imasu

gladly watch-ASP.POL.NPST

'Taroo relaxes and enjoys watching the moon as an old habit.'

(6.16) *Sono*: bridging contexts (*sono* NP vs. bare NP)

PREAMBLE: *Hanako wants to watch a flim that was just released.*

Hanako-wa daigakusei-no toki-kara

Hanako-TOP since.university

{ *sono* kantoku-o vs. kantoku-o } kokoro kara totemo

{ SONO director-ACC vs. director-ACC } from.heart very much

sonkei-site-imasu.

respect-do-ASP.POL.PST

'Hanako has respected the director from the bottom of her heart since university.'

The other major revision was the addition of several new types of filler. Particularly, there were four types prepared in two minimally different versions. These items were designed so that the target sentence be rated significantly higher in one condition (acceptable) than the other (unacceptable) (the details of these fillers are not provided because they are irrelevant here. However, they will be when Pilot SPRT ver. 2 is reported in the next section).

The test items were divided into 2 lists with each containing 96 items: 40 target items (8 items × 5 contexts), 32 control fillers (8 items × 4 contexts), and 24 unacceptable fillers (similar to those used in Pilot AJT ver. 1). The numbers of acceptable and unacceptable items were balanced, and each set of test items was presented in pseudo-randomised orders for each participant, as in Pilot AJT ver. 1.

6.4.1.3 Procedure

The participants went through the same procedure as in Pilot AJT ver. 1. They completed the task in approximately 45 minutes on average. Their participations were rewarded with shopping vouchers.

6.4.2 Results

6.4.2.1 *Sono*

Table 6.3 and Figure 6.2 summarise the mean ratings of the two NP forms (*sono* NP vs. bare NP) in the three contexts related to the properties of *sono*. All mean ratings are higher than the mid-point in the anaphoric and bridging contexts, with *sono* NPs being descriptively higher than bare NPs in the anaphoric context, but there being almost no difference in the bridging context. In the non-anaphoric contexts, s*ono* NPs were rated clearly lower than bare NPs.

Table 6.3 Mean acceptability ratings for sono *items: Pilot AJT ver. 2 (SD in parentheses)*

	Anaphoric	Non-anaphoric	Bridging
Sono NP	4.55 (1.61)	2.14 (2.00)	4.09 (1.91)
Bare NP	3.79 (1.94)	3.60 (2.05)	4.10 (2.01)

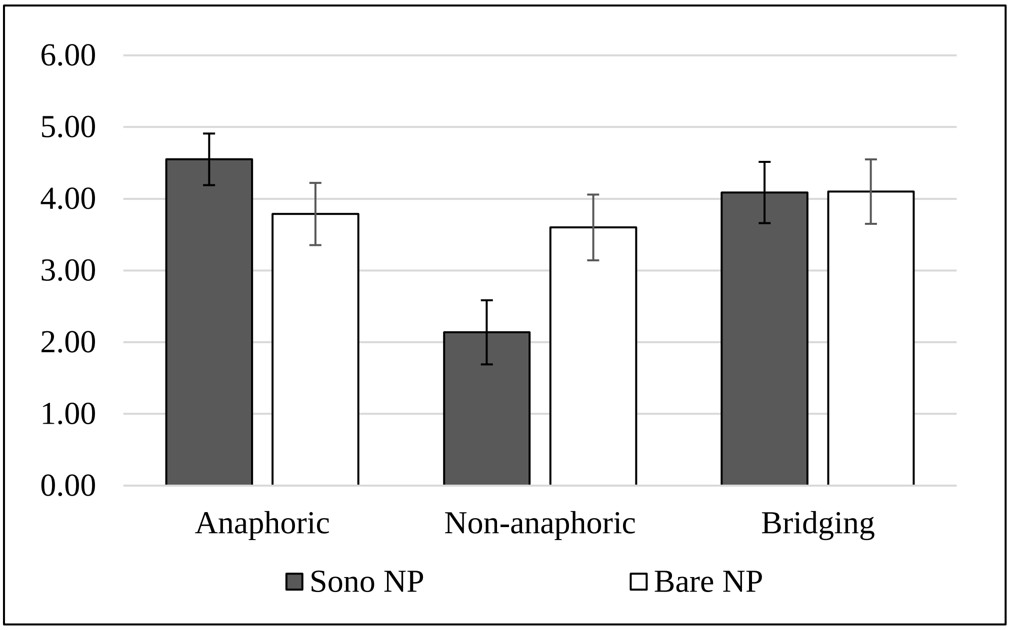


Figure 6.2 Mean ratings for sono *items (error bars = SE)*

Statistical significance testing through the comparison of the crucial mixed-effects models with null models confirmed the above observations: it was revealed that the fixed effect of NP type (*sono* NP vs. bare NP, coded as 0.5 vs. −0.5) was significant in the anaphoric and non-anaphoric conditions but not in the bridging context (the outputs of each crucial model and model comparison are provided in Tables 6.4).

Table 6.4 Fixed effect estimates of critical linear mixed-effects models and likelihood-ratio testing results of critical vs. null model comparisons for sono *items: Pilot AJT ver. 2*

Fixed effects	β	*SE*	*t*	χ^2	*p*
Anaphoric					
(intercept)	−0.554	0.136	−4.073		
NP type	0.331	0.118	2.796	5.431	.020*
Non-anaphoric					
(intercept)	0.030	0.089	0.340		
NP type	−0.615	0.234	−2.624	5.223	.022*
Bridging					
(intercept)	−0.508	0.150	−3.384		
NP type	>−0.01	0.164	−0.007	1.040	.308

Note. Formula: z-rating ~ NP type + (NP type | participant) + (NP type | item). Coding: NP type: *sono* NP = −0.5 vs. bare NP = 0.5. * $p < .05$. *P* values were calculated for χ^2 values obtained through comparisons of the full vs. null model in terms of the fix effect of *NP type* by means of the likelihood-ratio testing.

These results are compatible with the agreed properties of *sono* in terms of the infelicity of *sono* in the non-anaphoric context, the felicity thereof in the bridging context, and the preference for overt anaphoric definite marking with *sono* in the anaphoric context. As to the disputed property, the results suggest that bare NPs are comparable to *sono* NPs in acceptability in the bridging context.

6.4.2.2 NQ constructions

The mean ratings of the two types of NQ construction are presented in Table 6.5 and Figure 6.3. The results were generally in line with the predictions: floating NQs were rated clearly lower than post-nominal NQs by about 1.5 in the [+definite] context whereas both types of NQ were acceptable to similar degrees (4.2 or higher) in the

[−definite] context.

Table 6.5 Mean acceptability ratings for NQ construction *items: Pilot AJT ver. 2 (SD in parentheses)*

	[+def, (+spec)]	[−def, (−spec)]
Post-nominal	4.91 (1.68)	4.67 (1.68)
Floating	3.45 (1.86)	4.21 (1.65)

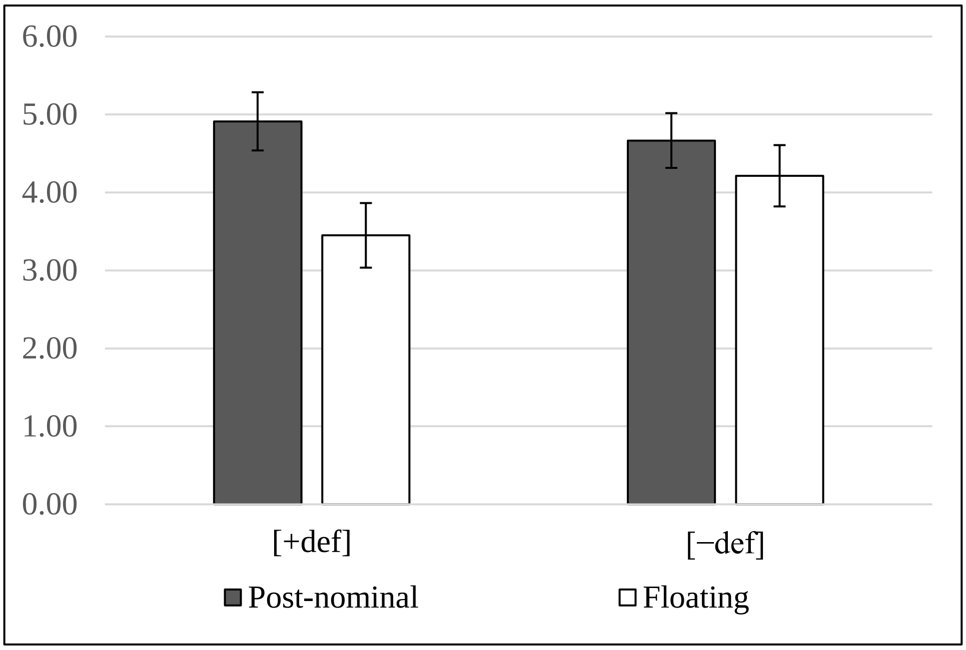


Figure 6.3 Mean ratings for NQ construction items (error bars = SE)

The effect of NQ type on acceptability ratings in each context was analysed using linear mixed-effects modelling following the procedures above. The statistical significance of the fixed effect of NQ type on rating was computed through comparison of relevant full and null mixed-effects models. The outputs for each critical model and model comparison can be found in Tables 6.6. The results showed that the effect of NQ type was significant in the [+definite] context and marginally so in the [−definite] context.

Table 6.6 Fixed effect estimates of critical linear mixed-effects models and likelihood-ratio testing results of critical vs. null model comparisons for NQ construction items: Pilot AJT ver. 2

Fixed effects	β	*SE*	*t*	χ^2	*p*
[+definite]					
(intercept)	−0.548	0.089	−6.166		
quantifier type	0.639	0.132	4.840	11.163	<.001***
[−definite]					
(intercept)	−0.649	0.096	−6.783		
quantifier type	0.277	0.135	2.047	3.202	.0074 †

Note. Formula: z-rating ~ quantifier type + (quantifier type | participant) + (quantifier type | item). Coding: quantifier type: post-nominal = −0.5 vs. floating = 0.5. *** $p < .001$, † $p < .10$. *P* values were calculated for χ^2 values obtained through comparisons of the full vs. null model in terms of the fix effect of *quantifier type* by means of the likelihood-ratio testing.

Therefore, the results of the previous AJT were replicated in that floating NQs were clearly rated lower than post-nominal NQs in the [+definite] context. As to the [−definite] context, however, although the difference here did not reach the threshold of statistical significance (α = .05), the marginal level of difference was not expected. A potential source of this unexpected small rating gap will be discussed in the next section.

6.4.3 Discussion

The results support the properties of *sono* and NQs agreed in the literature, namely *sono*'s incompatibility with non-anaphoric contexts, preference for *sono* NPs over bare NPs in anaphoric contexts, *sono*'s felicity in bridging contexts, and the forced indefiniteness interpretation of floating NQs. Regarding the disputed property of *sono,* namely whether *sono* NPs are preferred over bare NPs in bridging contexts, there was virtually no difference between the two NP types in felicity.[48] On the other hand, there was one unpredicted result obtained for NQs, namely the slight preference for post-nominal NQs over floating NQs in the [−definite] context. A thorough examination of mean ratings of individual items in the relevant context revealed that the unexpected results can be attributed to an effect of particular vocabulary choices. The relevant items were revised in terms of word choice and collocations in the hope of minimizing such effects. After revising the problematic elements, native Japanese informants reported that they would rate the sentences as predicted. Therefore, I conclude that the pilot AJT findings generally supported the descriptions of the agreed target linguistic properties in the literature. That is, these properties seem on the whole to be reliably measurable at least through an offline judgement task.

6.5 Pilot SPRTs ver. 1 & ver. 2

As mentioned earlier, two pilot versions of SPRT were conducted. The primary goal of these experiments was to examine whether native Japanese speakers show sensitivity to the Japanese properties in question in an implicit way, namely in the form of reading times (RTs). The first pilot SPRT did not capture the target reading effects. This motivated revisions and the second pilot SPRT.

In the remainder of the section, after briefly stating some key assumptions in the self-paced reading paradigm and presenting characteristics of Japanese orthography, I will describe the postulated processing mechanisms for the target linguistic properties, and give an overview of the designs of the two SPRTs and their findings.

48 Pilot AJT ver. 1 had essentially similar results except that the preference for *sono* NPs over bare NPs in anaphoric contexts were not statistically supported (not reported for conciseness).

6.5.1 Key assumptions in self-paced reading

In an SPRT, participants are asked to read sentences for comprehension at a natural speed. Sentences are presented in small segments such as words or phrases on a computer screen one at a time. Participants press a key to read the next segment and repeat this to the end of the sentence. The present study employs the moving widow paradigm, whereby the segments appear left to right across the screen (Just, Carpenter, & Woolley, 1982). RTs in each segment are recorded and compared across test conditions. The crucial assumption is that some part of the test sentence can cause reading slowdown due to unexpectedness such as ungrammaticality or a semantic anomaly (i.e., infelicity). For example, Pearlmutter, Garnsey, and Bock (1999) found from their SPRT that it took native speakers of English significantly longer to process the word *rusty* in the ungrammatical sentence (6.19b) compared to its grammatically correct version (6.19a). That is, the English natives were aware of number agreement violation between the subject NP and the verb while reading the ungrammatical sentence, which resulted in a delayed RT. The occurrence of the slowdown on the word *rusty* after the ungrammatical segment *were* is known as a spillover effect: slowdowns in SPRTs may occur at the point of anomaly or one or two segments after it.

(6.19) a. The key to the cabinet <u>was</u> rusty from many years of disuse.

b. The key to the cabinets <u>were</u> rusty from many years of disuse.

Such a slowdown effect is taken as evidence of real-time processing of the linguistic anomaly (a subject-verb number agreement error, in this case). Therefore, in the case of the present study, the incompatibility of *sono* NPs with non-anaphoric contexts and the incompatibility of the floating NQ construction with definite contexts attested in

Pilot AJT ver. 1 may materialise as reading slowdowns in an SPRT.

6.5.2 Japanese orthography

Before detailing the method of the pilot SPRTs, this section provides some information about Japanese script that is relevant to all the SPRTs reported in the present thesis, where test items were all displayed in Japanese script. Japanese has two modes of orthographic system: *kana* (syllabic system like English alphabet) and *kanji* (Chinese-origin ideographic system). *Kana* are further divided into two sub-types: *hiragana* (cursive *kana* e.g., あ) and *katakana* (square *kana* e.g., ア). The former is used for highly frequent morphemes such as case particles, postpositions, and inflectional endings whereas the latter is used for loan words except those from Chinese, onomatopoetic words and so on. Kanji are used for the roots of nouns, adjectives, adverbs, and verbs. Normally both kana and kanji are mixed together in Japanese writing as illustrated in (6.20) (note that there is no space between symbols in actual Japanese writing).

	%	#	*	#	* #	
(6.20)	ジョン	が	本	を	読む。	(* = kanji # = hiragana % = katakana)
	Jon	-ga	hon	-o	yomu.	
	jon	-NOM	book	-ACC	read	
	‘John reads books.’					

For uncommon kanji, a reading aid called *furigana* is often provided in the form of small hiragana written above kanji. For example, in the case of がくせい 学生, 学生 is kanji for the Japanese word meaning ‘student’, and the small symbol above the kanji, がくせい is the furigana which provides the pronunciation of the kanji, /ga-ku-se-i/. In principle, furigana is optional. In the present study, all kanji were presented with furigana in all the experiment items for learner-friendliness, though native Japanese participants were expected to be highly familiar with the reading of all relevant kanji.

6.5.3 Processing of *sono*

Processing of *sono* seems relatively straightforward. In principle, the task is to establish whether the entity denoted by the *sono* NP has a proper antecedent in the previous discourse. That is, when the target NP cannot find an antecedent in the previous discourse, participants may slow down at, or just after, this point. For example, in [−anaphoric] contexts as in (6.21), RTs are predicted to be longer when the target noun is presented in the form of *sono* NP than its corresponding bare NP.

(6.21) Non-anaphoric contexts

PREAMBLE: *Taroo likes walking at night. He has many ways to enjoy it. On a cloudless night...*

Taroo-wa	yoku	{#sono tuki-o vs. tuki-o}	uresisoni
Taroo-TOP	often	{SONO moon-ACC vs. moon-ACC}	gladly

yukkurito	nagamemasu.
for.a.long.time	gaze

‘Taroo often enjoys taking a long look at the moon. ’

6.5.4 Processing of NQ constructions

Based on the results of Pilot AJT ver. 1, I assume, following previous studies categorised as Prediction C (e.g., Kobayashi & Yoshida, 2001; Nakanishi, 2007; Shin, 2017), that a post-nominal NQ modifies the host noun within a nominal structure (6.22a) whereas the NQ functions as a verbal modifier in the floating construction

(6.22b) (as discussed in Chapter 3).[49]

(6.22) a. Structure of post-nominal NQ construction

[$_{DP}$ (external argument)] [$_{VP}$ [$_{V'}$ [$_{DP}$ [$_{N'}$ N (internal argument) **NQ**]] V]]

b. Structure of floating NQ construction

[$_{DP}$ (external argument)] [$_{VP}$ [$_{DP}$ (internal argument)] [$_{V'}$ **NQ** [$_{V'}$ V]]

(Based on Nakanishi, 2007)

In the processing of the NQ constructions, the position of the case marker seems to play a crucial role. Given the structural analyses in (6.22), the position of the case marker determines whether an NQ is post-nominal or floating since it informs the comprehender of where a nominal projection closes. For example, when the NQ comes between the host noun and the accusative case particle *-o* as in (6.23a), one can confirm that the NQ and the host noun form a constituent, hence post-nominal. In this case, the host nominal can be interpreted as either indefinitely or definitely. By contrast, when the NQ follows the case marker as in (6.23b), it can be analysed as a floating NQ, thus the indefiniteness constraint applies.

(6.23) a. Post-nominal

Taroo-wa moo [$_{VP}$ [$_{DP}$ *hon* [*san-satu*]-*o*] [$_{V}$ yonda]].

Taroo-TOP already book 3-CL-ACC read

'Taroo has already read (the) three books.'

---

49 As discussed in Chapter 3, it is still disputable whether a post-nominal NQ is an NP-adjunct (Nakanishi, 2007) or a nominal head taking an NP as its complement (Shin, 2017). Although this question is beyond the scope of the present thesis, I adopt Nakanishi's (2007) structure for convenience.

b. Floating

Taroo-wa moo [VP [DP hon-o] [V' [san-satu] [V yonda]]].

Taroo TOP already book-ACC 3-CL read

'Taroo has already read three books.'

The earliest point possible at which the indefiniteness constraint on floating NQs could take effect (in the form of reading slowdown) would be then at the NQ after the case marker. This prediction is based on the structural analyses in (6.22) as well as the assumption that native Japanese speakers can develop some kind of under-specified representation for a VP without knowing what exactly the verb is so that they can accommodate case-marked NPs already recognised. This assumption is supported by the well-documented observation from processing studies that native Japanese speakers start parsing sentences via case marking information even before encountering verbs (e.g., Kamide & Mitchell, 1999; Miyamoto, 2002).[50]

6.5.5 Method

6.5.5.1 Participants

Twenty-six native speakers of Japanese participated in Pilot SPRT ver. 1. The majority of them were international students on a 16-week short course at a UK university. Another group of twenty Japanese-speaking participants were recruited for Pilot SPRT ver. 2. They were mostly university students attending undergraduate, post-graduate programmes or a 4-week course at the same university as the Pilot SPRT ver. 1 participants.

---

50 This also appears to be the case for speakers of such languages as German (e.g., Bader & Lasser, 1994), Turkish (e.g., Kahraman, Sato, Ono, & Sasaki, 2010), and Finnish (Hyönä & Hujanen, 1997).

6.5.5.2 Test instruments

The test sentences used in Pilot SPRT ver. 1 were essentially similar to those used in Pilot AJT ver. 1, incorporating the revisions outlined in section 6.3.3. The test items for Pilot SPRT ver. 2 were essentially an SPRT adaptation of Pilot AJT ver. 2.

Both SPRTs were designed to examine the processing of the relevant properties of the demonstrative *sono* and the NQ constructions, using *PsychoPy* (Peirce et al., 2019). For each trial, participants read a short passage that established context and proceeded to read a test sentence as a continuation of the preceding passage. Test sentences were divided into six segments (NP—adjunct—NP—adjunct—adjunct—VP) in Pilot SPRT ver. 1 and into 7 segments (NP—adjunct—NP + adjunct—VP + adjunct—adjunct—adjunct—VP) in Pilot SPRT ver. 2. The third segment in both versions was designed to be the critical segment, where the predicted reading effect (slowdown) may start to emerge, whereas the immediately following two (ver. 1) or three (ver. 2) segments served as spillover regions. The test battery included five critical contexts: anaphoric, non-anaphoric, bridging contexts for *sono*, and [+definite] and [−definite] contexts for the NQ constructions. (6.24)–(6.28) are sample test items from Pilot SPRT ver. 2 ("|" indicates a segment boundary). Reading slowdown was predicted for the *sono* NP condition in the non-anaphoric context (6.25) and for the floating condition in the [+definite] context (6.27), in the form of greater RTs compared to the alternative condition. On the other hand, based on the slight preference of *sono* NPs over bare NPs in the anaphoric context found in Pilot AJT ver. 2, it was predicted that reading facilitation (i.e., shorter RTs compared to the alternative condition) might take place in the *sono* NP condition in that context (6.24).

(6.24) *Sono*: anaphoric (*sono* NP vs. bare NP)

PREAMBLE: *There is a restaurant popular among local people nearby Taroo's house.*

| Taroo-wa | mukasi kara |

Taroo-TOP from.long.ago

{ sono resutoran-o vs. resutoran-o} dare-yori-mo |

{ SONO restaurant-ACC vs. restaurant-ACC } more.than.anybody

totemo kiniitteiru to | mawari ni | yoku | katatte-imasu. |

very like that surroundings to often tell-ASP.POL.NPST

'Taroo often tells people around him that he has liked the restaurant more than anybody, for a long time.'

(6.25) *Sono*: non-anaphoric (*sono* NP vs. bare NP)

PREAMBLE: *Taroo likes walking at night. He has many ways to enjoy it. On a cloudless night...*

| Taroo-wa | mukasi kara |

Taroo-TOP from.long.ago

{ sono tuki-o vs. tuki-o } yukkuri | nagamete tanosindeiru-no-o |

{ SONO moon-ACC vs. moon-ACC } slowly watch enjoying-COMP-ACC

mawari-ni-wa | amari | hanasite-imasen. |

people.around-to-TOP much tell-ASP.POL.NEG.NPST

'Taroo has not told many people around him that, for years now, he has been enjoying taking a long look at the moon.'

(6.26) *Sono*: bridging (i.e., implicit anaphoric) (*sono* NP vs. bare NP)

PREAMBLE: *There is a novel that Hanako is really into right now.*

| Hanako-wa | guuzen | { sono tyosya-o vs. tyosya-o } Tokyo de |

Hanako-TOP accidentally {SONO author-ACC vs. author-ACC } Tokyo in

issyun mikaketa koto-ga-aru to | mawari ni |

moment saw thing-NOM-have.done that surroundings to

uresisooni | hanasite-imasu.|

happily tell-ASP.POL.NPST

'Hanako is happily telling people around her that she happened to catch sight of the author in Tokyo.'

(6.27) NQs: [+definite] (post-nominal vs. floating)

PREAMBLE: *Taroo is a good tennis player. He played games with his friends, Takasi, Hirosi, and Goroo yesterday.*

| Taroo-wa | itumo-no-yoni |

Taroo-TOP as.always

|{tomodati san-nin-o kantanni vs. tomodati-o kantanni san-nin} |

{friend 3-CL-ACC easily vs. friend-ACC easily 3-CL }

sugu makasite-simatta to | tenisu-kooti ni | uresisooni | ii-masita. |

quickly beat-finished that tennis.coach to happily tell-POL.PST

'Taroo happily told his tennis coach that he, as always, beat the three friends quickly and easily.'

(6.28) NQs: [−definite] (post-nominal vs. floating)

PREAMBLE: *Taroo works as a vet at an animal hospital. Although he mostly treats dogs, ...*

| Taroo-wa | senzitu |

Taroo-TOP the.other.day

| { neko san-biki-o mezurasiku vs. neko-o mezurasiku san-biki } |

{ cat 3-CL-ACC unusually vs. cat-ACC unusually 3-CL }

tuzukete tiryoo-sita to | yuuzin ni tanosisooni ii-masita. |

in.a.row treatment-did that friend to delightedly say-POL.PST

'Taroo delightedly told his friends that, unusually for him, he treated three cats in a row the other day.'

One main difference between the two versions of SPRT lies in the lengths of the critical and the spillover segments: ver. 2 had longer segments than ver. 1, extended with extra adverbials or adjuncts. This revision was motivated by the generally null results of Pilot SPRT ver. 1 (to be outlined in the next subsection section). Specifically, the longer segmentation in ver. 2 than ver. 1 was based on the assumption that the longer the segment is, the easier it might be to observe discourse-sensitive properties such as those tested in the present thesis, because it allows participants to read stimuli in a way that better approximates their standard reading behaviour.

Another important difference between the two SPRTs is that ver. 2 included four new types of filler, as mentioned above for AJT Pilot ver. 2. These fillers were presented in two minimally different versions, acceptable vs. unacceptable. Among those filler types were (i) NPs with wrong classifiers (unacceptable) vs. NPs with appropriate classifiers (acceptable) (6.29); (ii) lack of an associative plural reading (i.e., *x* and others associated with *x*) for NPs-*tati* in the pre-nominal NQ construction (unacceptable) vs. availability thereof in the post-nominal NQ construction (acceptable) (6.30); (iii) bare NPs in associative plural contexts (unacceptable) vs. NPs-*tati* (acceptable) (6.31); and (iv) (scopal) specificity constraint on NPs-*tati* (unacceptable) vs. absence thereof on bare NPs (acceptable) (6.32).[51] These fillers were intended to serve as control conditions to check the validity of the experimental design in case Pilot SPRT ver. 2 failed to find statistical evidence for the target properties of *sono* and NQs. Slowdown effects were predicted in these fillers in each unacceptable condition (marked by "#").[52]

(6.29) Noun-numeral classifier agreement: (correct vs. # wrong)

PREAMBLE: *Taroo likes keeping small creatures.*

| Taroo-wa | saikin | kaeru-o ie-de {ni-hiki vs. huta-ri} |

Taroo-TOP recently frog-ACC house-at {2-CL vs. 2-CL}

---

[51] It has been observed that in a referentially opaque context, where an NP is situated in a sentence with a scope-taking element (i.e., an operator) such as an intensional verbs (e.g., want, look for), an NP suffixed by -*tati* necessarily takes wide scope (=scopally specific) (e.g., Nakanishi & Tomioka, 2004). For example, in (i), *kangohu-tati* 'nurses' refers to specific nurses (i.e., wide scope reading) and thus sounds infelicitous when interpreted as non-specific nurses (i.e., narrow scope reading).

(i) Sono byooin-wa kangohu-tati-o sagasi-teiru
SONO hospital-TOP nurse-TATI-ACC look for-PROG

narrow scope reading: ✓look-for > nurse(s)
'That hospital is looking for a nurse/nurses (to hire).'

wide scope reading:*? nurse(s) > look-for
'There is a nurse/are nurses that hospital is looking for.'

(Adapted from Nakanishi & Tomioka, 2004, p.115)

[52] Note that the test items for the third control filler (associative reading of common nouns with -*tati*) (6.31) were minimally different within the preamble rather than in the segmented sentence.

kossori | katte-ite | mainiti | takaramono-no yooni |

secretly keep-ing everyday like.treasure

taisetu-ni | site-imasu. |

importantly do-ing

'Taroo has recently been keeping two frogs at home secretly, and he treasures them.'

(6.30) Associative reading of proper names with *-tati* (post-nominal vs. #pre-nominal)

PREAMBLE: *Hanako went to see her uncle living in the countryside, with her parents by train.*

| Oji-wa | yuugata-kara |

uncle-TOP evening-since

| { Hanako-tati san-nin-o vs. san-nin-no Hanako-tati-o } umi-ni | doraibu-ni

{ Hanako-TATI 3-CL-ACC vs. 3-CL-GEN Hanako-TATI-ACC } sea-to drive-for

turete-itte-kurete | sonomama | ie-made | okutte-kure-masita. |

take-go-gave directly home-to send-gave-did

'Her uncle took the group of three, Hanako and the others (her parents), for a drive to the seaside in the evening and drove them home from there.'

(6.31) Associative reading of common nouns with *-tati*: (collective vs. #singular)

PREAMBLE: *Taroo has a beloved* {*wife and daughter* vs. *wife*}.

| Taroo-wa | kono natu | tuma-tati-o kaigai ni | ryokoo ni

Taroo-TOP this summer wife-TATI-ACC abroad trip to turete-iku yotei da to | yuuzin ni | uresisooni | hanasite-imasu. |

take-go plan COP that friend to happily speak-ASP.POL.NPST

'Taroo is happily telling a friend that he is going to take his wife and others (his daughter) on a trip abroad this summer.'

(6.32) Specificity of NP-*tati*: (bare NP vs. # NP-tati)

PREAMBLE: *Hanako is dating a man and they are going to marry soon.*

| Hanako-wa | deki-reba | { kodomo-o vs. kodomo-tati-o } huta-ri |

Hanako-TOP if.possible { child-ACC vs. child-TATI-ACC } 2-CL

sanjyus-sai-made-ni umi-tai to | senzitu | yuuzin ni | hanasi-masita. |

30-year.old-until-by want.to.have that the.other.day friend to tell-POL.PST

'Hanako told a friend the other day that she wants to have two children by the age of 30, if possible.'

Here, I give details of two of the filler types because of their relevance to later discussions. The first concerns the agreement between NPs and classifiers (CLs). In (6.29), native Japanese speakers were predicted to read the target sentence significantly slower when the object noun (*kaeru* 'frog') was followed by the wrong CL (*ri,* CL for humans) ("wrong" condition) than when associated with the correct one (*hiki,* CL for small animals) ("correct" condition). The second control filler in question is related to the (un)availability of an associative reading of the Japanese plural suffix -*tati* in the pre-nominal and post-nominal NQ constructions. It has been pointed out in the literature that an associative reading of an NP-*tati* (i.e., *x* and others associated with *x*) is available when the NP is embedded in the post-nominal NQ construction but not so in the pre-nominal construction, where only a regular plural reading is possible (Ochi, 2012). In (6.30), *Hanako-tati san-nin* (Hanako-TATI 3-CL) (i.e., post-nominal) refers to a group of three people, namely Hanako and two others (=her parents) whereas *san-nin-no Hanako-tati* (3-CL-GEN Hanako-TATI) (i.e., pre-nominal) cannot refer to Hanako and two others; it must refer to three people named *Hanako* (i.e., plural reading). Thus, in a context where an associative reading is favoured (i.e., *Hanako and the others* (=her parents), greater reading time (slowdown) is expected in the pre-nominal condition than in the post-nominal condition.

The test items with two conditions were divided into two lists with each containing only one version of a given item in order to prevent participants from directly comparing the two types under the same context, as in the AJTs. Pilot SPRT ver. 1 had 64 items in total for each list, consisting of 40 critical items (8 items × 5 types) and 24 unacceptable fillers (6 types × 4 items), whereas in Pilot SPRT ver. 2 each list contained 96 items, namely 40 critical items (8 items × 5 types), 32 control fillers (8 items × 4 types), and 24 ungrammatical fillers (4 items × 6 types), as in Pilot AJT ver. 2.[53] The test items were pseudo-randomised for each participant in both SPRTs.

Comprehension questions were asked at the end of trials, in order to encourage participants to read for meaning, and to avoid focusing attention on the target linguistic properties. In Pilot SPRT ver. 1, a question was given after every test item. However, in Pilot SPRT ver. 2, the number of comprehension questions was reduced to one third of the test items (roughly one third of the questions were asked about the critical items and the rest were asked about the fillers) with a view to

[53] Details of the unacceptable fillers are not given due to space limitations. They were the same across the item lists and were included for balancing the number of the acceptable and unacceptable items, as in the AJTs.

reducing the task load in light of the increased number of test items (i.e., from 64 to 96).[54] The comprehension questions took the form of a sentence completion as in (6.33) following Smith (2016), where participants were presented with two statements from which they chose the accurate description of the contents of the test item. For the items with slowdown effects predicted (i.e., those with an unacceptable target sentence), the question concerned the contextual information rather than the target sentence, as exemplified in (6.33). Otherwise, the question was asked about the target sentence. The display position of the target answer (top or bottom option) was evenly distributed in both SPRTs.

(6.33) Question about the context in (6.27)

Kono bunsyoo ni yoruto···('According to this passage,…')

(▲) Taroo-wa yuuzin ni tenisu de katimasita

'Taroo beat his friends at tennis.'

(▼) Taroo-wa yuuzin ni gorufu de katimasita

'Taroo beat his friends at golf.'

6.5.5.3 Procedure

Participants were tested individually on a computer in quiet locations (e.g., library study rooms). They were instructed to read for comprehension at a natural speed. Before the main trials, participants completed four practice items to familiarize themselves with the task format. The participants completed the whole task in approximately 25 minutes in ver. 1 and 30 minutes in ver. 2. The participants received

54 Incorporating comprehension questions after just one third of the of the test items is a common practice in processing research (e.g., Millin, Divjak, & Baayen, 2017; Roberts & Felser, 2011; Trueswell, Tanenhaus, & Kello, 1993).

thanks payment in cash for participation.

6.5.6 Results

6.5.6.1 SPRT data preparation

The collected RTs in both SPRTs went through a series of treatments, as below.

Response screening

In an attempt to remove unreliable data potentially due to participants' low commitment to the task, responses were first screened on the basis of accuracy on the comprehension questions. 70% accuracy was chosen as the cut-off, following common practice in processing studies (e.g., Foote, 2011; Lago, Shalom, Sigman, Lau, & Phillips, 2015; Tucker, Idrissi, & Almeida, 2015). All participants in both SPRTs scored above this cut-off.

Data trimming

Raw RTs were trimmed by removing outlying values in order to prevent extreme data points from adversely altering the data distribution. RTs lower than 100 milliseconds (ms) were eliminated because such short reading times are generally impossible (Luce, 1986): they most likely suggest that the participant simply did not read the text or unintentionally pressed the button (Jegerski, 2014). RTs over 3,000 ms were also excluded because such data often reflect a lack of attention (e.g., Havik, Roberts, Van Hout, Schreuder, & Haverkort, 2009; Roberts & Felser, 2011).[55] The elimination of relevant RTs affected 3.23 % (< 100 ms: 3.04 %; > 3,000 ms: 0.19 %) in ver. 1 and

[55] Conventionally, residual RTs beyond 2–3 standard deviations from the mean for a given condition and position for each participant are often removed. However, this method was not applicable to the present study: because each participant was allocated only 4 tokens for each condition, all RTs fell within two standard deviations from the mean.

affected 5.72 % of the data (< 100 ms: 5.37 %; > 3,000 ms: 0.35 %) in ver. 2.

Data transformation

The remaining raw RTs were converted into residual reading times (RRTs) in order to control for differences in word-length between stimuli and to minimise the influence of individual variation in reading speed (e.g., Ferreira & Clifton, 1986; Trueswell, Tanenhaus, & Garnsey, 1994). The number of morae (the standard phonological unit to measure Japanese words) was used as the unit of word length rather than the number of syllables. This was because there is some evidence that the number of morae can better predict RTs compared to the number of syllables (Sawasaki, 2007).[56] Specifically, expected RTs per mora were first calculated for each participant by using the linear regression formula in (6.34a). The parameters in the formula, *a* (= intercept) and *b* (= slope) vary among the participants. Then, the RRTs were obtained by subtracting the expected RTs from the raw RTs as shown (6.34b)

(6.34) a. Expected RT = a + b * (number of morae)

b. Residual RT (RRT) = Raw RT – Expected RT

Sawasaki (2007, p. 57)

Using RRTs instead of RTs is particularly important for the comparison of *sono* and bare NPs within *sono* items (6.24–6.26) because, with *sono* NPs being two morae longer than their bare NP counterparts, potential longer reading times for *sono* NPs may simply be attributed to a word length effect. (This residualisation method was also used in the main study reported in Chapter 7.)

[56] For example, a Japanese word such as *syuutyuu* 'concentration' constitutes four phonological units in morae, syu-u-tyu-u, while the same word is considered only two phonological units long in syllables, syuu-tyuu.

6.5.6.2 Summary of results

For both SPRTs, visual inspections of RRT data did not detect any noticeable evidence of the target properties, namely higher RRTs for *sono* NPs than bare NPs in the non-anaphoric context and higher RRTs for floating NQs than post-nominal NQs in the definite context (i.e., reading slowdown). The potential reading facilitation effects, namely lower RRTs for *sono* NPs than bare NPs in the anaphoric context were not observed, either. For further analysis, the effect of condition on RRTs was examined by means of linear mixed-effects modelling for each context and each relevant segment. A comparison of the critical model with a maximal random effect structure with its null model (= the critical model minus the fixed effect of condition) did not find any significant effect for any property in either ver. 1 or ver. 2. However, in Pilot SPRT ver. 2, the native Japanese speakers showed a predicted slowdown to statistically marginal levels for two of the control filler types, namely the ones that were highlighted above, concerning noun-classifier agreements (6.29) and the (un)availability of the associative reading of NPs-*tati* (6.30).[57]

The results of both SPRTs were not at odds with the predictions, in that no counter-evidence was found for the predictions. However, no statistically significant data were obtained to confirm the measurability of the target properties of *sono* and NQs. This suggests that the revised segmentation and other minor adjustments applied to Pilot SPRT ver. 2 were not successful. Nevertheless, the fact that marginally significant predicted slowdown effects observed with some control fillers could suggest that the general design of the SPRTs is appropriate, but that the SPRTs may not be suitable for measuring the target properties of *sono* and the NQ constructions.

[57] Since these null findings are all from pilot studies, it was decided it would not be good use of space to include all of the relevant reading time plots and model outputs here. Hence, they are not presented here. However, the relevant RRTs plots in the Pilot SPRTs can be found in Appendix 4.

Alternatively, the absence of statistical evidence for the predicted differences may be partially due to low statistical power resulting from the relatively small numbers of participants and of test items. The implications of these null pilot study results for the main study SPRT are considered in the discussion section, following.

6.6 General discussion

6.6.1 Key findings of the pilot studies

First and foremost, the crucial finding of the series of the AJT studies is that the target properties repeated as (6.35–6.36) were attested.

(6.35) Agreed properties of the demonstrative *sono*

- The use of *sono* is preferred in anaphoric contexts over bare forms.
- The use of *sono* is infelicitous in non-anaphoric (i.e., unique) definite contexts.
- The use of *sono* is felicitous in bridging definite contexts.

(6.36) Agreed properties of the NQ constructions

- Floating NQs are allowed only in [−definite] contexts.
- Post-nominal NQs are compatible with [+definite] contexts.

In terms of the disputed properties repeated below (6.37–6.38), the following conclusion can be drawn. Regarding the preference for definiteness-marking by *sono* in bridging anaphoric contexts, there was no evidence for the preference for *sono* NPs over corresponding bare NP: both NP types seem equally acceptable. This is incompatible with the proposal that the use of *sono* is strongly preferred in bridging anaphoric contexts (e.g., Tsutsumi, 2012).[58] As to the NQs, the data from the first AJT show that floating NQs are highly acceptable in [+specific, −definite] contexts, and so are post-nominal NQs in [−specific, −definite] contexts. This pattern is compatible only with Prediction C, the floating-NQs-as-adverbial account (e.g., Kobayashi & Yoshimoto, 2001; Nakanishi, 2007; Shin, 2017) (see Chapter 3), which suggests NQs are constrained by definiteness but not specificity. Therefore, it can be argued that indefiniteness rather than non-specificity is covertly realised in the floating NQ construction in Japanese.

(6.37) Disputed properties of the demonstrative *sono*

- Whether the use of *sono* is preferred over bare forms in bridging contexts.

(6.38) Disputed properties of the NQ constructions

- Whether floating NQs are acceptable in [+specific, −definite] contexts.
- Whether post-nominal NQs are acceptable in [−specific, −definite] contexts.

In contrast to the clear-cut AJT results, the self-paced reading experiments did not produce robust evidence of the target properties above. Particularly, the expected slowdown effects were not observed, although the data were not in contradiction to the AJT results. Nevertheless, native-Japanese-speaking participants were sensitive at marginal levels to some non-target properties (the control fillers regarding the semantic agreement between nouns and classifiers and the availability

[58] Unlike the pilot AJTs, statistically significant preference for *sono* was found in bridging contexts in the native Japanese data in main studies (to be reported in Chapters 7 and 8). However, this preference was not so strong (similar to the preference of *sono* in the anaphoric context in Pilot AJT ver. 2) and bare nominals were still highly acceptable, contra the proposal.

of the associative reading of NPs-*tati*). Those marginally significant results indicate that the test instrument itself is valid, but the discourse-dependent target properties may be difficult to capture by means of the self-paced reading paradigm. This suggests that an AJT is a more suitable method than an SPRT for the investigation of the target properties at least as far as native speakers are concerned.

However, with further improvements on the test items (see below) and larger data sets, there is still a chance to obtain solid evidence for the target properties. Furthermore, even if the native SPRT data remains unclear, L2 learners from different L1 backgrounds (i.e., English and Korean) might reveal potentially interesting crosslinguistic differences. Additionally, previous findings that L2 learners can demonstrate clearer effects in an SPRT than in an AJT (e.g., Hopp, 2009; Ionin, Choi, & Liu, 2019; Orfitelli & Polinsky, 2017) motivate combining the two methods for the L2 research in the present study, too. Therefore, the test instruments (Pilot AJT ver. 2 and Pilot SPRT ver. 2) were considered appropriate for main L2 studies, following some further revisions discussed in the next section.

6.6.2 Further revisions for main studies

First, it was decided that identical test sentences were to be used for both the AJT and the SPRT in the hope of reducing the risk of unqualifiable variation in data. (Recall that the test sentences of Piot AJT ver. 2 and Pilot SPRT ver. 2 were essentially similar but slightly different, as can be seen through a comparison of example items of the two experiments above (e.g., (6.15) and (6.25)): target sentences were slightly longer and structurally more complexed in Pilot SPRT ver. 2.) Second, the items with relatively low ratings in the acceptable condition in the AJT were revised, through extensive discussions with some native Japanese informants, in terms of collocations and word choices, in order to make them sound more natural. A third change concerned extra adverbials or adjuncts used to extend the test sentences in Pilot AJT ver. 2 and Pilot SPRT ver. 2. They were removed for the purpose of reducing the overall task load, given that the desired outcome of this modification was not achieved (i.e., it did not lead to clear reading slowdown effects in SPRT ver. 2). Finally, the test vocabulary was adjusted, using a vocabulary list for Japanese language education, named *nihongo kyoiku goi hyo ver. 1.0* 'a list of words for Japanese language education ver. 1.0' (http://jhlee.sakura.ne.jp/JEV/) (Nihongo Gakushu Jisho Shien Group [Japanese learning support group], 2015), for learner-friendliness.[59] Specifically, when a given word was not found on the list, it was replaced with a listed word with a similar meaning. Turning to the issue of statistical power, the main studies aimed to recruit more participants. However, the number of items was not increased because it was already relatively large (96 items per list). For example, if the present study followed Keating and Jegerski's (2015) guideline of 16–24 items for each critical context, the total number would rise to at least double the number (i.e., 192 or more items). Given that participants have to read a preamble in addition to the target sentence, this large number of items would cause too much burden on participants, which might compromise the validity of data.

6.6.3 Implications for acquisition tasks and learnability predictions

The key properties of the target phenomena were all confirmed in the AJT data. This means that no modification is needed to the acquisition tasks and learnability predictions presented in Chapter 5. However, native intuitions turned out to be relatively subtle with what were designed as unacceptable conditions, particularly the floating NQ construction in definite contexts. Specifically, the Japanese native

[59] This list contains 18,000 words sorted by part of speech and target proficiency levels, compiled based on *Balanced Corpus of Contemporary Written Japanese* (Maekawa et al., 2014) and 100 Japanese language textbooks.

speakers showed indeterminate judgments, rating around the mid-point on the scale. This suggests that the target intuitions are themselves quite subtle. Such subtlety might mean greater difficulty in acquiring the properties and/or in judging their acceptability in experimental settings for L2 learners, particularly English-speaking learners, who are predicted to tackle feature reassembly tasks.

Chapter 7: Main study 1 (non-web based SPRT and AJT with L2 Japanese learners)

7.1 Introduction

This chapter reports on a study with L2 learners of Japanese as well as newly recruited native Japanese speakers as controls that involves a self-paced reading task (SPRT) and an acceptability judgement task (AJT), which were both developed through the careful piloting documented in the previous chapter. The next section outlines the overall experimental design, followed by descriptions of the participants and the testing instruments. Subsequently, results of each task will be presented in the order of administration. Finally, the discussion section provides a detailed examination of the results in comparison with those of the pilot studies, which will identify some potential methodological problems with the AJT that makes it difficult to interpret the collected data reliably. The discussion will lead to the conclusion that in order to effectively test the predictions about the L2 acquisition of the target properties, an additional AJT should be conducted, avoiding the methodological problems.

7.2 Method

7.2.1 Overall design

This study is comprised of three main tasks: an SPRT, an AJT, and a cloze task (i.e., proficiency test). Along with these tasks, participants were asked to complete a language background questionnaire (see Appendix 2 for the actual questionnaire). The AJT and SPRT were designed to test the target definiteness properties of the Japanese demonstrative *sono* and NQ constructions detailed in Chapter 3. The tasks were administered in the following order:

(7.1) Administration order

SPRT => language background questionnaire => cloze task => AJT

The SPRT was conducted first because its objective was to measure implicit knowledge, which could be harder to observe if preceded by other tasks more likely to induce explicit meta-linguistic knowledge (i.e., the cloze test and the AJT). The SPRT and the AJT were constructed and administered using *PsychoPy* (Peirce et al., 2019) whereas the cloze task and the background questionnaire were implemented in paper-and-pencil formats. Note that the present study ran both the SPRT and the AJT with the same participants rather than different participants. This is primarily because testing with the same individuals would make it easier to compare participants' performance between the tasks by controlling a number of potential sources of variation (e.g., proficiency) (although this thesis does not particularly address questions concerning untimed vs. real-time behaviour).

7.2.2 Participants

Thirty-eight L2 Japanese learners participated in the study, 12 with L1-Korean and 26 with L1-English.[60] The participant recruitment targeted learners with intermediate or higher Japanese language proficiency. Specifically, the study was advertised for learners who have passed N3 or higher levels on the Japanese Language Proficiency Test (JLPT) (5 levels with N5 the lowest and N1 the highest). Although there is no one-to-one equivalence between the JLPT and standardised set of proficient levels such as Common

[60] The original plan had been to have an equal number of Korean and English speakers, and data collection from Korean speakers was scheduled in both February and May 2020. However, the May data collection had to be cancelled indefinitely due to the COVID-19 pandemic.

European Frame of Reference (CEFR) (Council of Europe, 2001), N3 roughly corresponds to B1 (i.e., intermediate) on the CEFR scale (The Japan Foundation, 2017). It was considered that at least intermediate level Japanese proficiency would be needed to comprehend the test items (with both context and target sentence, as illustrated in Chapter 6). The English-speaking participants were university students in the UK who majored in the Japanese language whereas the Korean-speaking counterparts were undergraduate or graduate students at Japanese universities with a variety of majors. Table 7.1 summarises demographic information about the L2 participants.

Table 7.1 Participant information: Main study 1

Group	Age (yrs old)	Age of onset (yrs old)	Formal Japanese education (yrs)	Length of residency in Japan (yrs)
Korean speakers	23.00	17.17	5.33	2.41
(*n* = 12)	(2.98)	(2.72)	(3.15)	(1.86)
English speakers	21.46	17.19	3.85	0.92
(*n* = 26)	(1.61)	(2.56)	(2.42)	(0.62)

Note. () = *SD*.

Twenty-six native speakers of Japanese were recruited as a control group. They were undergraduate students on short-term courses at a UK university except for one postgraduate student at the same university (*M* age = 21.1, *SD* = 5.08).

7.2.3 Test instruments

7.2.3.1 Proficiency test

A cloze test adapted from Marsden (2005) was used as a measure of Japanese language proficiency.[61] Participants were asked to fill 42 blanks in a passage, by choosing the right answer from four options (see Appendix 3 for the actual cloze test). This task was administered to the native Japanese participants as well as L2 learners for reference purposes. The performance of each group is summarised in Table 7.2.

Table 7.2 Scores on cloze test (0–42 points): Main study 1

Group	*M*	*SD*	Range
Native controls (*n* = 26)	38.62 (92%)	2.17	34–42 (81–100%)
Korean speakers (*n* = 12)	36.92 (88%)	2.71	31–40 (74–95%)
English speakers (*n* = 26)	25.35 (60%)	5.57	13–31 (31–74%)

The results suggest that the L1-Korean participants can be considered generally more advanced than the L1-English participants given the large difference between the mean scores and hardly overlapping score ranges. The Korean-speaking group performed similarly to the native controls, judging from the largely overlapping score ranges and the relatively small difference between the mean scores. A one-way between-participants ANOVA confirmed that the effect of L1 on proficiency score was significant (F (2, 61) = 79.04, $p < .001$). Results of post hoc multiple comparisons by the Tukey HSD test showed that the English-speaking group differed significantly from both Korean-speaking and Japanese-speaking groups (Japanese − English: *M-Diff* = 13.27, *95% CI* [10.60, 15.93], $p < .001$; Korean – English: *M-Diff* = 11.57, *95% CI* [8.22, 14.92], $p < .001$). On the other hand, the Korean group did not significantly differ from the Japanese group. These results confirm that the Korean group were relatively advanced, scoring as high as the native control group; and the English group was less proficient than the Korean group.[62]

---

61 This method was chosen mainly due to the unavailability of concise proficiency tasks for Japanese which offer assessment comparable to a standardised set of proficiency levels such as the CEFR. It has been shown that cloze tests offer a reliable L2 proficiency measurement (e.g., Jonz, 1990).

62 As noted above, more L1-Korean participants would have been sought, had the COVID-19

7.2.3.2 SPRT

The SPRT was conducted with the same experimental design and test materials as Pilot study ver. 2 with some revisions as described in the previous chapter (Section 6.6.2). The main features of the task were as follows. Target sentences were presented in six segments non-cumulatively (the moving window paradigm). The critical region came third with two immediately following segments designed to catch spillover effects. There were 96 items per each of two lists, consisting of 40 critical items (8 items × 5 contexts), 32 control fillers (8 items × 4 types), and 24 unacceptable fillers (4 items × 6 types) to balance the acceptable-to-unacceptable ratio across the test (the same as in Pilot SPRT ver. 2). The test items were pseudo-randomised for each participant. Comprehension questions were asked for a third of the test items (32/96). Sample items from each context of the critical items are provided in (7.2)–(7.5) (segment boundaries are indicated by "|") (see Appendix 5 for the full set of items). (Note that as in the pilot, the preambles were presented in Japanese script in the actual test, although presented here in English for convenience).

(7.2) *Sono*: anaphoric (*sono* NP vs. bare NP)

PREAMBLE: *There is a very popular restaurant which serves great food for reasonable prices near Taroo's house.*

| Taroo-wa | mukasi-kara | {sono resutoran-o vs. resutoran-o} |

Taroo-TOP past-from { SONO restaurant-ACC vs. restaurant-ACC }

dare-yori-mo | yoku | riyoo-site-i-masu |

more.than.anyone often use-do-ASP-POL.NPST

'Taroo has eaten at the restaurant more often than anyone since long ago.'

pandemic not prevented this. In particular, L1-Korean participants with lower proficiency would have been sought.

(7.3) *Sono*: non-anaphoric (*sono* NP vs. bare NP)

PREAMBLE: *Hanako had a pleasant dream recently. In that dream, ...*

| Hanako-wa | tori-no yooni | { sono sora-o vs. sora-o } | jiyuuni |
Hanako-TOP like.a.bird { SONO sky-ACC vs. sky-ACC } freely

kimotiyoku | tobi-masita |
pleasantly fly-POL.PST

'Hanako enjoyed flying freely in the sky like a bird.'

(7.4) *Sono*: bridging (*sono* NP vs. bare NP)

PREAMBLE: *Hanako is now into a popular novel.*

| Hanako-wa | guuzen | { sono tyosya-o vs. tyosya-o } |
Hanako-TOP accidentally { SONO author-ACC vs. author-ACC }

Tokyo de | issyun dake | mi-ta koto-ga ari-masu |
Tokyo in one.moment only have.seen

'Hanako has caught sight of the author in Tokyo once.'

(7.5) NQs: [+definite] (post-nominal vs. floating)

PREAMBLE: *Taroo has one son and daughter and always takes care of them all by himself in the morning. ...*

| Taroo-wa | kesa mo | { kodomo huta-ri-o vs. kodomo-o huta-ri } |
Taroo-TOP this.morning too { child-2-CL-ACC vs. child-ACC 2-CL }

itumo-no yooni | siti-zi ni | okosi-masita |
as.always 7-o'clock at wake.up- POL.PST

'Taroo, as always, woke the two children up at 7 this morning.'

(7.6) NQs: [−definite] (post-nominal vs. floating)

PREAMBLE: *A new restaurant just opened near Hanako's house and she wants to go there. But, since she does not want to go there alone, ...*

| Hanako-wa | kinoo | { yuuzin huta-ri-o vs. yuuzin-o huta-ri } |

Hanako-TOP yesterday { friend-2-CL-ACC vs. friend-ACC 2-CL }

sassoku | ranti ni | sassote-mi-masita |

immediately lunch for ask.out-try.to-POL.PST

'Hanako went ahead and asked two friends out for lunch yesterday.'

The predictions for native speakers' reading patterns were formulated based on the pilot AJT findings, under the assumptions that semantic or pragmatic infelicity could lead to increased processing load (i.e., greater reading times (RTs) compared to the alternative condition); and that preference for one form over another could facilitate processing (i.e., shorter RTs compared to the less-preferred form), as described below.

- SPRT prediction 1:

  Reading facilitation might take place in the *sono* NP condition in anaphoric contexts (7.2).

- SPRT prediction 2:

  Reading slowdown was expected in the *sono* NP condition in non-anaphoric contexts (7.3).

- SPRT prediction 3:

  No RT difference was expected between the *sono* NP and bare NP conditions in bridging contexts (7.4).

- SPRT prediction 4:

  Reading slowdown was expected in the floating condition in [+definite] contexts (7.5).

- SPRT prediction 5:

  No RT difference was expected between the post-nominal and floating conditions in [−definite] contexts (7.6).

Note, however, that all of these predictions are essentially theoretical ones based on the relative (un)acceptability in the pilot AJT data because none of them were attested in the pilot SPRT data.

7.2.3.3 AJT

The AJT was conducted using the same test items as the SPRT with the following experimental design. Participants were asked to read test sentences and rate the acceptability of the target sentences as continuations of the preceding context on a 7-point scale from 0 (=complete odd) to 6 (= completely natural) with an *I don't know* option, which was added for L2 learners who might have difficulty comprehending test sentences. Unlike the pilot AJTs, this AJT was administered using a non-web-based tool (i.e., PsychoPy rather than Qualtrics): participants completed the AJT on the same computer (provided by the researcher) that they had completed the SPRT on, beforehand. Similarly to the SPRT, the test items were divided into two lists with each containing 96 items and presented pseudo-randomly. Examples (7.2)–(7.6) from the SPRT serve to illustrate the AJT items, too. The only difference was that in the AJT, the target sentences were presented as full sentences, rather than phrase-by-phrase (the full list of items can be found in Appendix 5). The context and target sentences were viewed together. Additionally, the participants could work through the test in their own time: there was no time limit for making the judgements.

The predictions for native controls' responses based on the pilot AJTs were the following.

- AJT prediction 1:

  Sono NPs would be rated significantly higher than or as high as bare NPs in anaphoric contexts (7.2).

- AJT prediction 2:

  Sono NPs would be rated significantly lower than bare NPs in non-anaphoric contexts (7.3).

- AJT prediction 3:

  Sono NPs would be rated as high as bare NPs in bridging contexts (7.4).

- AJT prediction 4:

  Floating NQs would be rated significantly lower than post-nominal NQs in [+definite] contexts (7.5).

- AJT prediction 5:

  Floating and post-nominal NQs would be rated equally high in [−definite] contexts (7.6).

7.2.4 Procedure

The participants were tested individually in quiet places (e.g., library study spaces; rental meeting rooms). After filling out the consent form preceded by the researcher's explanation about the study, they completed the three main tasks along with the language background questionnaire in the order in (7.1). The experiments were completed on a computer provided by the researcher. The SPRT and the AJT each included four practice examples. Between the cloze task and the AJT, the participants were offered to have a short break (up to 10 minutes), if they needed. The whole process took 50–60 minutes for the native Japanese speakers and 100–120 minutes for the L2 learners. Three English-speaking participants did not manage to complete the AJT due to time constraints and/or technical glitches with the computer used for the experiment (thus they were excluded from the AJT data analysis). The participants were compensated for their time with cash payment.

7.3 Results

7.3.1 SPRT results

7.3.1.1 SPRT data preparation

The RT data were prepared for analysis following similar procedures to those in the pilot studies though a modification was made to the data transformation method, as outlined below.

Response screening

Participants were first screened based on their comprehension scores, to check whether they met the 70% accuracy cut-off point. All groups performed at ceiling (native controls: M = 31.12/32 (97.3%), SD = 1.03; Korean speakers: M = 31.33/32 (97.9%), SD = 0.98; English speakers: M = 30.81/32 (96.3%), SD = 1.39). All participants scored 84.4 % (27/31) or higher, thus nobody was excluded.

Data trimming

Next, outlying RTs were removed using the following criterion. With the native controls, data points below 100 ms or beyond 3,000 ms were eliminated. As to the L2 learners, below 100 ms or beyond 6,000 ms cut-off was used.[63] Table 7.3 summarised the amount of data affected by group.

Table 7.3 Amount of data affected

Group	< 100 ms	> 3,000/6,000 ms
Native controls	11.87 %	0.09 %
Korean speakers	0.55 %	0.06 %
English speakers	0.04 %	1.10 %

Data transformation

Since the trimmed RT data had a positively skewed distribution as illustrated in Figure 7.1, they were log-transformed (with base 10) to approximate a normal distribution for linear mixed effects model analysis reported below.[64]

63 This 6,000 ms cut-off was motivated by the fact that the average task completion time of the L2 learners was roughly double that of the native controls (section 7.2.4).

64 Note that log-transformation was not applied to the RT data in the pilot SPRTs (Chapter 6), which were also likely to be positively skewed. It was used in the main SPRT, mainly as a result of development in my knowledge of statistical testing.

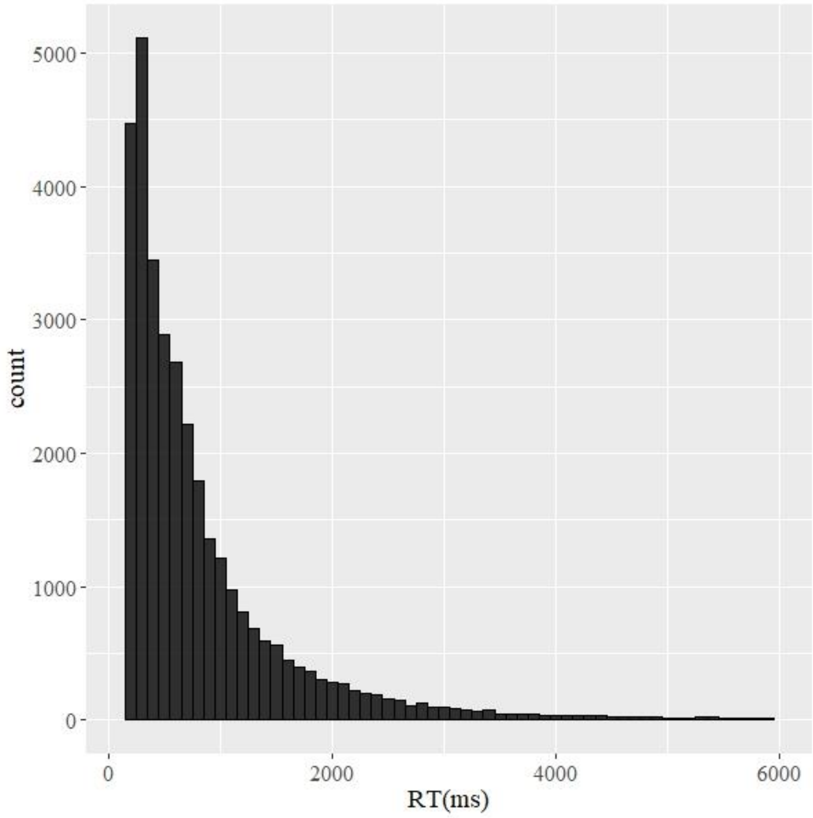


Figure 7.1 Histogram of the trimmed RT data

The log RTs were then residualised in order to minimise the effects of individual variation in reading speed and the different lengths for *sono* items across the conditions (*sono* NP vs. bare NP), using the method described in Chapter 6 (Section 6.5.6.1).

7.3.1.2 *Sono*: SPRT results

Tables 7.4–7.6 present mean residual log-transformed RTs for each region in the *sono* items for the native controls, the Korean-speaking learners, and the English-speaking learners, respectively. Figures 7.2–7.4 provide graphical summaries for each context type (anaphoric, non-anaphoric, and bridging). The rectangle frames indicate the regions to be statistically analysed.

Table 7.4 Sono*: mean log-transformed RRTs: native controls (SD in parentheses)*

Type	#1 NP	#2 Adjunct	#3 NP	#4 Adjunct	#5 Adjunct	#6 VP
Anaphoric						
Sono NP	−0.016 (0.187)	−0.003 (0.233)	−0.033 (0.205)	0.004 (0.172)	0.012 (0.155)	0.099 (0.239)
Bare NP	−0.041 (0.154)	−0.062 (0.158)	−0.048 (0.152)	−0.054 (0.162)	−0.013 (0.156)	0.029 (0.200)
Non-anaphoric						
Sono NP	−0.023 (0.175)	−0.020 (0.218)	−0.040 (0.216)	−0.008 (0.171)	0.021 (0.222)	0.066 (0.234)
Bare NP	−0.010 (0.176)	−0.040 (0.170)	0.035 (0.191)	−0.016 (0.183)	−0.011 (0.155)	0.043 (0.251)
Bridging						
Sono NP	−0.047 (0.149)	−0.038 (0.197)	−0.058 (0.177)	−0.009 (0.138)	−0.035 (0.164)	0.035 (0.168)
Bare NP	−0.040 (0.159)	−0.026 (0.178)	−0.026 (0.167)	−0.002 (0.172)	−0.006 (0.157)	0.041 (0.194)

Table 7.5 Sono*: mean log-transformed RRTs: Korean speakers (SD in parentheses)*

Type	#1 NP	#2 Adjunct	#3 NP	#4 Adjunct	#5 Adjunct	#6 VP
Anaphoric						
Sono NP	−0.112 (0.196)	−0.031 (0.215)	−0.053 (0.213)	0.016 (0.176)	0.003 (0.178)	−0.009 (0.194)
Bare NP	−0.106 (0.219)	−0.035 (0.180)	−0.015 (0.187)	−0.023 (0.136)	0.032 (0.178)	−0.030 (0.185)
Non-anaphoric						
Sono NP	−0.070 (0.203)	0.023 (0.197)	0.037 (0.164)	0.101 (0.187)	0.018 (0.159)	−0.016 (0.210)
Bare NP	−0.115 (0.192)	0.007 (0.228)	0.033 (0.165)	0.051 (0.241)	0.065 (0.198)	−0.010 (0.237)
Bridging						
Sono NP	−0.108 (0.205)	0.016 (0.246)	0.114 (0.198)	0.006 (0.114)	0.022 (0.195)	−0.067 (0.186)
Bare NP	−0.094 (0.220)	0.020 (0.239)	0.222 (0.204)	0.049 (0.171)	0.037 (0.172)	−0.027 (0.214)

Table 7.6 Sono*: mean log-transformed RRTs: English speakers (SD in parentheses)*

Type	#1 NP	#2 Adjunct	#3 NP	#4 Adjunct	#5 Adjunct	#6 VP
Anaphoric						
Sono NP	−0.132 (0.195)	−0.004 (0.190)	−0.081 (0.149)	−0.001 (0.182)	0.005 (0.176)	0.025 (0.244)
Bare NP	−0.139 (0.228)	0.001 (0.201)	−0.064 (0.136)	0.012 (0.203)	0.053 (0.196)	−0.011 (0.217)
Non-anaphoric						
Sono NP	−0.101 (0.255)	0.046 (0.206)	0.068 (0.200)	0.078 (0.178)	0.024 (0.170)	−0.032 (0.219)
Bare NP	−0.082 (0.204)	0.015 (0.212)	0.079 (0.209)	0.034 (0.216)	0.027 (0.176)	−0.029 (0.217)
bridging						
Sono NP	−0.069 (0.253)	0.037 (0.191)	0.178 (0.177)	−0.005 (0.171)	0.063 (0.244)	−0.058 (0.194)
Bare NP	−0.050 (0.260)	0.036 (0.199)	0.242 (0.249)	0.014 (0.192)	0.064 (0.220)	−0.026 (0.200)

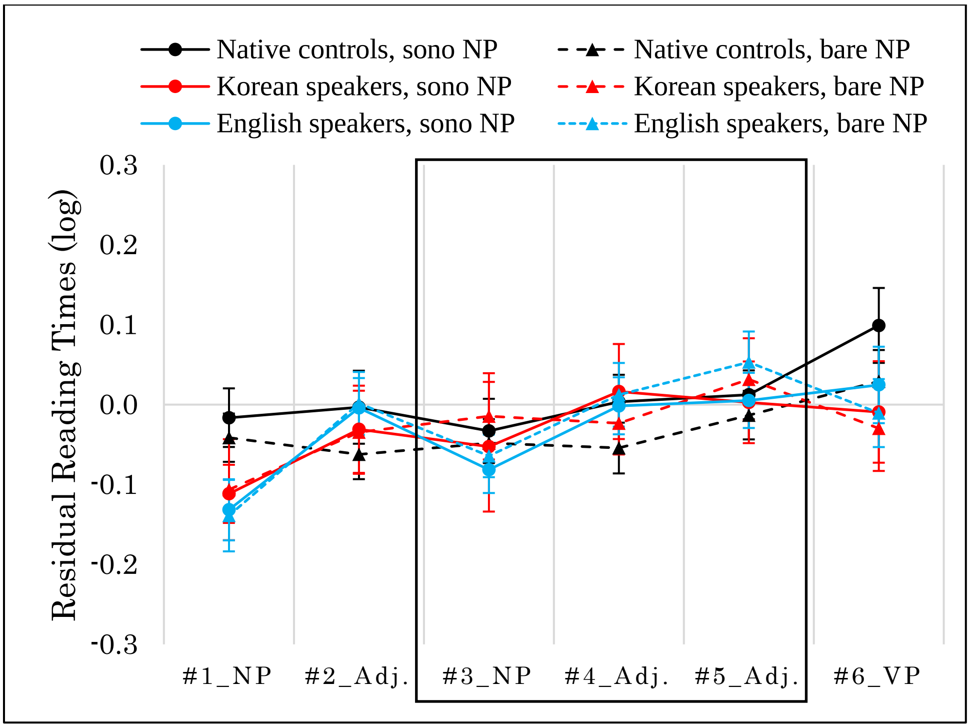


Figure 7.2 Mean RRTs (log) for sono *items in the anaphoric context (error bars = SE)*

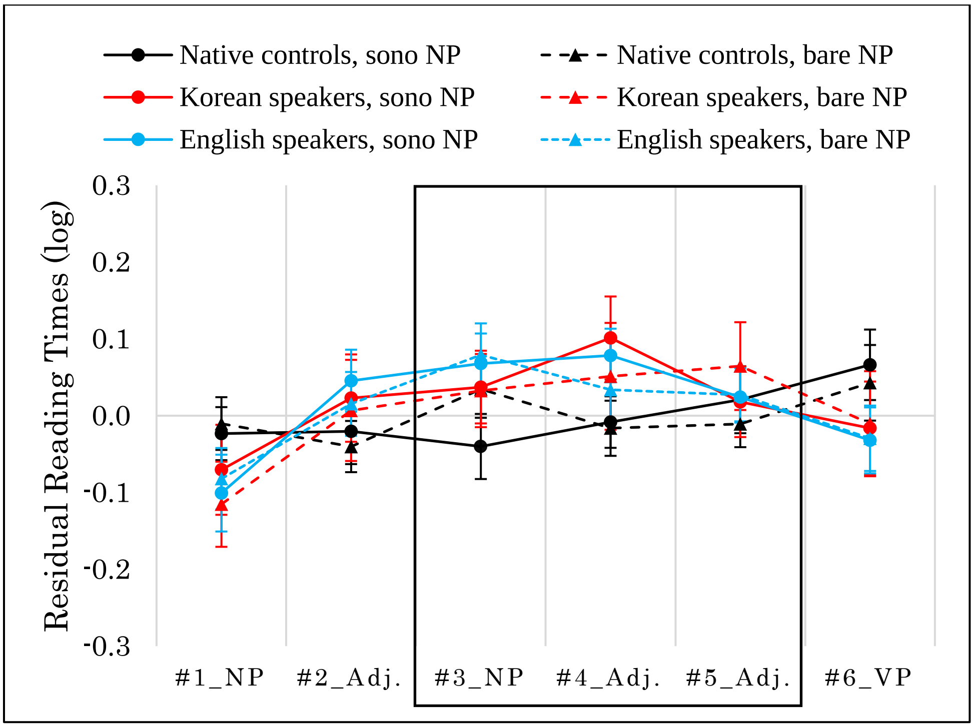


Figure 7.3 Mean RRTs (log) for sono *items in the non-anaphoric context (error bars = SE)*

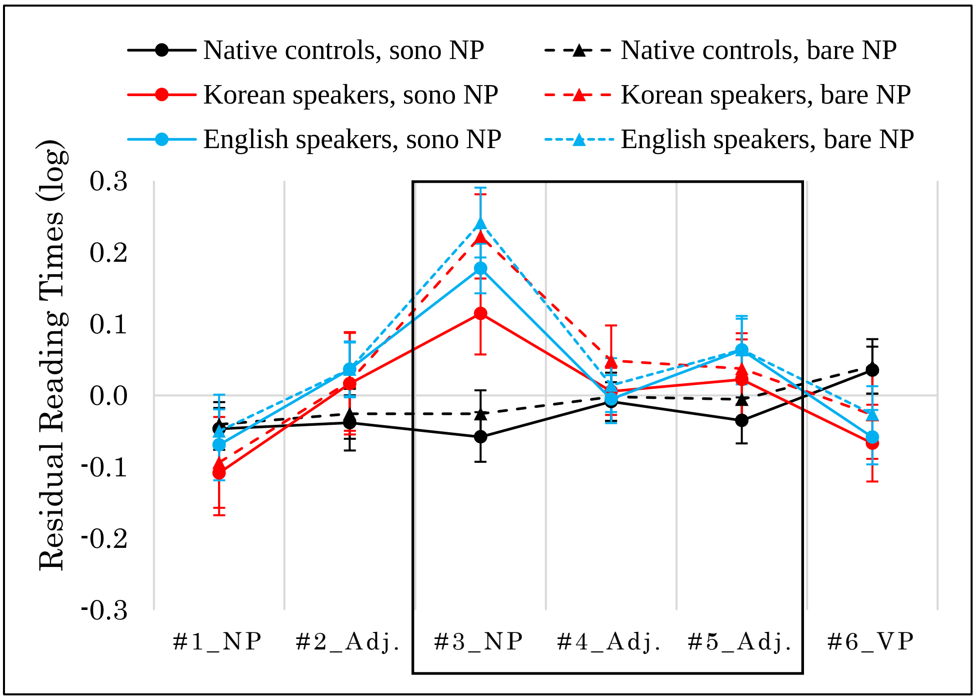


Figure 7.4 Mean RRTs (log) for sono items *in the bridging context (error bars = SE)*

The log-transformed residual RTs were analysed by means of the *lme* function in R. Separate linear mixed-effects models were constructed for each region of interest (Regions 3–5: critical region and two spillover regions). The fixed effects of these models were CONTEXT (*anaphoric* vs. *non-anaphoric* vs. *bridging*), NP TYPE (*sono* NP vs. bare NP), L1 (Japanese vs. Korean vs. English) and the interactions between them with the maximal possible random effects structure (only the final model that converged is reported). CONTEXT and NP TYPE were sum-coded (CONTEXT: anaphoric = [0, 0.5], non-anaphoric = [0.5, 0], bridging = [−0.5, −0.5]; NP TYPE: *sono* NP = 0.5 vs. bare NP = −0.5) whereas L1 was Helmert-coded so that the native Japanese group was compared to the L2 learners (the Korean and English groups combined), and the two L2 groups were compared to each other (L1: Japanese = [1, 0], Korean = [−0.5, 1], English = [−0.5, −1]) (the same Helmert coding scheme is used for the rest of the thesis for L1).[65] When

[65] The sum coding of CONTEXT and the Helmert coding of L1 can be shown in more comprehensible ways using tables in (i) and (ii), respectively.

interactions were found in the main model, they were examined further using relevant nested models.[66]

Table 7.7 shows the results of the omnibus linear mixed effects model analysis in the critical region. There were effects of CONTEXT (context 1 & context 2), NP TYPE, and the contrast between the native control and the L2 groups. There were also two-way interactions between CONTEXT (context 1 & context 2) and the native vs. L2 contrast, between CONTEXT (context 1) and the Korean vs. English contrast. Additionally, there was a marginal two-way interaction between CONTEXT (context 2) and NP TYPE. Particularly important in confirming whether the predicted reading effects occurred is the absence/presence of interactions involving CONTEXT (context 1 & context 2) and NP TYPE. Thus, the absence of relevant interactions means that none of the three groups read the two types of NP significantly differently in any context, despite the marginal interaction between CONTEXT (context 2) and NP TYPE.

(i) Sum coding of CONTEXT

Level	Context 1 (LVL 1 vs. Mean)	Context 2 (LVL 2 vs. Mean)
1. non-anaphoric	0.5	0
2. anaphoric	0	0.5
3. bridging	−0.5	−0.5

(ii) Helmert coding of L1

Level	L1 1:J vs. (K & E) (LVL 1 vs. LVLs 2 & 3)	L1 2: K vs. E (LVL 2 vs. LVL 3)
1. Japanese	1	0
2. Korean	−0.5	1
3. English	−0.5	−1

Note. J = Japanese, K = Korean, E = English.

For CONTEXT, the first contrast (context 1) compared level 1 (non-anaphoric) with the grand mean of all levels, and the second (context 2) compared level 2 (anaphoric) with the grand mean. Level 3 (bridging) was not compared to the other levels. As to L1, the first contrast compared level 1 of the variable (Japanese) with all the subsequent levels (levels 2 & 3: Korean & English) and the second contrast compared level 2 (Korean) with level 3 (English).

66 Given the significant difference between the Korean and English groups in proficiency (i.e., cloze test scores), the effect of proficiency was planned to be examined if three-way interactions between CONTEXT, NP TYPE, and the contrast between the Korean vs. English were found significant in order to tease apart the effect of L1 from that of proficiency. However, such interactions were not found, as shown below.

Table 7.7 Fixed effects estimates for linear mixed effects model of log-transformed RRTs for sono *items in the critical region*

Fixed Effects	β	*SE*	*t*	*p*
(intercept)	0.036	0.012	2.653	.014 *
context 1	−0.163	0.035	−4.728	<.001 ***
context 2	0.160	0.035	4.607	<.001 ***
NP type	−0.036	0.012	−3.088	.005 **
L1(1): Japanese vs. (Korean & English)	−0.062	0.011	−5.609	<.001 ***
L1(2): Korean & English	−0.007	0.007	−1.034	.305
context 1 × NP type	0.045	0.033	1.377	.181
context 2 × NP type	−0.062	0.033	−1.879	.073 †
context 1 × L1(1): Japanese vs. (Korean & English)	0.142	0.031	4.566	<.001 ***
context 2 × L1(1): Japanese vs. (Korean & English)	−0.187	0.031	−5.984	<.001 ***
context 1 × L1(2): Korean & English	0.054	0.020	2.714	.008 **
context 2 × L1(2): Korean & English	−0.030	0.020	−1.472	.146
NP type × L1(1): Japanese vs. (Korean & English)	0.005	0.013	0.340	.734
NP type × L1(2): Korean & English	−0.007	0.013	−0.555	.580
context 1 × NP type × L1(1): Japanese vs. (Korean & English)	0.046	0.038	1.213	.226
context 2 × NP type × L1(1): Japanese vs. (Korean & English)	0.059	0.038	1.545	.124
context 1 × NP type × L1(2): Korean & English	−0.005	0.036	−0.151	.880
context 2 × NP type × L1(2): Korean & English	−0.025	0.037	−0.676	.500

Note. Formula: RRT ~ context * NP type * L1 + (context * NP type | participant) + (NP type + L1 | item). Coding: context: anaphoric = [0, 0.5], non-anaphoric = [0.5, 0], bridging = [−0.5, −0.5]; NP type: *sono* NP = 0.5 vs. bare NP = −0.5; L1: Japanese = [1, 0], Korean = [−0.5, 1], English = [−0.5, −1]. *** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$.

The main model output for the spillover 1 region is presented in Table 7.8. The model revealed that there were effects of the contrast between the native and L2 groups and a two-way interaction between the effects of CONTEXT (context 2) and NP TYPE. Given the absence of three-way interactions, it is suggested that although the effect of NP TYPE differed across the contexts, the three groups were not significantly different from one another in terms of the effect of NP TYPE within each context.

The effect of NP TYPE within each context type was examined further with a nested model. Its results (Table 7.9) revealed that the effect of NP TYPE was not significant within any contexts. The two-way interaction in the main model apparently resulted from the coefficient for the effect of NP TYPE being in the negative for the bridging context but in the positive for the anaphoric and non-anaphoric contexts.

Table 7.10 provides the results of the main model for the spillover 2 region. There was an effect of the contrast between the native control and the L2 groups, as in the preceding two segments. There was a marginal two-way interaction between the effects of CONTEXT (context 2) and the native vs. L2 contrast. Additionally, there was a three-way interaction between the effects of CONTEXT (context 2), NP TYPE and the native vs. L2 contrast, which suggests that the effect of NP TYPE depends on the context and the group (native or L2). To identify the source of the interaction, separate nested models were constructed to examine the effect of NP TYPE within each group for each context. The models (Table 7.11) found no significant effect of NP TYPE within each group for any context despite a marginal effect in the English group for the anaphoric context in the predicted direction (i.e., *sono* NP < bare NP). It seems that the three-way interaction reflects the fact that the polarities of NP TYPE effects differ in the native control vs. the L2 learners (the Korean and English groups combined), but the relation varies according to the context as well (anaphoric & non-anaphoric vs. bridging).[67]

Overall, the series of statistical analyses above found no solid evidence for the predicted reading patterns (SPRT Predictions 1−3 in Section 7.2.3.2) for any groups, namely reading slowdown in the *sono* NP condition in the non-anaphoric context and reading facilitation in the *sono* NP condition in the anaphoric and bridging context.

[67] It can be seen in Table 7.11 that in the bridging model, the Korean group matches the native control group rather than the English group in terms of the direction of the effect of NP type. However, the English group inevitably represents the L2 pattern due to the sample size discrepancy (Korean, $n = 12$ vs. English, $n = 26$).

Table 7.8 Fixed effects estimates for linear mixed effects model of log-transformed RRTs for sono *items in the spillover 1 region*

Fixed Effects	*β*	*SE*	*t*	*p*
(intercept)	0.014	0.013	1.072	.294
context 1	−0.043	0.034	−1.269	.218
context 2	−0.010	0.035	−0.288	.776
NP type	0.012	0.009	1.323	.186
L1(1): Japanese vs. (Korean & English)	−0.029	0.014	−2.145	.041 *
L1(2): Korean & English	0.005	0.009	0.562	.577
context 1 × NP type	0.032	0.025	1.265	.206
context 2 × NP type	−0.070	0.025	−2.765	.006 **
context 1 × L1(1): Japanese vs. (Korean & English)	0.021	0.036	0.595	.558
context 2 × L1(1): Japanese vs. (Korean & English)	0.029	0.037	0.801	.431
context 1 × L1(2): Korean & English	−0.019	0.021	−0.920	.364
context 2 × L1(2): Korean & English	0.012	0.022	0.561	.578
NP type × L1(1): Japanese vs. (Korean & English)	0.006	0.012	0.531	.595
NP type × L1(2): Korean & English	0.007	0.012	0.585	.559
context 1 × NP type × L1(1): Japanese vs. (Korean & English)	0.047	0.034	1.369	.171
context 2 × NP type × L1(1): Japanese vs. (Korean & English)	0.019	0.034	0.570	.569
context 1 × NP type × L1(2): Korean & English	0.039	0.033	1.201	.230
context 2 × NP type × L1(2): Korean & English	−0.037	0.033	−1.124	.261

Note. Formula: RRT ~ context * NP type * L1 + (context + NP type | participant) + (NP type + L1 | item). Coding: context: anaphoric = [0, 0.5], non-anaphoric = [0.5, 0], bridging = [−0.5, −0.5]; NP type: *sono* NP = 0.5 vs. bare NP = −0.5; L1: Japanese = [1, 0], Korean = [−0.5, 1], English = [−0.5, −1]. *** p <.001, * p < .05.

Table 7.9 Fixed effects estimates for nested linear mixed effects model of log-transformed RRTs for sono *items in the spillover 1 region: effects of NP type within each context*

Fixed Effects	*β*	*SE*	*t*	*p*
(intercept)	0.010	0.013	0.772	.447
context 1	−0.038	0.034	−1.128	.272
context 2	−0.010	0.034	−0.295	.771
context: anaphoric / NP type	0.025	0.016	1.521	.132
context: bridging / NP type	−0.020	0.015	−1.290	.198
context : non-anaphoric / NP type	0.029	0.020	1.482	.144

Note. Formula: RRT ~ context / NP type + (context * NP type | participant) + (NP type | item). "*x* / *y*" represents the effect of variable *y* with variable *x* held constant. Coding: context: anaphoric = [0, 0.5], non-anaphoric = [0.5, 0], bridging = [−0.5, −0.5]; NP type: *sono* NP = 0.5 vs. bare NP = −0.5.

Table 7.10 Fixed effects estimates for linear mixed effects model of log-transformed RRTs for sono *items in the spillover 2 region*

Fixed Effects	β	SE	t	p
(intercept)	0.021	0.014	1.531	.139
context 1	−0.013	0.039	−0.331	.744
context 2	0.006	0.039	0.158	.876
NP type	−0.013	0.010	−1.386	.166
L1(1): Japanese vs. (Korean & English)	−0.027	0.007	−3.749	<.001 ***
L1(2): Korean & English	−0.006	0.007	−0.795	.429
context 1 × NP type	−0.006	0.027	−0.234	.815
context 2 × NP type	−0.004	0.027	−0.158	.874
context 1 × L1(1): Japanese vs. (Korean & English)	0.019	0.019	0.999	.319
context 2 × L1(1): Japanese vs. (Korean & English)	−0.036	0.019	−1.923	.056 †
context 1 × L1(2): Korean & English	−0.001	0.018	−0.067	.946
context 2 × L1(2): Korean & English	−0.023	0.018	−1.250	.213
NP type × L1(1): Japanese vs. (Korean & English)	0.020	0.013	1.593	.111
NP type × L1(2): Korean & English	−0.007	0.012	−0.551	.582
context 1 × NP type × L1(1): Japanese vs. (Korean & English)	0.043	0.036	1.173	.241
context 2 × NP type × L1(1): Japanese vs. (Korean & English)	−0.074	0.036	−2.057	.040 *
context 1 × NP type × L1(2): Korean & English	0.032	0.035	0.913	.361
context 2 × NP type × L1(2): Korean & English	>0.001	0.035	0.008	.994

Note. Formula: RRT ~ context * NP type * L1 + (context + NP type | participant) + (NP type | item). Coding: context: anaphoric = [0, 0.5], non-anaphoric = [0.5, 0], bridging = [−0.5, −0.5]; NP type: *sono* NP = 0.5 vs. bare NP = −0.5; L1: Japanese = [1.0, 0], Korean = [−0.5, 1], English = [−0.5, −1]. *** p <.001, * p < .05, † p < .10.

Table 7.11 Results of nested models for RRTs for each context for sono *items in the spillover 2 region*

Fixed Effects	*β*	*SE*	*t*	*p*
anaphoric				
(intercept)	0.016	0.019	0.835	.429
L1(1): J vs. (K & E)	−0.016	0.021	−0.769	.464
L1(2): K vs. E	−0.006	0.016	−0.391	.705
L1: J / NP type	0.026	0.030	0.843	.426
L1: K / NP type	−0.028	0.040	−0.704	.500
L1: E / NP type	−0.047	0.024	−1.922	.070 †
Non-anaphoric				
(intercept)	0.024	0.027	0.905	.395
L1(1): J vs. (K & E)	−0.019	0.015	−1.299	.230
L1(2): K vs. E	0.006	0.013	0.502	.627
L1: J / NP type	0.028	0.029	0.975	.354
L1: K / NP type	−0.049	0.041	−1.205	.259
L1: E / NP type	−0.004	0.024	−0.166	.869
Bridging				
(intercept)	0.024	0.026	0.934	.381
L1(1): J vs. (K & E)	−0.045	0.027	−1.674	.137
L1(2): K vs. E	−0.017	0.019	−0.911	.391
L1: J / NP type	−0.029	0.027	−1.098	.287
L1: K / NP type	−0.014	0.036	−0.384	.701
L1: E / NP type	<0.001	0.027	0.012	.990

Note. Formula: RRT ~ L1 / NP type + (NP type | participant) + (NP type * L1 | item). J = Japanese, K = Korean, E = English. "*x* / *y*" represents the effect of variable *y* with variable *x* held constant. Coding: context: anaphoric = [0, 0.5], non-anaphoric = [0.5, 0], bridging = [−0.5, −0.5]; NP type: *sono* NP = 0.5 vs. bare NP = −0.5; L1: Japanese = [1.0, 0], Korean = [−0.5, 1], English = [−0.5, −1]. † $p < .10$.

7.3.1.3 NQ constructions: SPRT results

Tables 7.12–7.14 provide mean residual RTs for each region in the NQ construction items for the native controls, the Korean speakers, and the English-speakers, respectively. Figures 7.5–7.6 summarise the data for each context type (i.e., [+definite] and [−definite]) in graphical forms. The rectangle frames indicate the regions subject to statistical analysis.

Table 7.12 NQ constructions: mean log-transformed RRTs: native controls (SD in parentheses)

Type	#1 NP	#2 Adjunct	#3 NP	#4 Adjunct	#5 Adjunct	#6 VP
[+definite]						
Post-nominal	−0.027 (0.167)	−0.051 (0.171)	−0.001 (0.199)	0.023 (0.156)	0.008 (0.145)	0.043 (0.218)
Floating	−0.027 (0.152)	−0.046 (0.213)	−0.045 (0.207)	0.029 (0.186)	−0.006 (0.153)	0.017 (0.188)
[−definite]						
Post-nominal	−0.046 (0.187)	−0.065 (0.195)	−0.042 (0.221)	0.010 (0.196)	−0.033 (0.155)	0.029 (0.217)
Floating	−0.069 (0.164)	−0.056 (0.156)	−0.056 (0.234)	0.045 (0.179)	−0.023 (0.156)	0.012 (0.226)

Table 7.13 NQ constructions: mean log-transformed RRTs: Korean speakers (SD in parentheses)

Type	#1 NP	#2 Adjunct	#3 NP	#4 Adjunct	#5 Adjunct	#6 VP
[+definite]						
Post- nominal	−0.116 (0.159)	−0.061 (0.240)	−0.022 (0.238)	0.046 (0.168)	0.009 (0.154)	0.027 (0.231)
Floating	−0.138 (0.152)	−0.017 (0.182)	−0.041 (0.238)	0.061 (0.187)	0.032 (0.157)	0.048 (0.230)
[−definite]						
Post- nominal	−0.127 (0.126)	−0.033 (0.189)	0.100 (0.281)	0.055 (0.206)	0.074 (0.177)	0.005 (0.220)
Floating	−0.122 (0.168)	−0.041 (0.187)	0.089 (0.260)	0.052 (0.172)	0.049 (0.205)	−0.106 (0.169)

Table 7.14 NQ constructions: mean log-transformed RRTs: English speakers (SD in parentheses)

Type	#1 NP	#2 Adjunct	#3 NP	#4 Adjunct	#5 Adjunct	#6 VP
[+definite]						
Post- nominal	−0.163 (0.169)	0.032 (0.216)	0.054 (0.213)	0.003 (0.180)	0.001 (0.171)	−0.025 (0.223)
Floating	−0.163 (0.186)	0.019 (0.188)	0.017 (0.224)	0.083 (0.208)	0.022 (0.196)	−0.009 (0.205)
[−definite]						
Post- nominal	−0.162 (0.192)	0.085 (0.238)	0.127 (0.251)	0.080 (0.185)	0.030 (0.160)	−0.078 (0.201)
Floating	−0.170 (0.169)	0.069 (0.214)	0.127 (0.201)	0.101 (0.265)	0.086 (0.225)	−0.033 (0.163)

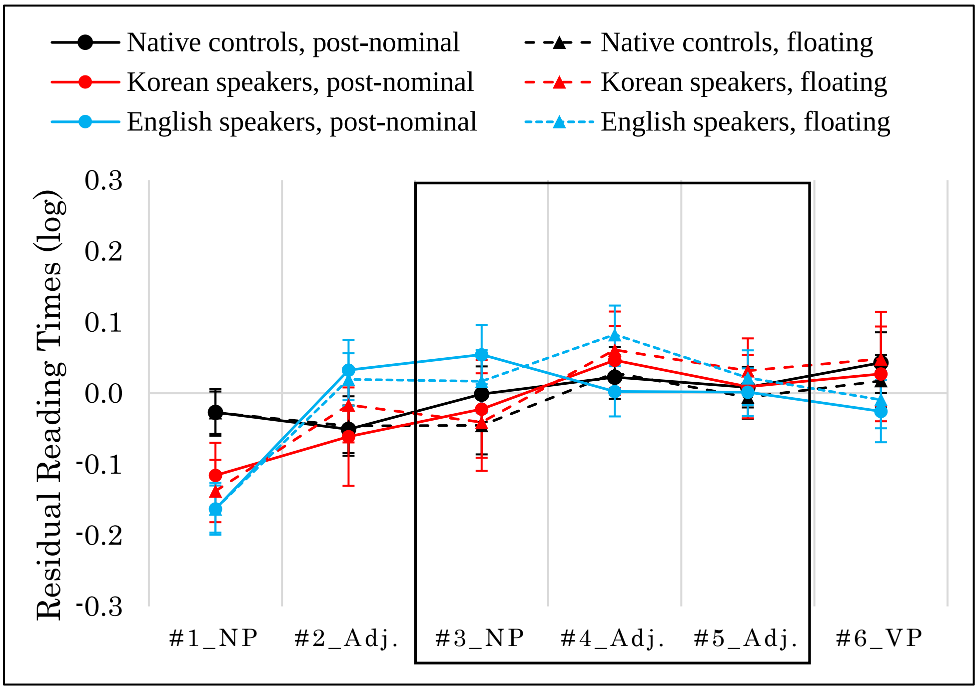


Figure 7.5 Mean RRTs (log) for NQ construction items in the [+definite] context (error bars = SE)

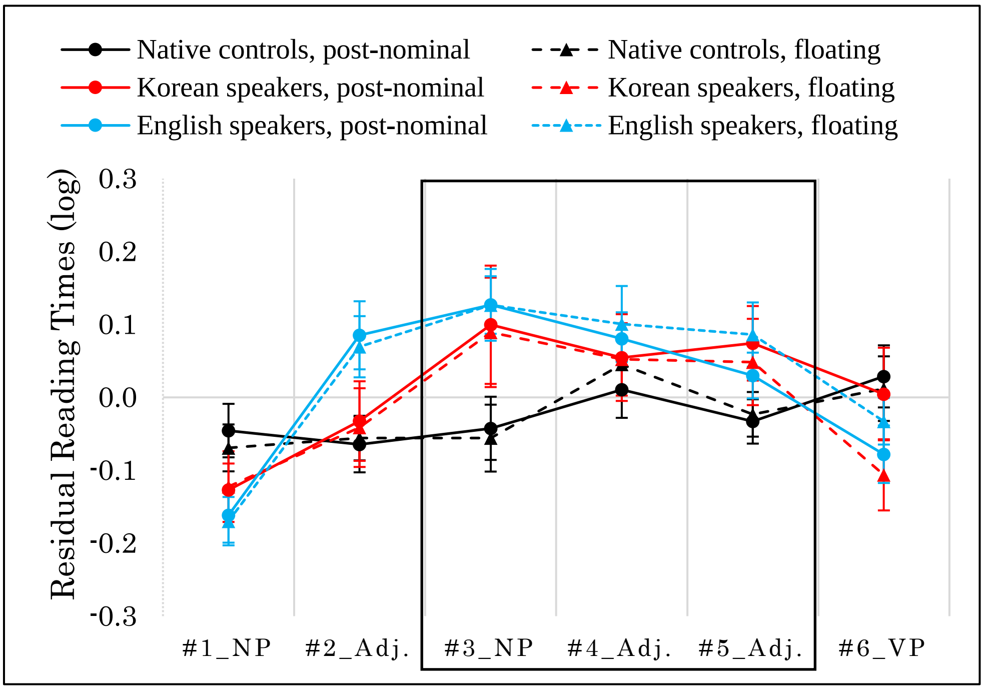


Figure 7.6 Mean RRTs (log) for NQ construction items in the [−definite] context (error bars = SE)

Analysis was conducted by means of a linear mixed-effects model, which including as fixed effects DEFINITENESS (*[+definite]* vs. *[−definite]*), QUANTIFIER TYPE (*post-nominal* vs. *floating*), L1(*Japanese* vs. *Korean* vs. *English*), and the interactions between them. DEFINITENESS and QUANTIFIER TYPE were sum-coded (DEFINITENESS: [+definite] = 0.5 vs. [−definite] = −0.5; QUANTIFIER TYPE: post-nominal = 0.5 vs. floating = −0.5) whereas L1 was Helmert-coded to compare between the native vs. the L2 learners and between the Korean speakers vs. the English speakers. The random effects structure was specified as maximal as possible (only the final models that converged are reported).

Table 7.15 presents the results for the linear mixed-effects analysis in the critical segment. There were effects of the contrast between the native controls and the L2 learners and a marginal effect of the contrast between the Korean and English group as well as a two-way interaction between DEFINITENESS and the native vs. L2 contrast. However, the fact that there were no interactions involving the effects of DEFINITENESS and QUANTIFIER TYPE means that no groups showed evidence of sensitivity to the definiteness constraint on floating NQs in terms of RTs.

Table 7.16 provides the linear mixed-effects model output for the fourth segment (spillover 1). There was an effect of the contrast between the native controls and the L2 learners. However, no interactions involving the effects of DEFINITENESS and QUANTIFIER TYPE were found, meaning no evidence for the relevant definiteness constraint across the groups.

The results for the spillover 2 model are summarised in Table 7.17. The model revealed an effect of the contrast between the native controls and the L2 learners. There was an interaction between DEFINITENESS and the native vs. L2 contrast, as in the critical segment. The absence of two- and three-way interactions concerning DEFINITENESS and QUANTIFIER TYPE suggests that no groups were sensitive to the definiteness constraint, as in the preceding two segments.

In sum, because there were no two-way interactions between DEFINITENESS and QUANTIFIER TYPE and no three-way interactions in any segment of interest, no group seems sensitive to the definiteness constraint on floating NQs in terms of RTs, which suggests no evidence of the predicted reading slowdown effect (SPRT Prediction 4 in Section 7.2.3.2).

Table 7.15 Fixed effects estimates for linear mixed effects model of log-transformed RRTs for NQ construction *items in the critical region*

Fixed Effects	*β*	*SE*	*t*	*p*
(intercept)	0.025	0.026	0.954	.355
definiteness	0.065	0.051	1.271	.224
quantifier type	0.020	0.014	1.389	.183
L1(1): Japanese vs. (Korean & English)	−0.064	0.022	−2.950	.009 **
L1(2): Korean vs. English	−0.026	0.014	−1.839	.080 †
definiteness ×quantifier type	−0.024	0.029	−0.842	.412
definiteness × L1(1): Japanese vs. (Korean & English)	−0.092	0.040	−2.273	.039 *
definiteness × L1(2): Korean vs. English	0.017	0.025	0.669	.514
quantifier type × L1(1): Japanese vs. (Korean & English)	0.005	0.018	0.320	.749
quantifier type × L1(2): Korean vs. English	−0.005	0.018	−0.274	.784
definiteness ×quantifier type × L1(1): Japanese vs. (Korean & English)	−0.012	0.037	−0.337	.736
definiteness ×quantifier type× L1(2): Korean vs. English	0.014	0.035	0.403	.687

Note. Formula: RRT ~ definiteness * quantifier type * L1 + (definiteness * quantifier type | participant) + (quantifier type + L1 | item). Coding: definiteness: [+definite] = 0.5 vs [−definite] = −0.5; quantifier type: post-nominal = 0.5 vs. floating = −0.5; L1: Japanese = [1.0, 0], Korean = [−0.5, 1], English = [−0.5, −1]. ** $p < .01$, * $p < .05$, † $p < .10$.

Table 7.16 Fixed effects estimates for linear mixed effects model of log-transformed RRTs for NQ construction items in the spillover 1 region

Fixed Effects	β	*SE*	*t*	*p*
(intercept)	0.049	0.016	2.994	.009 **
definiteness	0.017	0.033	0.525	.608
quantifier type	−0.026	0.016	−1.633	.123
L1(1): Japanese vs. (Korean & English)	−0.023	0.008	−2.675	.008 **
L1(2): Korean vs. English	−0.007	0.008	−0.826	.409
definiteness ×quantifier type	0.014	0.031	0.450	.659
definiteness × L1(1): Japanese vs. (Korean & English)	−0.014	0.017	−0.817	.414
definiteness × L1(2): Korean vs. English	−0.025	0.016	−1.527	.127
quantifier type × L1(1): Japanese vs. (Korean & English)	0.005	0.017	0.290	.773
quantifier type × L1(2): Korean vs. English	0.022	0.017	1.325	.191
definiteness ×quantifier type × L1(1): Japanese vs. (Korean & English)	−0.047	0.034	−1.380	.168
definiteness ×quantifier type× L1(2): Korean vs. English	−0.022	0.032	−0.666	.506

Note. Formula: RRT ~ definiteness * quantifier type * L1 + (definiteness * quantifier type | participant) + (quantifier type | item). Coding: definiteness: [+definite] = 0.5 vs. [−definite] = −0.5; quantifier type: post-nominal = 0.5 vs. floating = −0.5; L1: Japanese = [1.0, 0], Korean = [−0.5, 1], English = [−0.5, −1]. ** $p < .01$.

Table 7.17 Fixed effects estimates for linear mixed effects model of log-transformed RRTs for NQ construction items in the spillover 2 region

Fixed Effects	*B*	*SE*	*t*	*p*
(intercept)	0.021	0.010	2.065	.051 †
definiteness	0.020	0.019	1.055	.308
quantifier type	−0.011	0.011	−1.005	.315
L1(1): Japanese vs. (Korean & English)	−0.033	0.010	−3.436	.001 **
L1(2): Korean vs. English	0.003	0.009	0.335	.739
definiteness × quantifier type	−0.003	0.023	−0.148	.883
definiteness × L1(1): Japanese vs. (Korean & English)	−0.048	0.015	−3.170	.002 **
definiteness × L1(2): Korean vs. English	−0.003	0.015	−0.183	.855
quantifier type × L1(1): Japanese vs. (Korean & English)	0.014	0.015	0.900	.368
quantifier type × L1(2): Korean vs. English	0.020	0.015	1.371	.171
definiteness × quantifier type × L1(1): Japanese vs. (Korean & English)	−0.019	0.030	−0.632	.527
definiteness × quantifier type× L1(2): Korean vs. English	0.042	0.029	1.436	.151

Note. Formula: RRT ~ definiteness * quantifier type * L1 + (definiteness * quantifier type | participant) + (quantifier type | item). Coding: definiteness: [+definite] = 0.5 vs. [−definite] = −0.5; quantifier type: post-nominal = 0.5 vs. floating = −0.5; L1: Japanese = [1.0, 0], Korean = [−0.5, 1], English = [−0.5, −1]. ** $p < .01$, † $p < .10$.

7.3.2 AJT results

7.3.2.1 *Sono*: AJT results

Group mean acceptability ratings were calculated for each NP type for each context type, as summarised in Table 7.18. Figures 7.7 through 7.9 illustrate the data in bar graphs for the anaphoric, non-anaphoric and bridging contexts, respectively. *I don't know* responses were excluded, which affected 2.80 % of the data (41/1464: 1 from the native control group and 40 from the English group) with roughly a half of them found in the bridging context and the other half evenly spread across the anaphoric and non-anaphoric contexts.

Table 7.18 Mean acceptability ratings by context and by group : Main study 1 sono *items (SD in parentheses)*

Condition	NP type	Native controls	Korean speakers	English speakers
Anaphoric	*Sono* NP	5.30 (1.20)	4.50 (1.49)	4.47 (1.66)
	Bare NP	4.58 (1.86)	4.29 (1.24)	4.41 (1.82)
Non-anaphoric	*Sono* NP	4.06 (2.04)	3.69 (1.82)	3.91 (1.89)
	Bare NP	4.65 (1.82)	4.46 (1.50)	4.57 (1.78)
Bridging	*Sono* NP	5.17 (1.32)	4.46 (1.49)	3.83 (1.99)
	Bare NP	4.70 (1.81)	4.15 (1.53)	3.27 (2.02)

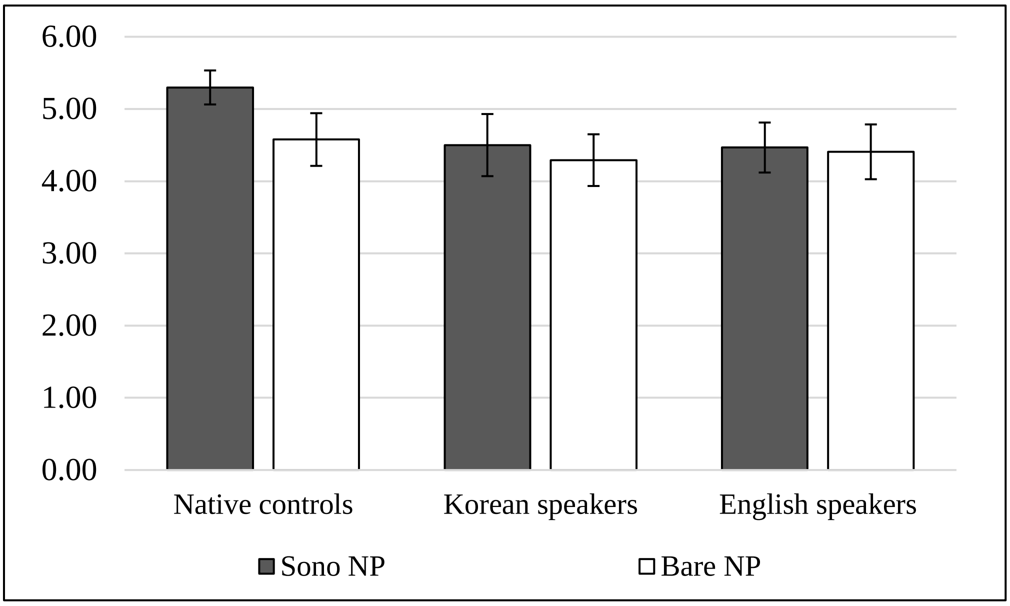


Figure 7.7 Mean acceptability ratings for sono *items in the anaphoric context (error bars = SE)*

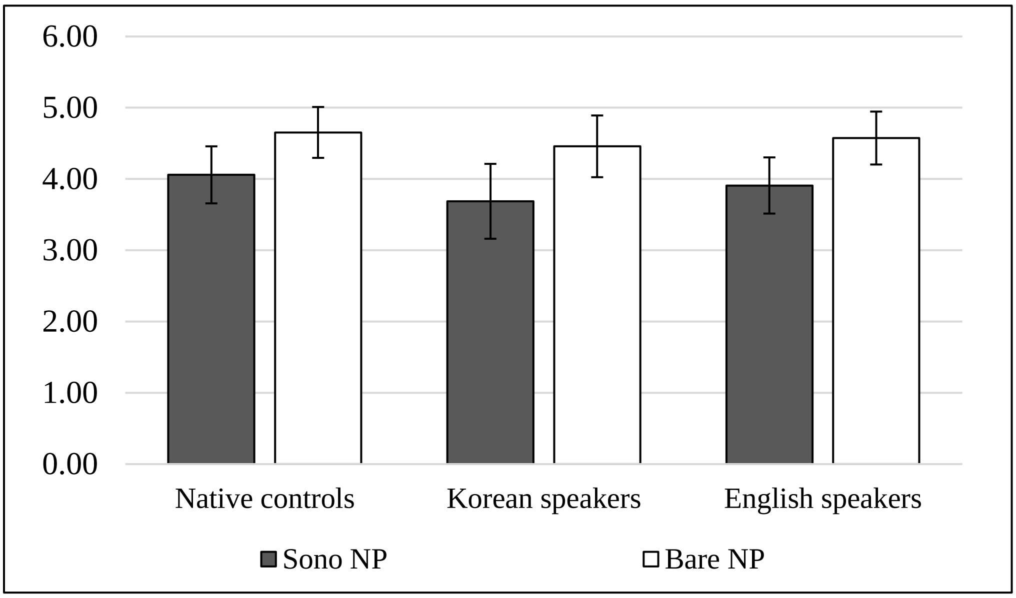


Figure 7.8 Mean acceptability ratings for sono *items in the non-anaphoric context (error bars = SE)*

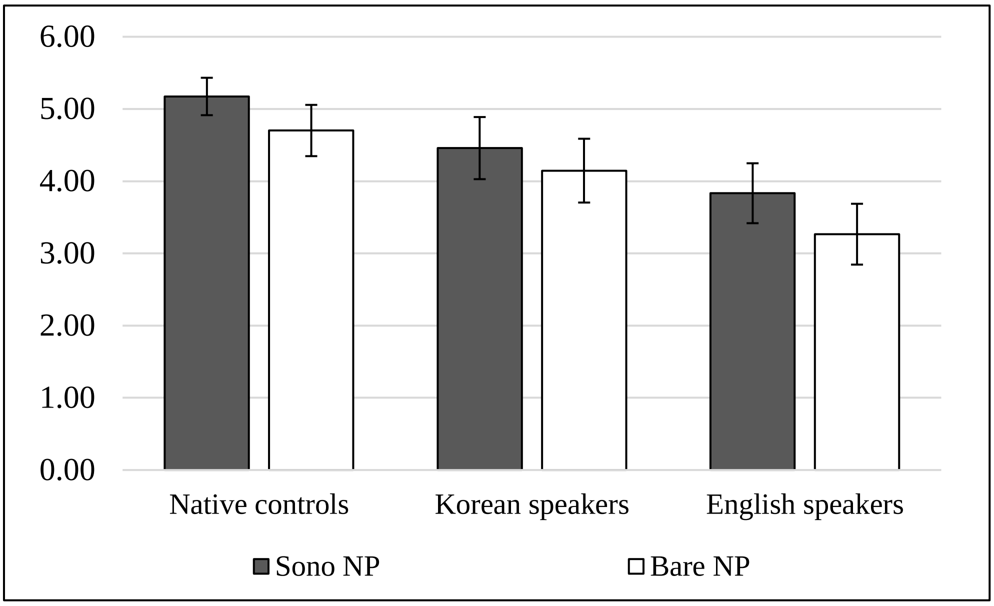


Figure 7.9 Mean acceptability ratings for sono *items in the bridging context (error bars = SE)*

Recall the predicted native judgement pattern based on the pilot studies (AJT predictions 1–3 in Section 7.2.3.3): s*ono* NPs would be rated significantly lower than bare NPs in the non-anaphoric context; whereas *sono* NPs would be rated higher than or equally high as bare NPs in the anaphoric context, and would be equally high as bare NPs in the bridging context. Figures 7.7 and 7.8 show that the native controls' responses were overall compatible with the predictions for the anaphoric and non-anaphoric contexts (Predictions 1 and 2), although the preference for *sono* NPs over bare NPs in the anaphoric context was notably clearer in the native control groups than the L2 groups. As to the bridging context, all the three groups showed a slight preference for *sono* NPs over bare NPs, against the prediction (AJT prediction 3).

For further analysis, an ordinal mixed effects model was fitted to the acceptability ratings using the *clmm* function in the *ordinal* package (Christensen, 2018) in R.[68] The fixed effects were CONTEXT (*anaphoric* vs. *non-anaphoric* vs. *bridging*), NP

[68] For the AJT data reported in this chapter and next, ordinal mixed-effects models are used rather than linear mixed-effects models such as those used in the pilot data analysis in the previous chapter. This is simply because it is the most reasonable for Likert ratings with natural ordering between the

TYPE (sono *NP* vs. *bare NP*), and L1 (Japanese vs. Korean vs. English), and their interactions. Random intercepts were included for participants and items, with CONTEXT and NP TYPE as random by-participant slopes and NP TYPE as random by-item slopes (i.e., the maximal possible model that converged). L1 was Helmert-coded whereas the other variables were sum-coded in the same way as in the SPRT data analysis above.

The omnibus model (Table 7.19) revealed an effect of the contrast between the native control and L2 groups and two-way interactions between CONTEXT (context 1 & context 2) and NP TYPE, and between CONTEXT (context 1) and the native vs. L2 contrast. There was also a marginal effect of CONTEXT (context 2). Of particular interest are the two-way interactions of NP TYPE with CONTEXT (context 1 & context 2), which suggest that the effect of NP TYPE differs according to the context. Furthermore, the fact that there were no three-way interactions means that the effect of NP TYPE within each context does not significantly differ between the native controls and the L2 groups or between the L2 groups.

values (= ordinal data). Furthermore, linear mixed-effects models rely on the assumption that the dependent variable is interval (or ratio) data, which is not met in Likert rating data hence there is always risk for systematic Type I and Type II errors (Liddell & Kruschke, 2018), and even z-transformed ratings are not completely free of such errors (Dillion & Wagers, forthcoming). I am grateful to Shayne Sloggett (experimental officer in Psycholinguistics at the University of York) for his advice on this issue.

Table 7.19 Results of the omnibus ordinal model for acceptability ratings for sono *items*

Fixed Effects	*β*	*SE*	*z*	*p*
Context 1	−0.327	0.336	−0.973	.331
Context 2	0.535	0.308	1.737	.082 †
NP type	−0.053	0.116	−0.459	.646
L1(1): Japanese vs. (Korean & English)	0.646	0.161	4.006	<.001 ***
L1(2): Korean vs. English	−0.018	0.160	−0.112	.911
Context 1 × NP type	1.675	0.329	5.085	<.001 ***
Context 2 × NP type	−0.731	0.328	−2.227	.026 *
Context 1 × L1(1): Japanese vs. (Korean & English)	−0.753	0.265	−2.842	.004 **
Context 2 × L1(1): Japanese vs. (Korean & English)	0.027	0.199	0.135	.892
Context 1 × L1(2): Korean vs. English	−0.390	0.259	−1.506	.132
Context 2 × L1(2): Korean vs. English	−0.268	0.186	−1.444	.149
NP type × L1(1): Japanese vs. (Korean & English)	−0.198	0.138	−1.434	.152
NP type × L1(2): Korean vs. English	−0.052	0.131	−0.401	.689
Context 1 × NP type × L1(1): Japanese vs. (Korean & English)	0.322	0.387	0.831	.406
Context 2 × NP type × L1(1): Japanese vs. (Korean & English)	−0.549	0.393	−1.398	.162
Context 1 × NP type × L1(2): Korean vs. English	0.027	0.370	0.074	.941
Context 2 × NP type × L1(2): Korean vs. English	−0.297	0.368	−0.805	.421

Note. Formula: rating ~ context * L1* NP type + (context + NP type | participant) + (NP type | item). Coding: context: anaphoric = [0, 0.5], non-anaphoric = [0.5, 0], bridging = [−0.5, −0.5]; NP type: *sono* NP = −0.5 vs. bare NP = 0.5; L1: Japanese = [1.0, 0], Korean = [−0.5, 1], English = [−0.5, −1]. *** p <.001, ** p <.01, * p < .05, † p < .10.

To analyse this further, a nested model was constructed to test the effect of NP TYPE within each context. The model (Table 7.20) found that the effect of NP TYPE was significant in all three contexts: *sono* NPs were rated higher than bare NPs in the anaphoric and bridging contexts whereas the opposite was true in the non-anaphoric context. Taken together with the absence of the three-way interactions between CONTEXT, NP TYPE, and L1-related contrasts in the omnibus model, it can be said that both the native control and L2 groups followed the same patterns within each context. Such a result seems surprising with the anaphoric context, where the relevant contrast appeared notably less clear in the descriptive data for the L2 groups than the native control group (Figure 7.7). To examine this absence of native vs. L2 contrast regarding the effect of NP TYPE, further analysis was conducted by running a post hoc nested model and checking individual response patterns focusing on the anaphoric context. The nested model output, provided in Table 7.21, suggests that the effect of NP TYPE was only significant within the native control group. This could indicate that the native controls tend to be more consistent in rating *sono* NPs higher than bare NPs (although this model result does not provide concrete evidence, since the interaction between NP TYPE and L1 was not significant in the omnibus model).

Table 7.20 Results of the nested ordinal model for acceptability ratings : the effect of NP type by context

Fixed Effects	*β*	*SE*	*z*	*p*
Context 1	−0.369	0.346	−1.065	.287
Context 2	0.553	0.311	1.780	.075 †
Context: anaphoric / NP type	−0.439	0.202	−2.169	.030 *
Context: non-anaphoric / NP type	0.794	0.214	3.709	<.001 ***
Context: bridging / NP type	−0.556	0.210	−2.653	.008 **

Note. Formula: rating ~ context / NP type + (context + NP type | participant) + (NP type | item). Coding: context: anaphoric = [0, 0.5], non-anaphoric = [0.5, 0], bridging = [−0.5, −0.5]; NP type: *sono* NP = −0.5 vs. bare NP = 0.5. *** p <.001, ** p <.01, † p < .10.

Table 7.21 Results of the nested ordinal model for acceptability ratings within each group for the anaphoric context

Fixed Effects	*β*	*SE*	*z*	*p*
L1(1): Japanese vs. (Korean & English)	0.644	0.185	3.483	<.001 ***
L1(2): Korean vs. English	−0.156	0.179	−0.871	.384
L1: Japanese / NP type	−0.889	0.315	−2.822	.005 **
L1: Korean / NP type	−0.382	0.375	−1.017	.309
L1: English / NP type	0.036	0.317	0.114	.909

Note. Formula: rating ~ L1 / NP type + (NP type | participant) + (NP type | item). Coding: NP type: *sono* NP = −0.5 vs. bare NP = 0.5; L1: Japanese = [1.0, 0], Korean = [−0.5, 1], English = [−0.5, −1]. *** p <.001, ** p <.01.

Turning to the individual response analysis, Table 7.22 categorises the individual participants in terms of three types of mean ratings pattern: (i) those with higher ratings for *sono* NPs than bare NPs, (ii) no difference between the meaning ratings of the two NP types, and (iii) those with lower ratings for *sono* NPs than bare NPs. It can be seen that that although the L2 groups had proportionally greater numbers of non-target responses (i.e., *sono* NP < bare NP) than the native controls, the majority of individuals in both L2 groups showed the target pattern (i.e., *sono* NP > bare NP). Thus, it can be said that despite the unclear distinction of the NP types in the group mean ratings (Figure 7.7) and less consistent responses by the L2 groups compared to the native control group, the target pattern was dominant at the individual level among the L2 speakers. This partially accounts for the absence of three-way interaction between CONTEXT, NP TYPE and the native vs. L2 contrast in the main model.

Table 7. 22 Distribution of response patterns within each group, in the anaphoric context

L1	*Sono* NP mean rating > bare NP mean rating		*Sono* NP = bare NP	*Sono* NP mean rating < bare NP mean rating	
	n	Size of difference	*n*	*n*	Size of difference
Japanese	19	0.25–2.25	2	5	0.25–1.25
Korean	7	0.25–1.50	1	4	0.25–0.50
English	12	0.25–1.75	2	9	0.25–2.00

Note. *n* = number of participants with the given response pattern. Size of difference is given in points on the rating scale.

In sum, the results show that all groups exhibit the predicted target responses by rating *sono* NPs lower than bare NPs in the non-anaphoric context and doing the opposite in the anaphoric context. However, they rated *sono* NPs higher than bare NPs in the bridging context, against the predicted comparable acceptability of the two NP types.

7.3.2.2 NQ constructions: AJT results

Group mean acceptability ratings are summarised by each quantifier type and context in Table 7.23. The bar graphs in Figures 7.10 and 7.11 illustrate the data for the [+definite] and [−definite] contexts, respectively. *I don't know* responses were removed. They accounted for 1.74 % of the data, distributed more or less equally across the contexts (17/976: 1 from the native control group and 16 from the English group).

Table 7.23 Mean acceptability ratings by context and by group: Main study 1 NQ construction items (SD in parentheses)

Condition	Quantifier type	Native controls	Korean speakers	English Speakers
[+definite]	Post-nominal	4.93 (1.52)	4.44 (1.35)	3.97 (2.01)
	Floating	4.71 (1.58)	3.19 (1.92)	3.62 (2.10)
[−definite]	Post-nominal	5.21 (1.23)	4.31 (1.50)	4.26 (1.73)
	Floating	4.91 (1.50)	3.96 (1.77)	3.88 (1.89)

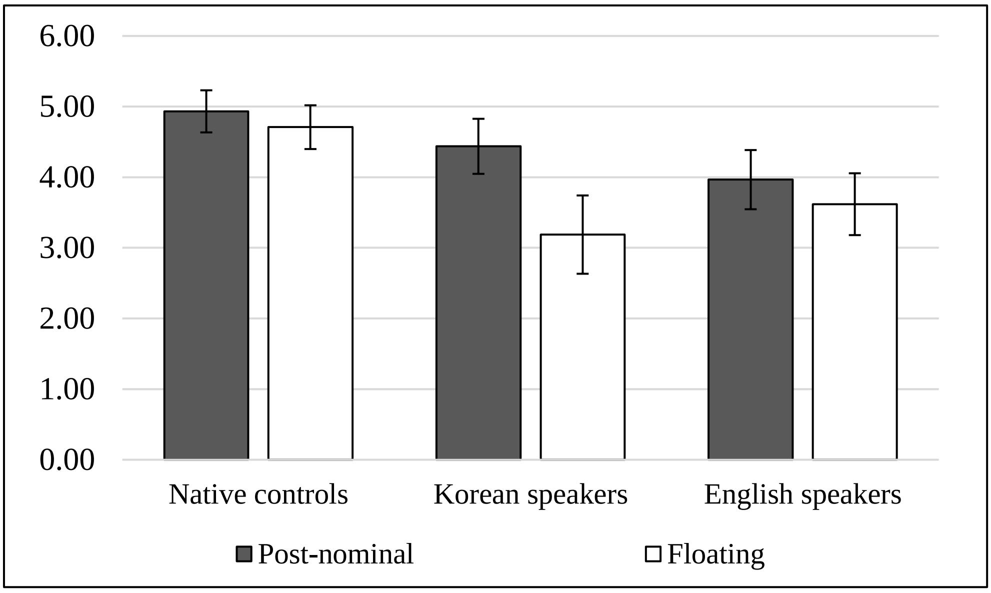


Figure 7.10 Mean acceptability ratings for NQ construction items in the [+definite] context (error bars = SE)

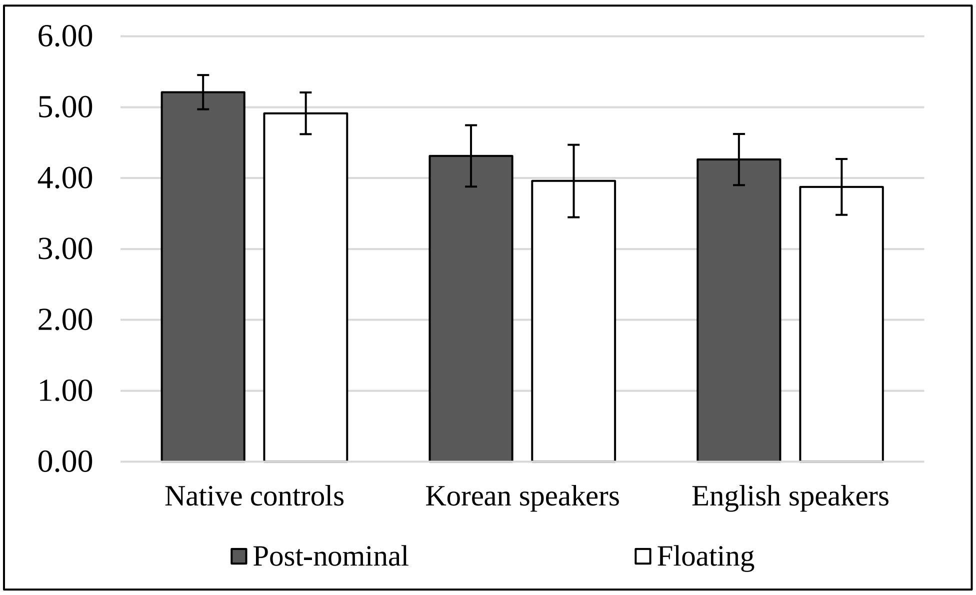


Figure 7.11 Mean acceptability ratings for NQ construction items in the [−definite] context (error bars = SE)

To recap, native controls were expected to yield lower ratings for floating NQs than post-nominal NQs in the [+definite] context but rate both NQs equally acceptable in the [−definite] context. Figures 7.10 and 7.11 suggest that only the Korean group responded in this manner, whereas the native control and English groups appear to have accepted both types of NQ more or less equally across the contexts.

An ordinal mixed-effects model was fitted to the acceptability ratings for further analysis. The fixed effects were DEFINITENESS (*[+definite]* vs. *[−definite]*), QUANTIFIER TYPE (*post-nominal* vs. *floating*), and L1 (*Japanese* vs. *Korean* vs. *English*), and their interactions with the maximal random effects structure (i.e., random intercepts for participants and items along with DEFINITENESS and QUANTIFIER TYPE and their interactions as random by-participant slopes; and QUANTIFIER TYPE, L1 and their interactions as random by-item slopes). L1 was Helmert-coded whereas DEFINITENESS and QUANTIFIER TYPE were sum-coded, in the same way as in the SPRT data analysis. The omnibus model output is provided in Table 7.24. The model shows that there were effects of QUANTIFIER TYPE and the contrast between the native control and L2 groups.

There was no two-way interaction between DEFINITENESS and QUANTIFIER TYPE, nor any three-way interactions, suggesting that no group was sensitive to the definiteness constraint (specifically, the model revealed that all the groups rated post-nominal NQs higher than floating NQs across the contexts).

The absence of three-way interaction between DEFINITENESS, QUANTIFIER TYPE and the Korean vs. English contrast seems surprising given the notably different sensitivity to the definiteness effect in the [+definite] context. To explore this further, the effects of QUANTIFIER TYPE were examined within each group, using nested ordinal models focusing on each context. The model outputs are provided in Table 7.25. The [+definite] model revealed that the effect of quantifier type was significant for the Korean group but not for the Japanese and English groups. The [−definite] model, on the other hand, found no effect of quantifier type for any group. These results suggest that the Korean group was indeed more consistent than the other groups in judging floating NQs as less acceptable than post-nominal NQs and rating both NQ types equally acceptable, though only tentatively given the absence of the relevant three-way interaction in the main model. The absence of two-way interaction seems at least partially attributable to the low statistical power due to the considerably small sample size of the Korean group.

However, the potential L1-related difference between the Korean and English groups could be the effect of proficiency rather than the L1 given their significantly different Japanese proficiency levels. An L2-groups-only version of the omnibus ordinal model with PROFICIENCY (centralised cloze test scores) added as a covariate (L1 was sum-coded with *Korean* as −0.5 and *English* as 0.5) found suggestive evidence that this might be the case: there was a marginal three-way interaction between DEFINITENESS, QUANTIFIER TYPE, and PROFICIENCY but no three-way interaction involving L1 in place of PROFICIENCY (see Table 7.26 for the full output). The model output indicates the tendency that as the proficiency increases, the effect of QUANTIFIER TYPE in the

[+definite] context becomes larger in the target direction (post-nominal > floating).

Overall, the results above suggest that there was no robust evidence of sensitivity to the definiteness constraint for any group. Although the Korean-speaking learners showed a numerically more target-like pattern than the English-speaking learners, this contrast seems attributable more to the difference in proficiency than the L1.

Table 7.24 Results of the omnibus ordinal model for acceptability ratings for NQ construction items

Fixed Effects	*β*	*SE*	*z*	*p*
definiteness	−0.269	0.286	−0.938	.348
quantifier type	0.530	0.143	3.704	<.001 ***
L1(1): Japanese vs. (Korean & English)	0.928	0.201	4.616	<.001 ***
L1(2): Korean vs. English	−0.020	0.201	−0.101	.919
definiteness ×quantifier type	0.234	0.263	0.891	.373
definiteness × L1(1): Japanese vs. (Korean & English)	−0.013	0.231	−0.056	.955
definiteness × L1(2): Korean vs. English	−0.015	0.238	−0.065	.948
quantifier type × L1(1): Japanese vs. (Korean & English)	−0.132	0.185	−0.715	.475
quantifier type × L1(2): Korean vs. English	0.177	0.182	0.974	.330
definiteness ×quantifier type × L1(1): Japanese vs. (Korean & English)	−0.325	0.340	−0.956	.339
definiteness ×quantifier type× L1(2): Korean vs. English	0.485	0.335	1.447	.148

Note. Formula: rating ~ definiteness * quantifier type * L1 + (definiteness * quantifier type | participant) + (quantifier type * L1 | item). Coding: definiteness: [+definite] = 0.5 vs [−definite] = −0.5 ; quantifier type: post-nominal = 0.5 vs. floating = −0.5; L1: Japanese = [1.0, 0], Korean = [−0.5, 1], English = [−0.5, −1]. *** $p < .001$.

Table 7.25 Results of the nested [+definite] ordinal model for acceptability ratings

Fixed Effects	β	SE	z	p
[+definite]				
L1(1): J vs. (K & E)	0.881	0.230	3.823	<.001 ***
L1(2): K vs. E	−0.035	0.216	−0.162	.871
L1: J / quantifier type	0.349	0.323	1.080	.280
L1: K / quantifier type	1.172	0.433	2.709	.007 **
L1: E / quantifier type	0.423	0.309	1.368	.171
[−definite]				
L1(1): J vs. (K & E)	0.963	0.233	4.139	<.001 ***
L1(2): K vs. E	−0.011	0.245	−0.043	.966
L1: J / quantifier type	0.466	0.285	1.638	.101
L1: K / quantifier type	0.371	0.365	1.017	.309
L1: E / quantifier type	0.482	0.309	1.560	.119

Note. Formula: rating ~ L1 / quantifier type + (quantifier type | participant) + (L1 * quantifier type | item). J = Japanese, K = Korean, E = English. Coding: quantifier type: post-nominal = 0.5 vs. floating = −0.5; L1: Japanese = [1.0, 0], Korean = [−0.5, 1], English = [−0.5, −1]. *** $p < .001$, ** $p < .01$.

Table 7.26 Results of the L2-only ordinal model for acceptability ratings for NQ construction items with proficiency effect considered

Fixed Effects	*β*	*SE*	*z*	*p*
definiteness	−0.139	0.481	−0.277	.782
quantifier type	0.235	0.448	0.525	.600
L1 (English vs. Korean)	0.468	0.770	0.608	.543
proficiency	0.050	0.052	0.952	.341
definiteness × quantifier type	−0.364	0.779	−0.467	.640
definiteness × L1	−0.863	0.816	−1.058	.290
quantifier type × L1	0.201	0.896	0.225	.822
definiteness × proficiency	−0.082	0.057	−1.438	.150
quantifier type × proficiency	0.035	0.061	0.582	.561
L1 × proficiency	0.075	0.102	0.738	.461
definiteness × quantifier type × L1	1.410	1.562	0.903	.366
definiteness × quantifier type × proficiency	0.194	0.106	1.826	.068 †
definiteness × L1 × proficiency	−0.002	0.105	−0.016	.987
quantifier type × L1 × proficiency	−0.105	0.122	−0.864	.388
definiteness × quantifier type × L1 × proficiency	−0.161	0.212	−0.757	.449

Note. Formula: rating ~ definiteness * quantifier type * L1 * proficiency + (definiteness * quantifier type * L1 | participant) + (quantifier type * L1 * proficiency | item). Coding: definiteness: [+definite] = 0.5 vs. [−definite] = −0.5; quantifier type: post-nominal = 0.5 vs. floating = −0.5; L1: Korean = −0.5, English = 0.5. Proficiency is centralised cloze test scores. † $p < .10$.

7.4 Discussion

7.4.1 Summary of the results

The results of the SPRT and the AJT can be summarised as follows. The SPRT did not find solid evidence for any of the predicted reading slowdown or facilitation effects (SPRT predictions 1–2 for *sono* and SPRT prediction 4 for NQ constructions, in Section 7.2.3.2) for any groups. These results mean that the pilot SPRT results were overall replicated in that no robust real-time processing evidence was found for native speakers' sensitivity to any target properties in both studies. As to the AJT, all groups showed only those related to *sono* in the anaphoric and non-anaphoric (AJT predictions 1–2, in Section 7.2.3.3). In the bridging context, the native Japanese group, along with the L2 groups, showed a slight preference for the use of *sono* over bare nominals, although this was not observed in Pilot AJT ver. 2.[69] More importantly, however, the native speakers' responses in this study are in sharp contrast with those in the pilot AJTs, which gained statistically robust evidence of the definiteness constraint on NQs consistently. That is, the pilot results were not replicated as far as the AJT was concerned. In the remainder of this section, I will discuss why such an asymmetry was found between the SPRT and the AJT by comparing the native Japanese speakers' data of the present study with those of the pilot studies in detail. Then, some methodological problems will be pointed out with the AJT, thereby partially accounting for the strikingly different AJT results in this study compared with the pilot study. Consequently, although the L2 data will be briefly discussed in relation to the predictions, a conclusion will be drawn only tentatively,

[69] The preference for *sono* NPs over bare NPs in the bridging context is not discussed further, since it is indeed not so surprising given the facts that the bridging context tested in the present thesis was anaphoric in nature (cf. non-anaphoric bridging) and that a numerically higher mean rating was observed in Pilot AJT ver. 1 (Appendix 4A). To accommodate this finding, the prediction for the bridging context will be revised for Main study 2 in the next chapter.

because of the methodological problems. The discussion in this section will set the scene for conducting a revised version of the main AJT.

7.4.2 Comparisons with the pilot studies

Starting with the SPRT data, although the present experiment seems overall similar to the pilot in terms of the target properties, this conclusion needs to be further verified by comparing native Japanese responses in control properties as well. Recall that there were some relatively clear target response patterns observed with some control fillers in the pilot SPRT ver. 2, namely those concerning the noun-classifier agreement and the contrast between post-nominal and pre-nominal NQ constructions regarding the interpretation of proper nouns suffixed by the Japanese plural suffix, -*tati*; and that these results were taken as evidence for the validity of the SPRT as a data collection tool. Indeed, similar patterns also were found in the present SPRT. The native control group that performed the present SPRT also seems sensitive to those filler properties. Linear mixed-effects model analyses of log-transformed residual RTs by each relevant segment and by each property revealed that the effect of condition was significant for the noun-classifier agreement items in the spillover 1 and spillover 2 regions and for the interpretive contrast of proper names with -*tati* in the spillover 1 region, as predicted (the model outputs are summarised in Table 7.27).[70] These results suggest that the present SPRT was overall as appropriately conducted as the pilot. Therefore, it seems that the self-spaced reading paradigm is not suitable for investigating the target interpretive properties of the demonstrative *sono* and NQ constructions.

[70] The native Japanese speakers did not show notable contrasts in the other types of control filler. In this respect too, the results were similar to the pilot SPRT ver. 2.

Table 7.27 Fixed effects estimates for linear mixed effects model of log-transformed RRTs for items regarding noun-classifier agreement and the interpretive contrast of proper nouns with -tati

Fixed Effects	*β*	*SE*	*t*	*p*
Noun-classifier agreement				
Spillover 1 (# 4)				
(Intercept)	0.098	0.021	4.697	.002 **
Condition	−0.140	0.030	−4.736	<.001 ***
Spillover 2 (# 5)				
(Intercept)	0.023	0.021	1.079	.317
Condition	−0.057	0.028	−1.994	.048 *
Interpretive contrast of proper nouns with -tati				
Spillover 1 (# 4)				
(Intercept)	0.161	0.022	7.326	<.001 ***
Condition	−0.062	0.029	2.102	.037 *

Note. Formula: RRT ~ condition + (condition | participant) + (condition | item). Coding: correct vs. wrong = −0.5 vs. 0.5 for noun-classifier agreement; post-nominal vs. pre-nominal = 0.5 vs. −0.5 for the interpretive contrast of proper nouns with *-tati*. *** $p < .001$, ** $p < .01$, * $p < .05$.

Looking at the AJT data, however, there is one striking contrast between the present AJT and the pilot version—the target contrasts between the acceptable and non-acceptable conditions were overall notably less clear in the present study than in the pilot. Table 7.28 compares the mean acceptability ratings of the Pilot AJT ver. 2 and the present AJT (Main study 1).

Table 7.28 Comparisons of native Japanese ratings in the target contexts between Pilot AJT ver. 2 and Main study 1 AJT (SD in parentheses)

	Context	Condition	Pilot ver. 2	Main study 1
Sono	anaphoric	*Sono* NP	4.55 (1.61)	5.30 (1.20)
		Bare NP	3.79 (1.94)	4.58 (1.86)
	non-anaphoric	*Sono* NP	2.14 (2.00)	4.06 (2.04)
		Bare NP	3.60 (2.05)	4.65 (1.82)
	bridging	Sono *NP*	4.09 (1.91)	5.17 (1.32)
		Bare NP	4.10 (2.01)	4.70 (1.81)
NQs	[+definite]	Post-nominal	4.91 (1.68)	4.93 (1.52)
		Floating	3.45 (1.86)	4.71 (1.58)
	[−definite]	Post-nominal	4.67 (1.68)	5.21 (1.23)
		Floating	4.21 (1.65)	4.91 (1.50)

Note. Shaded are unacceptable conditions.

Although the native Japanese speakers in Main study 1 gave generally higher ratings compared to those in Pilot ver. 2, they gave disproportionally higher ratings to the unacceptable conditions (shaded) than the acceptable conditions (unshaded). This is particularly the case with the property of NQ constructions. It is reflected in the fact that although the contrast of *sono* NPs vs. bare NPs in the non-anaphoric context were statistically significant in both studies, the contrast of post-nominal NQs vs. floating NQs in the [+definite] context was only so in Pilot ver. 2. This weaker contrast between the acceptable and unacceptable conditions in Main study 1 compared to Pilot ver. 2 can also be found with non-target items, namely the control fillers, where minimal acceptable vs. unacceptable pairs of target sentences were compared. Table 7.29 provides the meaning acceptability ratings for the control fillers in the two studies. It shows disproportionately greater increases in acceptability from Pilot ver. 2 to Main study 1 in the unacceptable conditions (shaded) than in the acceptable conditions (unshaded), except in the noun-classifier agreement items.

Table 7. 29 Comparisons of native Japanese ratings in the control fillers between Pilot AJT ver. 2 and Main study 1 AJT (SD in parentheses)

Control filler type	Condition	Pilot ver.2	Main study 1
Noun-classifier agreement	Correct	4.43 (1.67)	4.82 (1.51)
	Wrong	0.08 (0.57)	0.41 (0.90)
Specificity constraint of *-tati*	Bare NP	4.86 (1.39)	5.01 (1.74)
	NP-*tati*	3.04 (1.96)	4.62 (1.59)
Common nouns with *-tati*:	Associative	4.28 (1.74)	5.01 (1.55)
	Singular	2.03 (2.20)	3.15 (2.46)
proper nouns with *-tati*	Post-nominal	4.88 (1.52)	4.95 (1.56)
	Pre-nominal	0.64 (1.32)	1.79 (2.31)

Note. Shaded are unacceptable conditions.

In sum, these comparisons of the pilot and Main study 1 confirmed that the results of the Main study SPRT were more or less similar to those of Pilot SPRT ver. 2 : no evidence was found to suggest sensitivity to the target properties but native Japanese speakers showed relatively clear sensitivity to some of the control filler properties. In contrast, compared to Pilot AJT ver. 2, the main AJT showed considerably less clear contrasts between the acceptable and unacceptable conditions regarding the target and non-target properties alike. Such a difference is puzzling, considering the facts that the pilot AJT included essentially similar test items and that the sample sizes in both studies were similar although the present study had a slightly larger sample size (hence potentially a better statistical power) (i.e., Pilot AJT ver. 2, n = 20 ; the main study 1 AJT, n = 26). Next, I will discuss some potential causes for the discrepancy between the Pilot ver. 2 and Main study 1 AJTs.

7.4.3 Potential causes for the difference between the pilot and main AJTs

To identify potential problems regarding the Main study 1 AJT, let us consider two major differences between Pilot ver. 2 and Main study 1 as follows.

One concerns administration format: the Pilot ver. 2 AJT was conducted on a web-based application (Qualtrics) in the absence of the researcher whereas the Main study 1 AJT was conducted on a non-web-based application (i.e., PsychoPy, installed on the researcher's computer) in the presence of the researcher. However, this difference seems unlikely to be the cause of the rating effect in question. Since the nature of the tasks does not differ from one format to the other, given that essentially similar test items were used without time pressure in making judgments, the results should not differ considerably.[71] Particularly, it is not clear why the different administration methods led to the disproportionally higher ratings in the unacceptable conditions compared to the acceptable conditions in the main AJT. Indeed, Sprouse (2011), who evaluated the results of web-based acceptability judgment tasks and corresponding laboratory-based (i.e., non-web-based) versions for their compatibility in terms of participants' responses, reported no meaningful differences between the two formats.

The other difference lies in whether participants have been exposed to the test items in a different task (SPRT) prior to the AJT (Main study 1) or not (Pilot ver. 2). It

---

71 There was indeed one difference worth mentioning between the two AJTs regarding the test items for the definiteness constraint on floating NQs. In Pilot ver. 2, each test sentence included one more adverb than Main study 1. Specifically, in the floating condition, the extra adverb was placed between the NP and the NQ (ia), unlike in the main AJT, where the NQ immediately followed the NP (ib).

(i) *Comparison of the sentence structures in the floating condition*
  a. Pilot ver. 2
    …object NP(-CASE) <u>ADV</u> NQ ADV VP.
  b. Main study 1
    …object NP(-CASE) NQ ADV VP.

The position of the extra adverb in Pilot ver. 2 could possibly lower the acceptability of floating NQs across the contexts, presumably due to its greater structural complexity. However, this can account for the higher acceptability ratings in Main study 1 than Pilot ver. 2 only in the floating condition but not in the post-nominal condition. More importantly, it remains a mystery why the gain in acceptability was considerably bigger in the unacceptable context ([+definite]) than the acceptable context ([−definite]).

has been argued that prior exposition to test sentences can lead to at least two phenomena that could affect acceptability judgements, namely syntactic priming and syntactic satiation.

Syntactic priming is where processing of a certain structure is facilitated by a prior presentation of the same structure. For example, Luka and Barsalou (2005) found that exposure to certain structures in a reading task leads to higher acceptability ratings in a judgement task. They reported that "a single prior exposure to a similar sentence was sufficient to induce this structural facilitation effect." If this is true, the generally higher acceptability found in the Main study 1 AJT may be a result of syntactic priming due to the presentation of the test sentences in the preceding SPRT.

On the other hand, the absence of a contrast between the acceptable sentences and (some of) the unacceptable counterparts in the AJT results might be caused by something akin to syntactic satiation, the phenomenon whereby some sentences that seem ungrammatical at first appear less so after repeated exposures (Snyder, 2000). Snyder (2000, p. 680) argues that "satiation is not an across-the-board phenomenon affecting all sentence types equally", providing evidence that some structures may have higher satiability than others.[72] This seems compatible with the varied degrees of dulled sensitivity to unacceptability among different properties in the present study (e.g., the unacceptability related to the target definiteness properties of *sono* and NQs seems more susceptive to satiation than violations of noun-classifier agreement). Although syntactic priming and syntactic satiation have been originally proposed in the context of (morpho)syntactic felicity, they can seem to naturally extend to similar observations

[72] Sprouse (2009) argues that the satiation effect found in Snyder (2000) is due to a strategy in which participants try to balance the numbers of yes (acceptable) and no (unacceptable) responses when target responses are unevenly distributed on binary options. However, this proposal cannot account for the effect in the present study because the numbers of the acceptable and unacceptable were balanced and a gradable rating scale was used rather than a binary choice—it seems extremely difficult for participants to employ such a strategy in this study.

regarding semantic and pragmatic felicity. That is, it is possible that the rating asymmetry between the Pilot ver. 2 and Main AJTs resulted from a combination of syntactic priming and the satiation of unacceptability due to repeated presentations of the identical test sentences. In other words, the relevant effects might have been caused by the experimental design of Main study 1, namely the prior presentation of the test sentences in the SPRT. These potential methodological problems and the resulting absence of the predicted effects on the native control group have some serious consequences to the interpretation of the L2 learners' data, as discussed next.

7.4.4 Interpretation of L2 data

The AJT data showed that both Korean- and English-speaking learners were more sensitive to the anaphoric marking of *sono* than the definiteness constraint on NQs in the sense that they exhibited target response patterns with the former but not with the latter. This suggests that as the cline of difficulty predicts, the L2 learners have acquired the overt definiteness property of *sono* earlier than the covert definiteness property of NQs. However, these L2 learners' data must be interpreted with caution because syntactic priming and satiation, in principle, could happen to L2 learners as well as native speakers. Clearly, it is wrong to argue that native speakers do not have the relevant linguistic knowledge just because they do not show evidence in experimental settings (sensitivity to the definiteness constraint on NQs in this case): absence of something cannot constitute evidence of its non-existence. Crucially, this logic should be applied to L2 learners who behave similarly to native controls, which is the case with the present study. When evidence of the definiteness constraint cannot be found in the native data, it would be problematic to conclude that the L2 learners have not acquired the relevant properties based on the fact that they failed to provide relevant evidence. In other words, it is impossible to rule out the possibility that the L2 learners do possess the relevant linguistic knowledge but that that knowledge cannot be experimentally observed for a reason (e.g., satiation). The same can be said about the SPRT data—the absence of a reading slowdown/facilitation effect does not mean that the relevant linguistic knowledge is non-existent. However, because the SPRT was administered first, before the AJT, it can be said that the SPRT was free of syntactic priming and satiation effects. This further corroborates the earlier proposal that the target properties of *sono* and NQs are not amenable to detection through the SPRT.

There are some other aspects to the L2 data collected in the present study that potentially compromise their validity and reliability in terms of testing the L2 predictions. The most obvious problem is the small sample size of the Korean speakers (n = 12). The difference in sample size between the groups makes it difficult to compare them reliably.[73] The clear difference in Japanese language proficiency between the Korean and English groups is also potentially problematic in terms of examining the effect of L1 effectively, although no obvious difference was found between the Korean and English groups in the SPRT or the AJT, despite the proficiency gap.

Therefore, given the above discussions, it is hard to evaluate the predictions appropriately, based on the L2 data collected in the present study alone. Although the purpose of running both the SPRT and the AJT was to facilitate comparison of the results between the tasks by preventing potential sources of individual variation from adversely influencing the results, administering the online and offline tasks seems to give rise to a task-induced effect (decreased sensitivity to unacceptability, in particular). This motivates re-running the AJT as a stand-alone task — to be reported in the next chapter.

[73] As already stated, the sample size was due to the COVID 19 pandemic.

7.5 Conclusion

This chapter has reported on a study, where an SPRT and an AJT with the same set of test items were administered to Korean- and English-speaking L2 learners and native speakers of Japanese. The SPRT, which was conducted before the AJT, revealed that none of the groups showed concrete evidence of knowledge of the target interpretive properties of *sono* and NQs. Taken together with the similar pilot SPRT results, it can be said that the self-paced reading paradigm is not suitable for the investigation of those phenomena. The AJT data showed that all groups were sensitive to the target properties of *sono* but not of NQs. The native speakers' insensitivity to the definiteness constraint on NQs was in sharp contrast to the pilot AJT results of the native Japanese participants being clearly sensitive to the property. A detailed comparison of the pilot and present AJTs revealed that participants in the present study had the tendency to give generally higher ratings compared to those in the pilot; and that most of the unacceptable sentences were disproportionally rated higher than their acceptable counterparts, which resulted in general lack of relevant acceptable vs. unacceptable contrasts. I argued that this was due to a combination of syntactic priming and satiation to unacceptability, which were triggered presumably by the prior presentation of the test items in the preceding task (the SPRT). Although it is not clear why some properties satiated to greater degrees than others, it is clear that such an effect must be avoided for the purpose of the present study. Particularly, the considerably subtle or non-existent unacceptable vs. acceptable contrasts in the target properties with the native controls makes it difficult to reliably evaluate linguistic competence of the participants. This, in turn, becomes a serious obstacle in interpreting similarly unclear L2 data in terms of evaluating predictions for the L2 acquisition of the target properties. The findings of the present study and reasoning about them motivate a re-implementation of the AJT.

Chapter 8: Main study 2 (web-based AJT with L2 Japanese learners)

8.1 Introduction

This chapter documents an additional study targeting new groups of L2 learners and native Japanese speakers, with its main objective being a re-implementation of the AJT without a preceding SPRT using the same test sentences. This separate AJT was primarily motivated by the argument in the previous chapter that the participants' prior exposure to the target sentences in the SPRT could have affected their behaviour in the AJT. Specifically, it was suggested that syntactic priming could have increased ratings on the acceptable AJT sentences, and syntactic satiation could have attenuated judgements of unacceptable sentences.

In what follows, after a description of the experimental design and other relevant details, results of the present experiment (Main study 2, henceforth) will be compared to the results of Pilot ver. 2 and Main study 1. Particularly, it will be shown that the native responses in Main study 2 match those in Pilot ver. 2 rather than those in Main study 1, in terms of the acceptable vs. unacceptable contrasts in the target properties of *sono* and NQs. This will be taken as evidence of the alleged satiation effect on acceptability judgement in Main study 1, which, in turn, suggests that Main study 2 provides more reliable AJT data than Main study 1. Finally, Main study 2 L2 data will be used to evaluate the predictions based on the FRH and the cline of difficulty for the relative difficulty or ease of the acquisition of the target properties and the L1 difference therein.

8.2 Method

8.2.1 Participants

Twenty English-speaking and 20 Korean-speaking learners of Japanese participated in the study. The L1-English group included students enrolled at universities in the UK on Japanese-related programmes, Japanese language teachers working at UK and US institutions, and university lecturers at Japanese universities. All but one of the Korean-speaking participants were resident in Japan at the time of testing, being undergraduate students at Japanese universities or Korean language teachers based in Japan. The remaining one L1-Korean participant was studying in the US. As with Main study 1, those with intermediate or more advanced Japanese language proficiency (JLPT N3 or higher) were targeted. Table 8.1 summarises demographic information about the L2 participants.

Table 8.1 Participant information: Main study 2

Group	Age (yrs old)	Age of onset (yrs old)	Formal Japanese education (yrs)	Length of residency in Japan (yrs)
Korean speakers	25.20	14.85	6.03	3.34
(n = 20)	(6.48)	(3.60)	(5.58)	(4.46)
English speakers	29.90	17.65	9.03	6.19
(n = 20)	(12.77)	(3.44)	(8.00)	(10.37)

Note. () = *SD*.

Twenty native Japanese-speaking participants were recruited as controls. They were university students and professionals, resident in Japan (M age = 34.0, SD = 11.92). Only the L2 groups took a proficiency test (the cloze task reported in the next section) since the native data was already collected in Main study 1. The participants received compensation for their participation.

8.2.2 Test instruments

8.2.2.1 Proficiency test

The L2 learner groups took the same cloze test as used in Main Study 1, but rather than a pencil-and-paper format, it was converted into a web-based format, using Google Forms. The participants completed a passage by filling 42 blanks with the correct words each selected from four options. The scores of the native Japanese speakers from Main study 1 serve as a reference level. Performance is summarised in Table 8.2.

Table 8.2 Scores on cloze test (0–42 points): Main study 2

Group	*M*	*SD*	Range
Japanese controls (*n* = 26)	38.62 (92%)	2.17	34–42 (81–100%)
Korean speakers (*n* = 20)	36.70 (87%)	3.37	28–42 (67–100%)
English speakers (*n* = 20)	34.15 (81%)	4.63	26–42 (62–100%)

The results show that the L2 learners can be regarded as relatively at advanced levels in the sense that their mean scores seem relatively high and their score ranges largely overlap with those of the native speakers. The scores were analysed with a one-way between-participants ANOVA, which found a significant effect of L1 on proficiency score (F (2, 63) = 9.58, p < .001). The Tukey HSD multiple comparison test revealed that the L1-English group was significantly different from the Japanese group (*M-Diff* = 4.47, *95% CI* [2.02, 6.19], p < .001). However, the L1-Korean group did not significantly differ from either the native control group or L1-English group. Therefore, it can be said that the L2 groups have more or less matched proficiency, although the Korean group has more individuals scoring in the native speakers' range.

8.2.2.2 AJT

This AJT included the same set of the test sentences as the Main study 1 AJT but was conducted in a web-based format using Qualtrics (version Feb. 2020). The participants were asked to read contextualizing preambles and rate the test sentences as natural continuations on a 7-point Likert scale (0 = completely odd; 6 = completely natural) with an *I don't know* option. Two lists of test items were used and each list included 96 items, 40 of which were critical items (8 items × 5 contexts), 32 control fillers (8 items × 4 types), and 24 unacceptable fillers (4 items × 6 types) (identical to the main study 1 item lists). The critical items and control fillers were constructed in minimally different pairs. Only one member of each pair occurred within a given list. After making their judgement, participants pressed a button to proceed to the next item. It was impossible to go back and change answers, by design.

Sample test items for the critical items are presented below. The expected native speaker judgement pattern for *sono* items is to rate *sono* NPs lower than bare NPs in the non-anaphoric context (8.2); and *sono* NPs higher than bare NPs or equally accept both NP types in the anaphoric context (8.1). The prediction for the anaphoric context can also be applied to the bridging context (8.3), given the preference for *sono* NPs over bare NPs found in Main study 1 (Chapter 7).[74]

As to the NQ constructions, floating NQs were expected to receive significantly lower ratings than post-nominal NQs in the [+definite] context (8.4), whereas both types of NQ were rated equally acceptable in the [−definite] context (8.5). (As in the previous versions, the preambles were presented in Japanese using Japanese

---

[74] Recall that the pilot AJTs (Chapter 6) did not find any statistically meaningful difference in acceptability between *sono* NPs and bare NPs in the bridging context—the two types of NP were rated more or less equally high. However, since the preference for *sono* was observed in the bridging context as in the anaphoric context in Main study 1, the prediction for Main study 2 was revised to accommodate this finding.

script in the actual test, but are presented here in English for convenience.)

(8.1) *Sono*: anaphoric (*sono* NP vs. bare NP)

PREAMBLE: *There is a very popular restaurant which serves great food for reasonable prices near Taroo's house.*

Taroo-wa mukasi-kara { sono resutoran-o vs. resutoran-o }

Taroo-TOP past-from { SONO restaurant-ACC vs. restaurant-ACC }

dare-yori-mo yoku riyoo-site-i-masu

more.than.anyone often use-do-ASP-POL.NPST

'Taroo has eaten at the restaurant more often than anyone since long ago.'

(8.2) *Sono*: non-anaphoric (*sono* NP vs. bare NP)

PREAMBLE: *Hanako had a pleasant dream recently. In that dream, ...*

Hanako-wa tori-no yooni {sono sora-o vs. sora-o} jiyuuni

Hanako-TOP like.a.bird { SONO sky-ACC vs. sky-ACC } freely

kimotiyoku tobi-masita

pleasantly fly-POL.PST

'Hanako enjoyed flying freely in the sky like a bird.'

(8.3) *Sono*: bridging (*sono* NP vs. bare NP)

PREAMBLE: *Hanako is now into a popular novel.*

Hanako-wa guuzen { sono tyosya-o vs. tyosya-o }

Hanako-TOP accidentally { SONO author-ACC vs. author-ACC }

Tokyo-de issyun-dake mi-ta koto-ga ari-masu

Tokyo-in one.moment-only have.seen

'Hanako has caught sight of the author in Tokyo once.'

(8.4) NQs: [+definite] (post-nominal vs. floating)

PREAMBLE: *Taroo has one son and daughter and always takes care of them all by himself in the morning.*

Taroo-wa kesa-mo { kodomo huta-ri-o vs. kodomo-o huta-ri }

Taroo-TOP this.morning-too { child-2-CL-ACC vs. child-ACC 2-CL }

itumo-no yooni siti-zi ni okosi-masita

as.always 7-o'clock at wake.up- POL.PST

' Taroo, as always, woke the two children up at 7 this morning.'

(8.5) NQs: [−definite] (post-nominal vs. floating)

PREAMBLE: *A new restaurant just opened near Hanako's house and she wants to go there. But, since she does not want to go there alone, ...*

Hanako-wa kinoo { yuuzin huta-ri-o vs. yuuzin-o huta-ri }

Hanako-TOP yesterday { friend-2-CL-ACC vs friend-ACC 2-CL }

sassoku ranti-ni sassote-mi-masita

immediately lunch-for ask.out-try.to-POL.PST

'Hanako went ahead and asked two friends out for lunch yesterday.'

8.2.3 Procedure

Participation was by means of a web-based survey. For the L2 participants, the survey consisted of an information sheet and consent form; a short demographic questionnaire; the AJT, which started with instructions, then a training session with practice examples, followed by the main task; and finally, the cloze task. The whole process took 60–70 minutes for the L2 participants. The Japanese native controls completed all components but the cloze test within 45 minutes. The participants were compensated for their time with online vouchers.

8.3 Results

8.3.1 *Sono*

The group means of the acceptability ratings for each NP type are summarised in Table 8.3. Figures 8.1 through 8.3 illustrate the data for each context. *I don't know* responses were removed, which affected 1.25% of the data (anaphoric: 4/480 (all from the Korean group); non-anaphoric: 2/480 (both from the Korean group); bridging anaphoric: 12/480 (5 from the Korean group and 7 from the English group).

Table 8.3 Mean acceptability ratings by context and by group: Main study 2 sono *items (SD in parentheses)*

Context	Condition	Native controls	Korean speakers	English speakers
anaphoric	*Sono* NP	5.38 (1.26)	4.60 (1.93)	4.68 (1.76)
	Bare NP	3.90 (2.01)	4.26 (2.10)	4.03 (1.80)
Non-anaphoric	*Sono* NP	2.48 (2.18)	2.95 (2.36)	2.93 (2.20)
	Bare NP	4.65 (1.71)	4.16 (2.11)	4.60 (1.61)
Bridging	*Sono* NP	5.16 (1.16)	4.03 (2.16)	3.94 (2.08)
	Bare NP	4.38 (1.61)	3.58 (2.23)	3.33 (2.18)

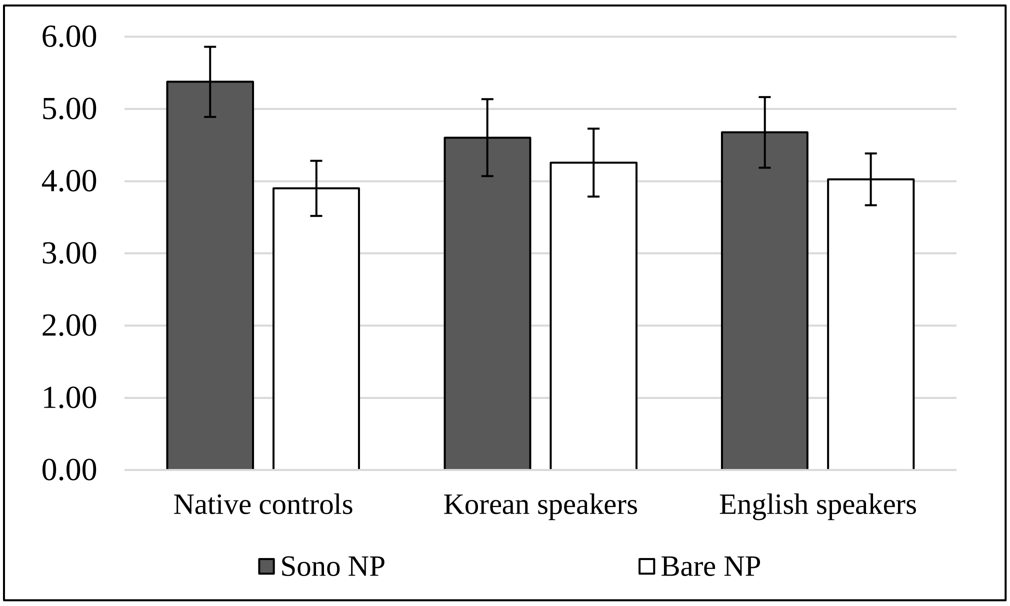


Figure 8.1 Mean acceptability ratings in the anaphoric context (error bars = SE)

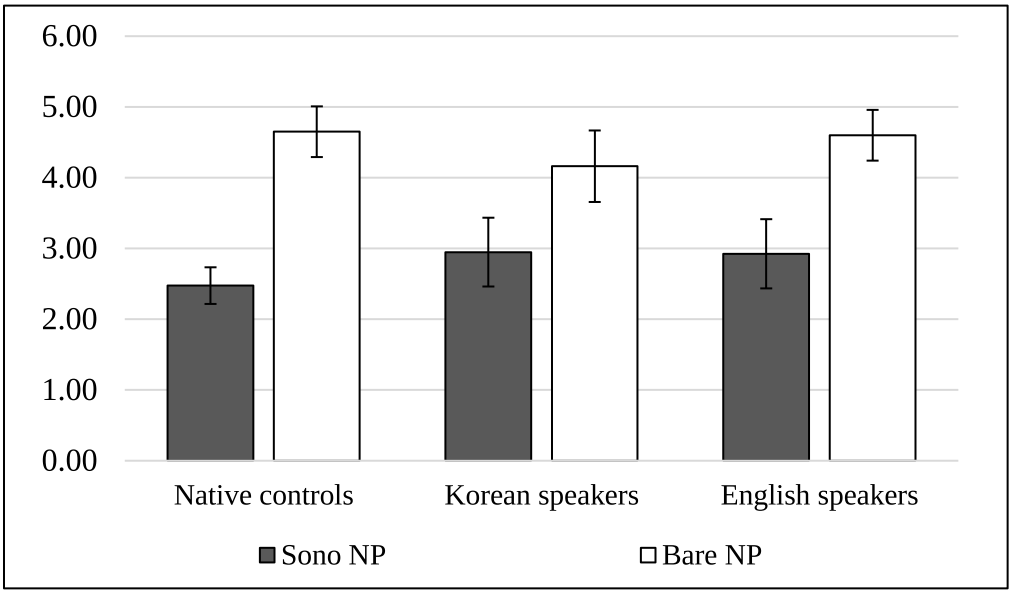


Figure 8.2 Mean acceptability ratings in the non-anaphoric context (error bars = SE)

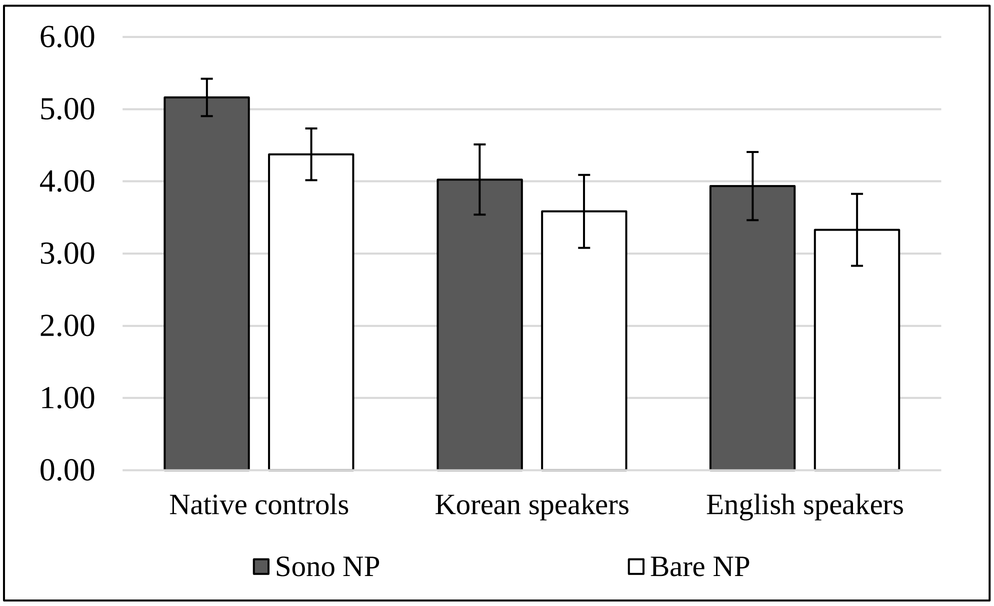


Figure 8.3 Mean acceptability ratings in the bridging context (error bars = SE)

In the anaphoric context (Figure 8.1), the native control group gave *sono* NPs lower ratings than bare NPs, as predicted. The same pattern was attested in the L2 groups. In the non-anaphoric context (Figure 8.2), the native control group rated *sono* NPs clearly lower than bare NPs, which is the predicted pattern. The learner groups also exhibited the same patterns. In the bridging anaphoric context (Figure 8.3), the native control group, in line with the prediction, rated *sono* NPs higher than bare NPs. The difference between the noun types in the L2 groups was also in the target direction.

For further analysis, an ordinal mixed-effects model was fitted to the acceptability ratings. The model included as fixed effects CONTEXT (*anaphoric* vs. *non-anaphoric* vs. *bridging*), NP TYPE (sono *NP* vs. *bare NP*), L1 (*Japanese* vs. *Korean* vs. *English*) and the interactions between the three variables. As random intercepts, participants and items were added as well as by-participant slopes for NP TYPE, CONTEXT and their interaction, and by-item slopes for L1 (the maximal possible model that converged). CONTEXT and NP TYPE were sum-coded whereas L1 was Helmert-coded so that the Japanese group (native controls) and the combination of the Korean and English groups (L2 groups) were compared first, and then the Korean and English groups were compared against each other (as in Main study 1).

The model output is presented in Table 8.4. There were effects of CONTEXT-related variables (contexts 1 & 2), whereas no effect was found for NP TYPE. However, the two-way interactions of CONTEXT (contexts 1 & 2) and NP TYPE were significant. There were also a marginal effect of the contrast between the native and L2 groups and a marginal interaction between this contrast and NP TYPE. The interaction between CONTEXT (context 1) and the native vs. L2 contrast was significant. Furthermore, there were three-way interactions between CONTEXT (contexts 1 & 2), NP TYPE and the native vs. L2 contrast. There was no effect of the Korean vs. English contrast or interactions involving this contrast. The three-way interactions suggest that the effect of NP TYPE depends on the context and whether the group is native or L2; whereas the absence of three-way interactions involving the Korean vs. English contrast means that the effect of NP TYPE within each context does not differ significantly between the two L2 groups.

To investigate further how the effect of NP TYPE differs from one context to another within each group, post hoc nested ordinal mixed effects models were constructed for each context. Table 8.5 presents the results of the nested model analyses (the model specifications can be found in the notes). The anaphoric and bridging models revealed that the native control group rated *sono* NPs significantly higher than bare NPs. The non-anaphoric model found that the native controls rated *sono* NPs significantly lower than bare NPs. On the other hand, neither of the L2 groups distinguished between the NP types at statistically significant levels in the bridging context. In the anaphoric context, the English-speaking learners rated *sono* nouns higher than bare nouns at a statistically significant level, but the Korean-speaking learners did not. However, both L2 groups rated *sono* NPs significantly lower than bare NPs in the non-anaphoric context. Commonly seen across the three contexts is the noticeably larger effect of NP TYPE for the native control group compared to the L2 groups, which is apparently the main source of the three-way interaction in the omnibus model between CONTEXT, NP TYPE and the native vs. L1 contrast. Despite the interaction, it can be said both L2 groups distinguished the NP types in the target direction relatively clearly in the non-anaphoric context, given the significant effects of NP TYPE in the nested model results. As to the other two contexts, the L2 groups differentiated the NP types less clearly than they did in the non-anaphoric context, although their judgment patterns were still compatible with the predicted target judgement. The non-significant vs. significant contrast found between the L2 groups in terms of the NP TYPE effect in the anaphoric context suggests that the English speakers rated *sono* NPs higher than bare NPs (i.e., the target pattern) more consistently than the Korean speakers. However, there is no concrete evidence of this L1 difference, because of the absence of three-way interactions between CONTEXT, NP TYPE and the Korean vs. English contrast in the omnibus model.

In sum, the native control group made the predicted judgements: they rated *sono* NPs higher than bare NPs in the anaphoric and bridging conditions and did the opposite in the non-anaphoric condition. The L2 groups also showed the predicted native judgement patterns across the contexts, although their distinction between the NP types was in general statistically different from the native controls'. Crucially, both Korean and English groups made a statistically significant distinction between the two NP types in the non-anaphoric context, which suggests knowledge of the target property of *sono*, namely overt anaphoricity marking.

Table 8.4 Results of the omnibus ordinal model for acceptability ratings for sono *items*

Fixed Effects	β	*SE*	*z*	*p*
context 1	−0.911	0.347	−2.625	.009 **
context 2	0.989	0.354	2.789	.005 **
NP type	−0.015	0.104	−0.145	.884
L1(1): Japanese vs. (Korean & English)	0.293	0.171	1.714	.087 †
L1(2): Korean vs. English	0.057	0.140	0.403	.687
context 1 × NP type	3.819	0.490	7.801	<.001 ***
context 2 × NP type	−2.298	0.358	−6.417	<.001***
context 1 × L1(1): Japanese vs. (Korean & English)	−0.710	0.281	−2.530	.011 *
context 2 × L1(1): Japanese vs. (Korean & English)	−0.075	0.298	−0.250	.803
context 1 × L1(2): Korean vs. English	−0.197	0.202	−0.974	.330
context 2 × L1(2): Korean vs. English	0.118	0.217	0.545	.586
NP type × L1(1): Japanese vs. (Korean & English)	−0.277	0.145	−1.911	.056 †
NP type × L1(2): Korean vs. English	0.048	0.125	0.382	.703
context 1 × NP type × L1(1): Japanese vs. (Korean & English)	1.351	0.677	1.997	.046 *
context 2 × NP type × L1(1): Japanese vs. (Korean & English)	−1.280	0.502	−2.549	.011 *
context 1 × NP type × L1(2): Korean vs. English	−0.369	0.588	−0.628	.530
context 2 × NP type × L1(2): Korean vs. English	0.386	0.429	0.900	.368

Note. Formula: rating ~ context * L1 * NP type + (context * NP type | participant) + (L1 | item). Coding: context: anaphoric = [0, 0.5] vs. non-anaphoric = [0.5, 0] vs. bridging = [−0.5, −0.5]; NP type: *sono* NP = −0.5 vs. bare NP = 0.5; L1: Japanese = [1.0, 0], Korean = [−0.5, 1], English = [−0.5, −1]. *** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$.

Table 8.5 Results of nested models for acceptability ratings for each context of sono *items*

Fixed Effects	*β*	*SE*	*z*	*p*
Anaphoric				
L1(1): J vs. (K & E)	0.204	0.209	0.977	.328
L1(2): K vs. E	0.167	0.180	0.926	.354
L1: J / NP type	−1.973	0.349	−5.651	<.001 ***
L1: K / NP type	−0.384	0.338	−1.135	.256
L1: E / NP type	−0.905	0.320	−2.831	.005 **
Non-anaphoric				
L1(1): J vs. (K & E)	−0.109	0.180	−0.603	.546
L1(2): K vs. E	−0.014	0.153	−0.092	.927
L1: J / NP type	2.291	0.590	3.880	<.001 ***
L1: K / NP type	1.434	0.620	2.311	.021 *
L1: E / NP type	1.856	0.598	3.102	.002 **
Bridging				
L1(1): J vs. (K & E)	0.514	0.194	2.648	.008 **
L1(2): K vs. E	0.052	0.167	0.309	.757
L1: J / NP type	−0.883	0.281	−3.139	.002 **
L1: K / NP type	−0.407	0.327	−1.244	.213
L1: E / NP type	−0.517	0.321	−1.607	.108

Note. Formula: *anaphoric*: rating ~ L1 / NP type + (NP type | participant) + (NP type | item); *non-anaphoric* & *briding* : rating ~ L1 / NP type + (NP type | participant) + (NP type * L1 | item). J = Japanese, K = Korean, E = English. "*x* / *y*" represents the effect of variable *y* with variable *x* held constant. Coding: context: anaphoric = [0, 0.5] vs. non-anaphoric = [0.5, 0] vs. bridging = [−0.5, −0.5]; NP type: *sono* NP = −0.5 vs. bare NP = 0.5; L1: Japanese = [1.0, 0], Korean = [−0.5, 1], English = [−0.5, −1]. *** $p < .001$, ** $p < .01$, * $p < .05$.

8.3.2 NQ constructions

Mean acceptability ratings are presented by context, condition, and group in Table 8.6. Figures 8.4 and 8.5 illustrate the data for the [+definite] context and the [−definite] context, respectively. *I don't know* responses were removed, which affected 1.04 % of the data (10/960: 3 from the English group and 7 from the Korean group split roughly evenly between the [+definite] and [−definite] contexts).

Table 8.6 Mean acceptability ratings by context and by group: Main study 2 NQ construction items (SD in parentheses)

Context	Condition	Native controls	Korean speakers	English speakers
[+definite]	Post-nominal	4.88 (1.67)	4.38 (2.16)	3.77 (2.25)
	Floating	3.73 (2.07)	3.21 (2.37)	3.62 (2.26)
[−definite]	Post-nominal	4.75 (1.55)	4.44 (1.97)	3.67 (2.19)
	Floating	4.95 (1.44)	3.86 (2.23)	4.14 (1.83)

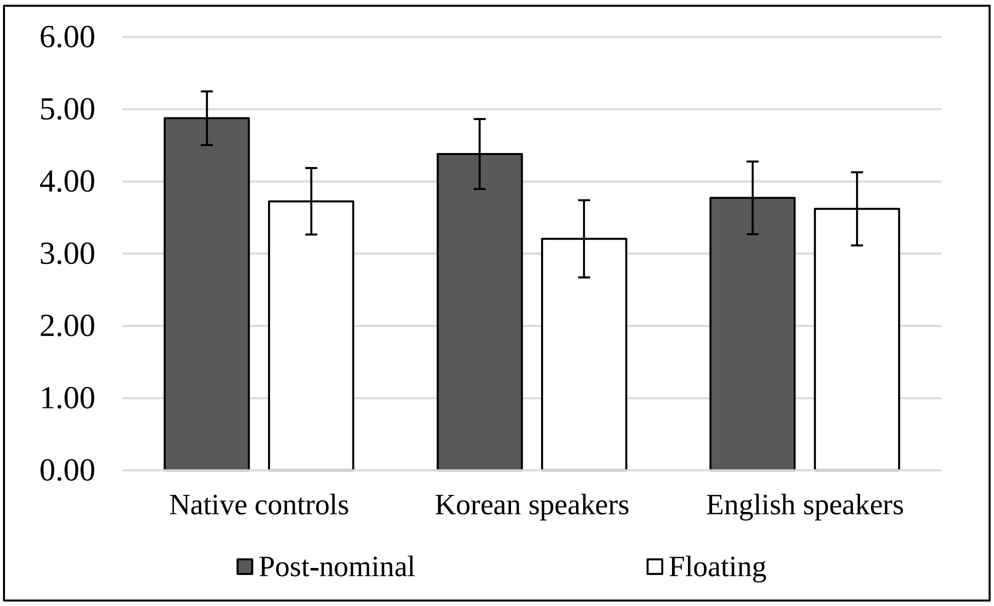


Figure 8.4 Mean acceptability ratings in the [+definite] context (error bars = SE)

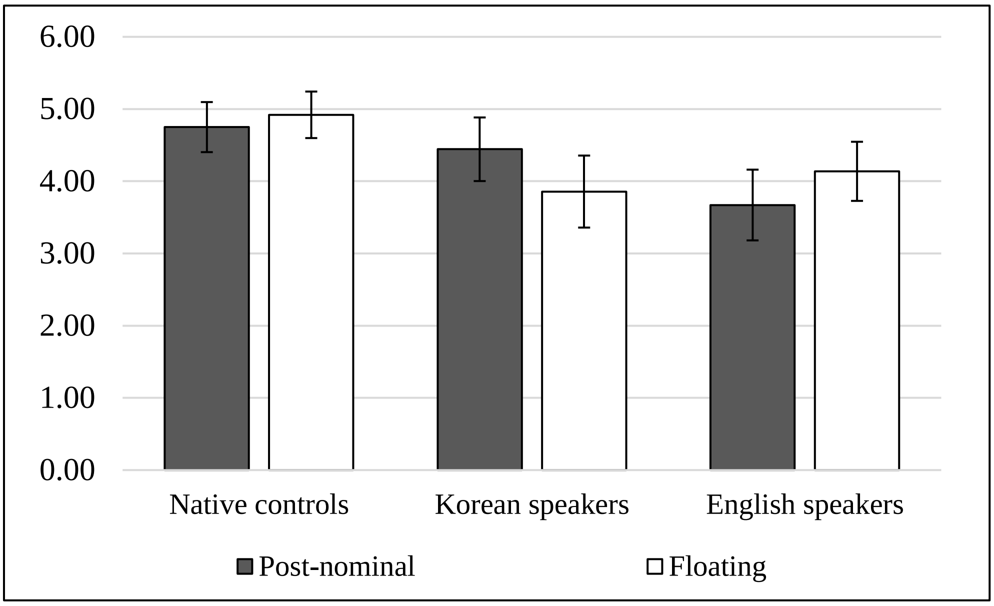


Figure 8.5 Mean acceptability ratings in the [−definite] context (error bars = SE)

Recall the predicted native Japanese responses are (i) significantly lower ratings for floating NQs than post-nominal NQs in the [+definite] context; and (ii) equally high ratings for both NQs in the [−definite] context. It can be seen that the native Japanese control group and the Korean group both gave clearly lower ratings to floating NQs than post-nominal NQs in the [+definite] context (Figure 8.4); whereas they gave more or less equally high ratings to both types of NQ in the [−definite] context (Figure 8.5). On the other hand, the English group accepted the two types of NQ roughly equally in both [+definite] and [−definite] contexts.

For analysis, an ordinal mixed-effects model was fitted to the acceptability ratings. The model was maximally specified: the fixed effects included DEFINITENESS (*[+definite]* vs. *[−definite]*), QUANTIFIER TYPE (*post-nominal* vs. *floating*), and L1 (*Japanese* vs. *Korean* vs. *English*), and their interactions; whereas the random effects included intercepts for participants and items along with by-participant slopes DEFINITENESS, QUANTIFIER TYPE, and by-item slopes for QUANTIFIER TYPE, L1 and their interactions. L1 was contrasted with Helmert coding (contrast 1: native vs. L2, contrast 2: Korean vs. English) whereas DEFINITENESS and QUANTIFIER TYPE were sum-coded (as in Main study 1). Table 8.7 shows the model output. The model suggests that there were effects of QUANTIFIER TYPE, the contrast between the native controls and the two L2 groups combined. There were two-way interactions between DEFINITENESS and QUANTIFIER TYPE and between QUANTIFIER TYPE and the Korean vs. English contrast. There were no other effects or interactions. The interaction between DEFINITENESS and QUANTIFIER TYPE provides evidence of sensitivity to the definiteness constraint in the form of lower ratings of floating NQs than post-nominal NQs in the [+definite] context. Crucially, the absence of any three-way interactions between DEFINITENESS, QUANTIFIER TYPE and L1-related contrasts mean that the sensitivity to the constraint does not significantly differ between the groups. This is surprising in that, as can be seen in Figure 8.4, the English group did not seem to distinguish clearly the two NQ types in the [+definite] context. To understand this puzzling result, I conducted post hoc comparisons of the effect of QUANTIFIER TYPE within each group, using nested ordinal mixed-effects models as well as analysis of individual response patterns for the [+definite] context.

The results of the nested comparisons are provided in Table 8.8. The [+definite] model found that the effect of QUANTIFIER TYPE was significant for the native control and Korean groups but not for the English group. These model results cannot provide solid evidence for any difference between the groups given the absence of three-way interaction in the omnibus model. However, they suggest that the Korean group is more consistent in rating floating NQs less acceptable than post-nominal NQs, compared to the English group. The [−definite] model found no significant effect of QUANTIFIER TYPE for the native control and English group, but a marginal effect for the Korean group (p = .058). Thus, overall, all groups rate both NQ types roughly equally despite the Korean group's tendency to rate the post-nominal NQs slightly higher than the floating NQs.[75]

Turning to the individual response patterns, Table 8.9 summarises (i) how many participants in each group rated floating NQs lower than post-nominal NQs (target pattern), (ii) how many showed no differences between the two, and (iii) how many had higher ratings for floating NQs than post-nominal NQs (non-target pattern), along with the ranges of difference sizes in each category. It can be seen that the English group had the greatest amount of non-target response patterns. Crucially, however, the majority of the English speakers fell into the target pattern, despite the apparent lack of the target contrast in the aggregate data in Figure 8.4. Therefore, the overall preferred pattern for all three groups is the target (lower ratings for floating NQs than post-nominal NQs) in the [+definite] context. In other words, the majority of participants in all three groups appear to be sensitive to the definiteness constraint. This accounts for the absence of a three-way interaction in the main model even though the pattern in the descriptive data (Table 8.6, Figure 8.4) suggests that the L1-English group may be different to the native Japanese and L1-Korean groups.

75 The Korean group's tendency to favour post-nominal NQs over floating NQs even in the [−definite] may indicate L1 influence in terms of the distribution of NQs. That is, post-nominal NQs seem to be used more frequently than floating NQs in Korean: Kim and Yang's (2006a, 2006b) corpus studies show that whereas post-nominal numerals account for about 15% of all instances of NQs (n = 694), floating NQs account for only about 5 %. However, the opposite appears to be true with Japanese: according to a Kim's (1995) survey of Modern Japanese texts from different genres, post-nominal and floating NQs occur about 6 % and about 21 % of the total number of all NQs (n = 858), respectively. This distributional difference may account for each group's mild preference for one type of NQ over the other in the [−definite] context. However, this Korean preference for post-nominal NQs cannot explain the considerably larger effect of QUANTIFIER TYPE in the [+definite] context than in the [−definite] context.

Table 8.7 Results of the omnibus ordinal model for acceptability ratings for NQ construction items

Fixed Effects	*β*	*SE*	*z*	*p*
definiteness	−0.223	0.356	−0.627	.531
quantifier type	0.545	0.176	3.099	.002 **
L1(1): Japanese vs. (Korean & English)	0.438	0.211	2.071	.038 *
L1(2): Korean vs. English	0.143	0.157	0.917	.359
definiteness ×quantifier type	0.948	0.297	3.186	.001 **
definiteness × L1(1): Japanese vs. (Korean & English)	−0.198	0.311	−0.637	.524
definiteness × L1(2): Korean vs. English	−0.020	0.192	−0.104	.917
quantifier type × L1(1): Japanese vs. (Korean & English)	0.060	0.250	0.242	.809
quantifier type × L1(2): Korean vs. English	0.607	0.209	2.905	.004 **
definiteness ×quantifier type × L1(1): Japanese vs. (Korean & English)	0.575	0.426	1.348	.178
definiteness ×quantifier type× L1(2): Korean vs. English	0.163	0.347	0.470	.638

Note. Formula: rating ~ definiteness * quantifier type * L1 + (definiteness * quantifier type | participant) + (quantifier type * L1 | item). Coding: definiteness: [+definite] = 0.5 vs. [−definite] = −0.5; quantifier type: post-nominal = 0.5 vs. floating = −0.5; L1: Japanese = [1.0, 0], Korean = [−0.5, 1], English = [−0.5, −1]. ** $p < .01$, * $p < .05$.

Table 8.8 Results of separate nested ordinal models for acceptability ratings for each condition of NQ construction items

Fixed Effects	*β*	*SE*	*z*	*p*
[+definite] condition				
L1(1): J vs. (K & E)	0.331	0.198	1.667	.095 †
L1(2): K vs. E	0.109	0.164	0.666	.506
L1: J / quantifier type	1.271	0.409	3.110	.002**
L1: K / quantifier type	1.450	0.492	2.950	.003**
L1: E / quantifier type	0.190	0.424	0.447	.655
[–definite] condition				
L1(1): J vs. (K & E)	0.604	0.331	1.826	.068 †
L1(2): K vs. E	0.168	0.211	0.799	.424
L1: J / quantifier type	−0.213	0.376	−0.568	.570
L1: K / quantifier type	0.726	0.384	1.893	.058 †
L1: E / quantifier type	−0.364	0.367	−0.992	.321

Note. Formula: rating ~ L1 / quantifier type + (quantifier type | participant) + (quantifier type * L1 | item). "*x* / *y*" represents the effect of variable *y* with variable *x* held constant. ** *p* < .01, † *p* < .10.

Table 8.9 Distribution of response patterns within each group in the [+definite] context

L1	Floating mean rating < post-nominal mean rating		Floating = post-nominal	Floating mean rating > post-nominal mean rating	
	n	Size of difference	*n*	*n*	Size of difference
Japanese	15	0.25–3.25	3	2	0.25
Korean	16	0.33–4.50	1	3	0.25–1.17
English	11	0.25–5.25	0	9	0.25–4.25

Note. *n* = number of participants with the given response pattern. Size of difference is given in points on the rating scale.

8.4 Discussion

In the first part of this section, the results of the Main study 2 AJT will be compared with those of the Main study 1 and Piot ver. 2 AJTs, focusing on native Japanese speakers' data, with a view to examining whether the revised experimental design was successful in avoiding satiation, suspected of causing the unclear target contrasts in the Main study 1 AJT. Once it is demonstrated that the experiment seems free of the relevant effect (hence reliable), the predictions for the L2 acquisition of the target properties will be evaluated in the second part.

8.4.1 Comparisons with Main study 1 and Pilot ver. 2

To recap, (8.6) and (8.7) states what had been observed in Main study 1.

(8.6) The ratings on both acceptable and unacceptable sentences were higher in Main study 1 than in Pilot ver. 2.

(8.7) Some contrasts in ratings between acceptable and unacceptable sentences in Pilot ver. 2 were no longer present in Main study 1.

I proposed that prior exposure to the test sentences via the SPRT could have caused these phenomena, and specifically, that syntactic priming via the SPRT could have led to the increased acceptance of the acceptable sentences, while satiation could have led to decreased sensitivity to infelicity. If the subtle contrasts between the acceptable vs. unacceptable conditions in Main study 1 had resulted from the repeated presentation of the test items due to participants completing the SPRT before the AJT, the relevant target contrasts observed in the Pilot ver. 2 AJT should re-emerge in Main study 2, where participants were not exposed to the same test sentences prior to the judgement task.

Table 8.10 compares the ratings for the target properties between the three studies. Firstly, note that the ratings for the acceptable conditions (unshaded) in Main study 2 are overall similar to those in Main study 1. Given that the exact same set of test items was used in these two studies, it seems that what appeared to be a syntactic priming effects (increased acceptability due to facilitated processing by repeated presentation of the same structures) did not occur in Main study 1. Thus, it seems more reasonable to attribute the increased acceptability to the revisions aimed to improve general acceptability in the acceptable conditions, rather than to syntactic priming. On the other hand, the ratings for the unacceptable sentences are lower than in Main study 1, more like in Pilot ver. 2. This suggests that satiation did indeed occur in Main study 1: the participants' sensitivity to infelicity was dulled by their prior exposure. Similar trends can be seen in the ratings of the control fillers. Table 8.11 summarises native speakers' responses to the control fillers. The ratings for the control fillers in Main study 2 have the same two characteristics as the target items: (i) clearer contrasts in ratings between the acceptable and unacceptable conditions than Main study 1, and (ii) ratings in the acceptable conditions generally matching those in Main study 1. All these results taken together, it seems natural to assume that a satiation effect did indeed take place in Main study 1 (but syntactic priming did not), and Main study 2 was successful in avoiding the problem. Therefore, I consider the results of Main study 2 to be a more reliable (less contaminated) reflection of participants' knowledge status compared to the Main study 1.

Table 8.10 Comparisons of native Japanese acceptability ratings in the target contexts between Pilot ver. 2, Main studies 1 and 2 (SD in parentheses)

Context		Condition	Pilot ver. 2	Main study 1	Main study 2
Sono	Anaphoric	*Sono* NP	4.55 (1.61)	5.30 (1.20)	5.38 (1.26)
		Bare NP	3.79 (1.94)	4.58 (1.86)	3.90 (2.01)
	Non-anaphoric	*Sono* NP	2.14 (2.00)	4.06 (2.04)	2.48 (2.18)
		Bare NP	3.60 (2.05)	4.65 (1.82)	4.65 (1.71)
	Bridging	*Sono* NP	4.09 (1.91)	5.17 (1.32)	5.16 (1.16)
		Bare NP	4.10 (2.01)	4.70 (1.81)	4.38 (1.61)
NQs	[+definite]	Post-nominal	4.91 (1.68)	4.93 (1.52)	4.88 (1.67)
		Floating	3.45 (1.86)	4.71 (1.58)	3.73 (2.07)
	[−definite]	Post-nominal	4.67 (1.68)	5.21 (1.23)	4.75 (1.55)
		Floating	4.21 (1.65)	4.91 (1.50)	4.95 (1.44)

Note. Shaded are unacceptable conditions.

Table 8.11 Comparisons of native Japanese acceptability ratings in the control fillers between Pilot ver. 2, Main studies 1 and 2 (SD in parentheses)

Control filler type	Condition	Pilot ver. 2	Main study 1	Main study 2
Noun-classifier agreement	Correct	4.43 (1.67)	4.82 (1.51)	4.94 (1.51)
	Wrong	0.08 (0.57)	0.41 (0.90)	0.03 (0.22)
Specificity constraint of *-tati*	Bare NP	4.86 (1.39)	5.01 (1.74)	5.06 (1.25)
	NP-*tati*	3.04 (1.96)	4.62 (1.59)	3.19 (2.15)
Common nouns with *-tati*	Associative	4.28 (1.74)	5.01 (1.55)	4.71 (1.76)
	Singular	2.03 (2.20)	3.15 (2.46)	2.11 (2.28)
Proper nouns with *-tati*	Post-nominal	4.88 (1.52)	4.95 (1.56)	5.11 (1.42)
	Pre-nominal	0.64 (1.32)	1.79 (2.31)	0.58 (1.46)

Note. Shaded are unacceptable conditions.

8.4.2 Evaluation of the predictions for L2 acquisition of the target properties

The predictions formulated about the L2 acquisition of *sono* and NQs are restated in (8.8–8.9).

(8.8) FRH approach

- L1-Korean: *sono* > L1-English: *sono*
- L1-Korean: NQs > L1-English: NQs

(8.9) Cline of difficulty approach

L1-Korean: *sono* > L1-English: *sono* > L1-English: NQs > L1-Korean: NQs

Considering, first, the overt realisation of definiteness with *sono*, both the FRH and cline of difficulty approaches predicted that the L1 Korean group would be more target-like than the L1 English group. However, there was no concrete evidence to support this prediction. The statistical results suggested the two L2 groups did not significantly differ from each other but they both differed from the native control group (Table 8.4). Nevertheless, both L2 groups showed numerically target-like contrast between *sono* NPs and bare NPs within each context. Crucially, in the non-anaphoric contexts, both L2 groups rated *sono* NPs significantly lower than bare NPs (Table 8.5), which suggests knowledge of *sono*'s incompatibility with non-anaphoric contexts. As to the anaphoric and bridging contexts, the lack of clear native-like preference for *sono* NPs is still compatible with the predicted native pattern. Therefore, it seems that the two L2 groups have acquired overt anaphoric definiteness marking of *sono*, without a noticeable difference between them.

Turning to the covert definiteness constraint on NQs, the main statistical model results suggested that the two L2 groups were not significantly different from each other or from the native controls in terms sensitivity to the relevant property at the group level (Table 8.7). However, there was also suggestive evidence that the Korean group was more consistent in giving lower ratings for floating NQs than the English group— in post hoc tests, the effect of QUANTIFIER TYPE was significant in the [+definite] context for the Korean group but not for the English group (Table 8.8).[76] In this regard, the L1 Korean group performance is more native-like, which is in favour of the FRH approach (8.8) rather than the cline of difficulty approach (8.9).

Finally, the prediction based on the cline of difficulty that *sono* would be acquired earlier, or more easily, than NQs (8.9) seems to be partially supported. On the one hand, the English group showed clearer sensitivity to the unacceptability of *sono* NPs in the non-anaphoric context than to the unacceptability of floating NQs in the [+definite] context, as predicted. On the other hand, however, the Korean group did not show such a difference between *sono* and NQs, which is not supportive of the prediction.

8.5 Conclusion

The native judgement data presented in this chapter generally replicated Pilot ver. 2. That is, the relevant contrasts regarding the target properties of *sono* and NQs were more clearly observed than in Main study 1, suggesting that the suspected attenuated sensitivity to unacceptability (i.e., satiation) was real. The L2 data partially supported the predictions based on the FRH and the cline of difficulty, in different ways. In the next chapter, the L2 results of Main study 2 will be discussed further in terms of their theoretical implications

[76] However, recall that there was no three-way interaction to fully motivate running the post hoc tests.

as well as acquisition mechanisms of each property by learners from each L1 background.

Chapter 9: General discussion

9.1 Introduction

In this chapter, the main findings of the present study will be first summarised in light of the research questions and predictions, followed by a discussion of implications of the findings for the cline of difficulty. Then, L2 acquisition mechanisms will be considered for each target property by each L1 group. Finally, some implications for the research on L2 acquisition of definiteness expressions will be discussed.

9.2 Summary of the main findings

This thesis set out to address the research questions restated in (9.1–9.3) through an investigation of L2 acquisition of overt definiteness marking by the demonstrative *sono* and a covert definiteness constraint on NQs by Korean- and English-speaking Japanese learners.

(9.1) Research question 1:

> Is a covert feature expression more difficult to acquire than an overt feature expression?

(9.2) Research question 2:

> Does the necessity of feature reassembly make the acquisition task more difficult?

(9.3) Research question 3:

> In which situation is the acquisition of a covert feature expression less difficult, (i) when the L1 has a functional morpheme that realizes the feature overtly or (ii) when

the L1 has a corresponding covert expression?

To briefly recap the linguistic properties under investigation, the demonstrative *sono* can optionally mark definiteness (anaphoricity) overtly in anaphoric contexts (directly or via bridging) but cannot do so in non-anaphoric (unique definite) contexts, as shown in (9.4)–(9.6).

(9.4) Anaphoric

John gave me a book yesterday.

(Sono) hon-o moo yomimasita.

SONO book-ACC already read

'I have already read the book.'

(9.5) Bridging (anaphoric)

I bought a book yesterday.

(Sono) tyosya-wa furansu-jin da.

SONO author-TOP French is.

'The author is French.'

(9.6) Non-anaphoric

When I got outside last night, …

#Sono taiyoo-ga kagayai-teita.

SONO *sun*-NOM shine-ing.

'The sun was shining.'

As to NQs, floating NQs must be used in [−definite] contexts hence cannot be used in [+definite] contexts, whereas post-nominal NQs are acceptable either in [+definite] or [−definite] contexts, as exemplified in (9.7) and (9.8).

(9.7) [−definite] (post-nominal NQ vs. floating NQ)

Taroo does online shopping almost every day.

Kinoo-wa	{ hon san-satu-o vs. hon-o san-satu }	katta.
yesterday-TOP	{ book 3-CL-ACC vs. book-ACC 3-CL }	bought

'He bought three books yesterday.'

(9.8) [+definite] (post-nominal NQ vs. # floating NQ)

Taroo has two little sisters.

Sensyuu	{ imooto huta-ri-o vs. # imooto-o huta-ri }	yuuenti-ni
last.week	{ sister 2-CL-ACC vs. sister-ACC 2-CL }	amusement.park-to

tureteitta.

took

'He took the two sisters to an amusement park last week.'

Within the cline of difficulty, the definiteness marked by *sono* is overt because it is realised through a functional morpheme, whereas the indefiniteness of the floating NQ construction is considered covert in that it is expressed through a non-lexical means, namely word order.

Predictions for the acquisition of these properties were formulated from the perspectives of the FRH (Lardiere, 2008, 2009) and the cline of difficulty in feature acquisition (Cho & Slabakova, 2014; Slabakova, 2009), and tested by means of an AJT and an SPRT. Within the FRH, both overt and covert definiteness properties in question were predicted to be acquired more easily by Korean-speaking learners, who have equivalent L1 properties (hence no reassembly is necessary), than by English-speaking learners, whose L1 does not have such properties (hence reassembly is necessary based on the closest L1 properties). The cline of difficulty predicted the same acquisition order between the L2 groups as the FRH in terms the acquisition of *sono*. However, the cline of difficulty, assuming facilitation effects of overt feature realisation, predicted that the overt definiteness property of *sono* would be easier to acquire than the covert definiteness property of NQs for both L2 groups; and that English-speaking learners would acquire the covert property of NQs more easily than Korean-speaking learners due to facilitation from the L1 overt realisation of definiteness (i.e., English articles).

The SPRT data (from Main study 1) did not provide any statistical evidence of the relevant linguistic properties for any participant groups (including the native Japanese controls). However, the AJT data (from Main study 2, in particular) yielded the following findings:

(9.9) Finding 1:

> The English-speaking learners acquired the overt definiteness property of *sono* more easily than the covert definiteness property of NQs (in that they were more consistent in rating *sono* NPs lower than bare NPs in the non-anaphoric context than in rating floating NQs lower than post-nominal NQs in the definite context).

(9.10) Finding 2:

> The Korean-speaking learners acquired both properties equally easily (in that they robustly showed target rating patterns across the properties: lower ratings for *sono* NPs than bare NPs in the non-anaphoric context, and lower ratings for floating NQs than post-nominal NQs in the definite context).

(9.11) Finding 3:

The Korean-speaking and English-speaking learners acquired the overt property of *sono* equally easily (in that both L2 groups reliably distinguished NP types in the non-anaphoric context by rating *sono* NPs as less acceptable than bare NPs).

(9.12) Finding 4:

The Korean-speaking learners acquired the covert property of NQs more easily than English-speaking learners (in that the Korean-speaking group rated floating NQs lower than post-nominal NQs in the definite context more consistently than the English-speaking group).

Findings 1 and 2 suggest that the answer to Research question 1 is a qualified yes. The English speakers' data testifies to the overt vs. covert contrast subsumed under the cline of difficulty, whereas the Korean speakers' data does not, although not falsifying the prediction of the cline of difficulty, either. Similarly, given Findings 3 and 4, the answer to Research question 2 is conditionally affirmative. In terms of the overt property of *sono*, it does not seem to affect the outcome whether some feature reassembly is necessary (L1-English) or not (L1-Korean) (as detailed in Chapter 5); however, as to the covert property of NQs, the absence of feature reassembly necessity seems to give an advantage to the L1-Korean learners over the L1-English learners, who have some reassembly to do (i.e., from overt realisation through the article system in the L1 to covert realisation through word order change between NQ constructions in the L2). Finally, Finding 4 means that the second option is more likely to be the answer to Research question 3—having a corresponding covert expression in the L1 is more facilitative than having an L1 functional morpheme with which the relevant feature is overtly realised. This is in favour of the FRH but contra the cline of difficulty. In sum, as illustrated in Table 9.1, the overall results seem more compatible with the FRH than the cline of difficulty in that contradictory results were found only for the cline of difficulty, for Finding 4. The two accounts were similar otherwise: both found results that were supporting and not supporting (but not falsifying), in terms of Findings 1–3.

Table 9.1 Assessment of predictions for each approach

	FRH	Cline of difficulty
Finding 1 (English: *sono* > English: NQs)	✓	✓
Finding 2 (Korean: *sono* = Korean: NQs)	?	?
Finding 3 (Korean: *sono* = English: *sono*)	?	?
Finding 4 (Korean: NQs > English: NQs)	✓	*

Note. ✓ = supported, ? = neither supported nor falsified, * = falsified. "*x* > *y*" means *x* is acquired faster/more easily than *y*. "*x* = *y*" means that *x* and *y* are acquired at equivalent rates.

9.3 Implications for the cline of difficulty

Recall that according to Cho and Slabakova's (2014) cline of difficulty, an acquisition task where the L1 expresses the relevant feature overtly but the L2 expresses it covertly (i.e., F_{overt} *to* F_{covert}) is predicted to be easier than a task where both L1 and L2 express the feature covertly in the same way (i.e., F_{overt} *to* F_{covert}*, reassembly not required*). In the present study, the L1-English learners tackle the former, whereas the L1-Korean learners tackle the latter. Finding 4 (9.12) suggests that it may be the reassembly necessity rather than the overt vs. covert contrast of the feature realisation in the L2 that plays a bigger role in determining the relative task difficulty. In Cho and Slabakova, the acquisition of covert definiteness property in L2 Russian necessitated feature reassembly for both the English-speaking and Korean-speaking learners. The results suggested that the relevant acquisition task was easier for the English-speaking learners than the Korean-speaking learners, which corresponded to the contrast between F_{overt} *to* F_{covert} (the third hardest task) and F_{covert} *to* F_{covert}*, reassembly required* (the hardest task) on the cline of difficulty. This finding from Cho and Slabakova together with Finding 4 from the present thesis suggest that F_{covert} *to* F_{covert}*, reassembly not required* should be relocated towards the easier end than when feature reassembly is necessary for acquisition of a covert property. The cline of difficulty assumes that overt features can be acquired more easily than covert features, regardless of feature reassembly. However, based on the present findings, it seems empirically more plausible to predict that acquisition involving feature reassembly is always harder than acquisition without it. Nevertheless, within each category, overt features may still be easier to acquire than covert features, and L1-L2 correspondences are more facilitative for overt than covert feature realisation. Therefore, I propose a revised cline of difficulty in Figure 9.1. The key revision point is that whereas the original cline assumes a bigger role for overt vs. covert feature realisation than for reassembly requirement, this revised cline has the reversed assumption—reassembly requirement is more influential than overt vs. covert feature realisation.

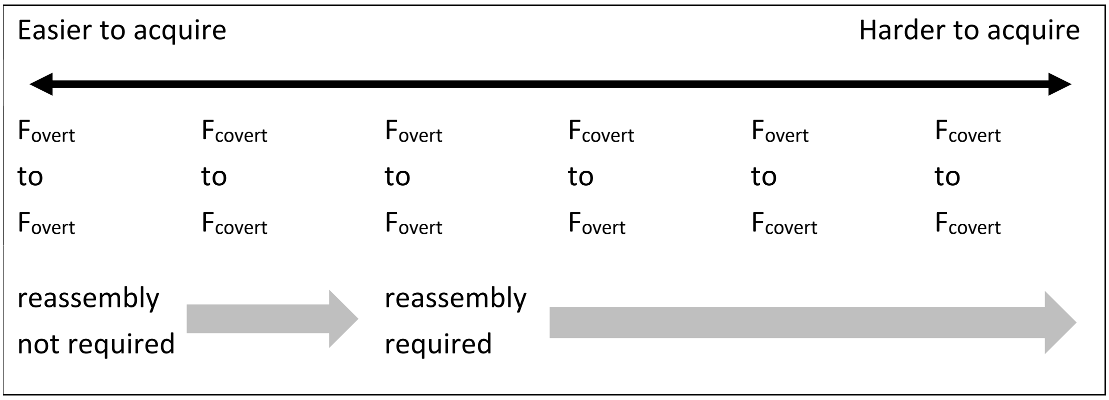


Figure 9.1 Revised Cline of Difficulty

Let us consider how compatible the present and previous findings are with the revised cline and in what aspects the cline is in need of empirical support. The relative difficulty of F_{overt} *to* F_{overt}*, reassembly required* compared to F_{overt} *to* F_{covert} on the scale is compatible with the L1-English group's apparently more successful acquisition of the overt property of *sono* than the covert property of NQs. However, the overt vs. covert contrast within the reassembly-not-necessary category has not been attested (this was true of Cho & Slabakova's original cline of difficulty too): the Korean-speaking group was equally successful in acquiring the relevant properties of both *sono* (overt) and NQs (covert) (Finding 2). Furthermore, the present study did not find evidence that a F_{covert} *to* F_{covert}*, reassembly not required* task is easier than F_{overt} *to* F_{overt}*, reassembly required*, either: it cannot be confidently said that the Korean-speaking learners' acquisition of the definiteness constraint was easier than the English-speaking learners' acquisition of *sono.* In order to assert that the F_{covert} *to* F_{covert}*, reassembly not required* task is easier than F_{overt} *to* F_{overt}*, reassembly required*, it must be demonstrated that the English-speaking learners are not sensitive to the property of *sono*, and at the same time, the Korean-speaking learners are sensitive to the definiteness constraint of floating NQs. However, because the English-speaking learners showed a clearly target-like judgement pattern for *sono* (significantly lower ratings of *sono* NPs than bare NPs in the non-anaphoric context), it does not seem clear whether the Korean speakers are more target-like in terms of the definiteness constraint than the English speakers are in terms of the overt definiteness marking of *sono*. Nonetheless, the revised cline in Figure 9.1 can be considered an improvement in the sense that it is compatible with the new findings from the present thesis and existing findings.

9.4 Acquisition mechanisms

I assume that the Korean-speaking learners of Japanese in the present study, given their target-like responses, acquired the definiteness properties of *sono* and NQs facilitated by full transfer of their representations of the corresponding L1 properties, as predicted by the FRH. However, there are questions to address regarding the L1-English learners' acquisition processes of the target properties, namely of (i) which scenario (presented in Chapter 5) seems the most probable in terms of their acquisition of *sono*; and (ii) how they could overcome the poverty-of-the stimulus problem predicted with their acquisition of the definiteness constraint on NQs.

9.4.1 Developmental stages of the L2 acquisition of *sono* by L1-English learners

Since the L1-English group in Main study 2 seems to have already acquired the property of *sono*, it is ultimately impossible to determine which developmental path postulated in Chapter 5 (previously in Table 5.2, repeated in Table 9.2) the English speakers went through. Additionally, because the learners are relatively advanced, all scenarios are, in principle, possible although some involve theoretically more complex learning tasks than others. Nevertheless, I speculate that the English-speaking learners initially mapped the feature of *that* ([+definite, +anaphoric, −bridging]) onto *sono* and undertook the reassembly task of adding [+bridging] (Scenario 2 or Scenario 3B(ii)), based on behaviour of the participants in Main study 1. Recall that the English-speaking participants in Main study 1 were clearly less proficient than their counterparts in Main study 2 in terms of the cloze test results. Moreover, they completed the AJT in an environment where their ability to distinguish between the acceptable and unacceptable sentences was potentially compromised (i.e., the satiation effect, discussed in Chapters 7 and 8). Nevertheless, they rated *sono* NPs statistically lower than bare NPs in non-anaphoric contexts (the target pattern), at levels comparable to the native control group and the significantly more proficient Korean-speaking group. Furthermore, they gave *sono* NPs higher ratings than bare NPs in bridging contexts, to a similar degree to the other two groups: an indication of their awareness of *sono* being possible in bridging contexts. This generally target-like judgement pattern suggests that even the less proficient English-speaking learners in Main study 1 had already successfully completed the relevant feature reassembly task (whichever scenario in Table 9.2 may hold true). Since feature reassembly tasks involving the restriction of feature distribution (i.e., [±anaphoric] => [−anaphoric] in Scenarios 1, 3A, and 3B (i)) typically require learners to have relatively advanced proficiency (e.g., Gabriele, 2009; Marsden, Whong, & Gil, 2018; Slabakova, 2006), the successful acquisition at the lower proficiency levels favours Scenarios 2 and 3B (ii), where the feature reassembly task is motivated by positive input, and hence easier (i.e., [−bridging] => [±bridging]).[77] Needless to say, however, further investigation with even less proficient learners is necessary to confirm this analysis.

[77] Conceptually, Scenario 2 seems even simpler than Scenario 3B (ii) (the easiest of all scenarios). That is, in Scenario 2, the initial mapping involves only one lexical entry, whereas in Scenario 3B (ii), learners first create two lexical entries, one for the features of *the* and one for the features of *that*, but discard the former at some point. However, it seems methodologically impossible to reliably tease these apart.

Table 9.2 Possible acquisition scenarios for L1-English learners

Scenario 1	<u>Initial mapping</u> L1 *the* = L2 *sono* [+definite, ±anaphoric, ±bridging]	<u>Feature reassembly</u> [±anaphoric] => [−anaphoric]
Scenario 2	<u>Initial mapping</u> L1 *that* = L2 *sono* [+definite, +anaphoric, −bridging]	<u>Feature reassembly</u> [−bridging] => [±bridging]
Scenario 3A	<u>Initial mapping</u> L1 *the/that* = L2 *sono* [+definite, ±anaphoric, ±bridging]	<u>Feature reassembly</u> [±anaphoric] => [−anaphoric]

Scenario 3B	<u>Initial mapping</u>
	(i) L1 *the* = L2 *sono*
	[+definite, ±anaphoric, ±bridging]
	
	(ii) L1 *that* = L2 *sono*
	[+definite, +anaphoric, −bridging]
		<u>Feature reassembly</u>
		If (i) is selected,
		[±anaphoric]
		=> [−anaphoric]
		
		If (ii) is selected,
		[−bridging]
		=> [±bridging]

9.4.2 How L1-English learners could overcome the poverty of stimulus regarding the definiteness constraint on NQs

Although the present results did not support Cho and Slabakova's (2014) cline of difficulty positioning of F_{overt} *to* F_{covert} as easier than F_{covert} *to* F_{covert}*, reassembly not required*, this does not necessarily mean an absence of the potential facilitative effect of overt realization of a feature in the L1 for acquisition of a covert L2 property involving that feature. That is, it is possible that the overt realization of definiteness in the L1 served as a heuristic to those English-speaking learners who successfully gave lower ratings to floating NQs than post-nominal NQs in the [+definite] context. One possibility is that they might have recognised nouns associated with floating NQs as corresponding to English DPs with the D coding [−definite]. Cho and Slabakova indeed propose a learning strategy akin to this for their English-speaking learners' acquisition of the covert expression of definiteness via word order in L2 Russian (Cho & Slabakova, 2014, p.183). Additionally, learners could avail themselves of distributional information about NQs and possible interpretations for each position, namely that floating NQs occur frequently in the [−definite] contexts but do not in the [+definite] context. Such information might have helped the English-speaking learners to inductively learn the definiteness constraint. However, this does not seem very feasible given the fact that other types of NQ (pre-nominal and post-nominal) are allowed in both [+definite] and [−definite] contexts. Following Boyd and Goldberg (2011, p. 55), statistical learning can occur in language acquisition when "an alternative formulation with the same function" is consistently witnessed in contexts where a pre-empted formulation is predicted to be appropriate. This suggests that in order to effectively pre-empt floating NQs in the [+definite] context, one specific type of NQ is required to be consistently witnessed instead of floating NQs. However, this condition is unlikely to be satisfied in the case of Japanese NQs. Thus, successful pre-emption of floating NQs seems unlikely in the [+definite] context.

An alternative account of the finding that the definiteness constraint on NQs seems acquirable for at least some English-speaking learners is the one employed by previous studies that reported successful acquisition of poverty-of-the-stimulus properties (e.g., Dekydtspotter et al. 2001; Okuma, 2019; Marsden, 2009), namely that the constraint itself does not have to be acquired based on input because it is made available through UG. In this view, the definiteness constraint on Japanese floating NQs is a result of the universal syntax-semantic computation of the floating construction rather than a purely lexical property of NQs alone. If this is true, the semantic constraint will be automatically activated in the L2 grammar, once learners acquire the essential properties of Japanese NQs, specifically that they must combine with classifiers, and that Japanese NQs can float, in contrast to English NQs. Note that such an account still predicts L1-Korean speakers to acquire the semantic constraint earlier than L1-English speakers, since Korean native speakers can transfer all relevant properties of floating NQs from their L1 counterparts, as soon as they identify the floating NQ structure in L2 Japanese input. English-speaking learners, on the other hand, may initially conceive of Japanese NQs as equivalents of English phrases consisting of a numeral and a classifier (e.g., *three flocks of birds*), which lack the ability to float. English-speaking learners must then at least adjust the L1-based representation of NQs towards the target, which would naturally require more time and effort to accomplish than the Korean-speaking learners' task, where no such adjustment, or reassembly, is required. I speculate that the key to the acquisition of the definiteness constraint on floating NQs could be the difference between Japanese NQs and English [numeral + classifier] structures (e.g., *three flocks*). Recall that Japanese classifiers are bound morphemes (Section 3.6.1). English classifiers such as *flocks* in *three flocks of birds*, on the other hand, are free morphemes. Kobuchi-Philip (2007) proposes that it is the morphological status of Japanese NQs as compound words (as opposed to phrases) that enables them to assume the floating quantifier position; by contrast, it is its phrase status that prevents a combination of [numeral + classifier] from floating in English.[78] If

[78] Kobuchi-Philip puts forward as evidence for the compound-word status of Japanese [Num + CL] combinations the fact that they exhibit word-internal voicing, *rendaku* (ia), which never occurs in the syntactic combination of two free morphemes (ib) (adapted from Kobuchi-Philip, 2007, p. 826).

(i) a. Num + CL: san + hon = san-bon '3-CL'
   b. Adj + N: omosiroi + hon = omosiroi hon 'interesting book'

this analysis is on the right track, English-speaking learners of Japanese must then acquire the compound-word morphology of Japanese classifiers in order that Japanese NQs can float would then arise automatically, and the definiteness constraint would ensue as a result of the universal syntax-semantic computation.

9.5 Implications for research on L2 acquisition of definiteness expressions

The present study offers some insights into L2 learnability of bridging and of definiteness properties as phenomena pertaining to linguistic-external interfaces. Recall that in Main study 2, the native Japanese controls showed a preference for *sono* NPs over bare NPs in the form of significantly higher ratings in bridging contexts, whereas L2 learners exhibited descriptively similar patterns yet there was no statistically significant distinction. Given the relatively advanced proficiency levels of the L2 groups, this suggests that ultimate attainment in bridging definiteness might pose a persistent problem in L2 acquisition, as observed in previous studies outlined in Chapter 4 (e.g., Cho, 2017; Feng, 2019). However, since the L2 groups did not show target-like preference for *sono* NPs in the non-bridging (direct) anaphoric context, either, then rather than a difficulty specific to bridging definiteness, it seems more likely that the kind of properties involved in discourse tracking for anaphoric reference may continue to be challenging even for advanced learners. This is along the lines of the Interface Hypothesis (e.g., Sorace, 2011; Sorace & Filiaci, 2006), which predicts persistent non-native-like variability in the L2 end-state with phenomena that necessitate a coordination of an internal component of the grammar (e.g., syntax, semantics), and an external component (i.e., pragmatics or discourse information). However, a question remains as to why the L2 groups were generally more sensitive to *sono*'s incompatibility with non-anaphoric contexts, than they were to *sono*'s felicity in anaphoric/bridging contexts; even though both non-anaphoric/unique definiteness and anaphoric/bridging definiteness concern the interface between internal components of the grammar (syntax-semantics) and an external component (pragmatics or discourse). That is, what could account for the learnability contrast between the properties that both involve internal-external interfaces, albeit with different configurations, (i) syntax-semantics-discourse (i.e., anaphoric and bridging contexts) and (ii) syntax-semantics-pragmatics (i.e., non-anaphoric contexts)?

One possibility is that the kind of definiteness established non-anaphorically (unique definiteness, or "out-of-the-blue definiteness" in Tuniyan's (2018) terms) may cause less difficulty to L2 learners than the kind that requires tracking discourse referents (anaphoric definiteness). This is presumably because unique definiteness can, in principle, be computed based on one's pragmatic knowledge (world or situational knowledge), independently of discoursal information, hence potentially less cognitively taxing. This suggests that not all external interfaces are equally problematic (e.g., White, 2011).

Alternative explanations for the contrast between non-anaphoric vs. anaphoric contexts can be made in terms of quality of input. *Sono*'s incompatibility with non-anaphoric contexts concerns a distinction in acceptability: *sono* NPs are unacceptable in such contexts whereas bare NPs are acceptable. However, in anaphoric contexts the distinction is about preference: *sono* NPs are preferred whereas bare NPs tend to be dispreferred but they are not unacceptable. That is, the former contrast is arguably more perceptible hence potentially easier to acquire than the latter. Furthermore, Japanese learners, in theory, will encounter only bare NPs, and not *sono* NPs, in non-anaphoric contexts. In anaphoric contexts, on the other hand, they will find more *sono* NPs than *bare* NPs yet will come across plenty of instances of bare NPs in those contexts, which may make it relatively difficult to rely on the input for form-meaning mappings. Note that these problems are not expected in L2 English. In English, the distributional contrast between *the* vs. *a* should be arguably more salient than the contrast between *sono* NPs vs. bare NPs in anaphoric contexts, which means more consistent input would be available. Consequently, L2 English learners are predicted to be more target-like in using *the* compared to L2 Japanese learners in using *sono* in anaphoric contexts. This is indeed compatible with the fact that the L2 Japanese learners in the present study seem more prone to deviate from the target patterns compared to the L2 learners in many previous L2 English article studies, at least in direct anaphoric contexts, where they perform in a generally target-like manner (i.e., accurately distinguishing *the* from *a*) (e.g., Cho, 2017; Feng, 2019; Ionin et al., 2004; Tuniyan, 2018).

It is of course premature to simply relegate the potential bridging problem in L2 Japanese to a matter of input. However, the present findings suggest that ambiguous input, in the sense of the optionality of anaphoric demonstratives in anaphoric contexts, might also play a role. This is what one needs to be cautious of in investigating L2 learnability of bridging in Japanese or other article-less languages such as Korean, because it could make it difficult to distinguish between the challenges posed by bridging and an ambiguous input effect. By contrast, in L2 English, for example, the definite article must be used in definite contexts, irrespective of whether in bridging or non-bridging contexts; therefore, the input for L2 English learners is considerably less ambiguous. This means that any differences found between bridging and non-bridging anaphoric contexts in L2 English can be effectively ascribed to other factors such as L1 transfer and external-interface status, where theoretically motivated. However, the potential adverse input effects are confounded with such factors in L2 Japanese, which makes it difficult to reliably confirm which factors are responsible for non-native-like behaviour. Nevertheless, based on the present findings, I tentatively conclude (i) that bridging contexts do not necessarily pose greater difficulty than non-bridging contexts as far as anaphoric contexts are concerned; and (ii) that external interface properties are not equally difficult: those related to pragmatics (unique definiteness in non-anaphoric contexts) seem less difficult than those related to discourse (anaphoric definiteness in direct and bridging contexts).

9.6 Summary of the chapter

In this chapter, some theoretical implications of the present findings were discussed for the cline of difficulty and L2 acquisition of definiteness properties. Firstly, a revision to Cho and Slabakova's (2014) cline of difficulty was proposed to accommodate the present findings as well as Cho and Slabakova's. For the learning processes of the target properties, I concluded that the Korean-speaking learners acquired both *sono* and NQs relatively easily, making use of their knowledge of the L1 corresponding properties, in line with the FRH. As to the English-speaking learners, I tentatively proposed that in acquiring *sono*, they took a route where they initially map the feature of *that* ([+definite, +anaphoric, −bridging]) onto *sono*, and subsequently add [+bridging] to complete the feature reassembly. Furthermore, I argued that some English-speaking learners may have acquired the definiteness constraint on NQs by means of the UG-guided syntax-semantics computation. Lastly, some potential input problems concerning L2 acquisition of overt definiteness marking by *sono* were identified as potential sources of challenge that are independent of other factors including L1 transfer and difficulty related to some external interface properties.

Chapter 10: Conclusion

10.1 Contributions

This thesis makes three key contributions: to second language acquisition theory, to L2 acquisition of definiteness, and to theoretical linguistic research on Japanese numeral quantifiers (NQs), respectively. Starting with L2 acquisition theory, the thesis offers a refinement to our understanding of what causes difficulty in acquiring functional features. This study was designed to tease apart the impact of overt vs. covert feature realisation and the reassembly vs. no-reassembly requirement. Particularly, the comparison of L1-English and L1-Korean learners' acceptability judgements in terms of the acquisition of the definiteness constraint on floating NQs provided crucial data to evaluate the untested assumption of Cho and Slabakova's (2014) cline of difficulty; namely, that overt vs. covert feature realisation plays a larger part than the need for feature reassembly in determining acquisition task difficulty. The results suggested that, at least for the acquisition of relationship between definiteness and floating NQ, the opposite seems more likely the case. This motived a proposed revision to the cline that can accommodate both previous and present results. The next section will detail how this opens up new directions for future research into learnability problems and the role of L1 in the non-native language acquisition of syntax-semantics mappings.

The second contribution concerns the originality of the focus on definiteness in L2 Japanese. While there has been a lot of research on L2 acquisition of the article system in English, L2 acquisition of definiteness properties in other languages, and particularly so-called article-less languages is rare. However, investigation in this domain is potentially informative in L2 research. For example, recall that the L2 Japanese learners, unlike the native controls, did not show a statistically reliable preference for overt definiteness marking using the demonstrative *sono* over covert marking through bare nominals in anaphoric contexts (bridging or non-bridging). This kind of challenge, involving the L2 acquisition of a native-like preference for optional overt definiteness-marking, seems to be a phenomenon that might offer new insights into L2 ultimate attainment, which could only be examined in article-less languages. On the other hand, the definiteness constraint on floating NQs can be considered a novel interesting poverty-of-the-stimulus (POS) phenomenon for learners whose L1 lacks the constraint (e.g., English). Specifically, the successful acquisition of the constraint by at least some English-speaking learners implies that L2 acquisition is guided by UG. These showcase the utility of L2 research into definiteness expressions in article-less languages such as Japanese and Korean.

The third contribution is directed towards theoretical Japanese linguistic research. The Japanese properties under investigation are ones for which varied informal native intuitions can be found in the literature but which, prior to the present thesis, had not been tested formally through experimental research. Particularly, the relatively clear native judgement data regarding the semantic constraint on Japanese NQs from Pilot AJT ver. 1 played a crucial role in determining the most valid characterisation of the constraint out of the three competing theoretical proposals: the results provided solid evidence for the proposal that the floating NQs are adverbials and the relevant constraint concerns definiteness rather than specificity. The present thesis makes a convincing case for the utility of experimental work in evaluating linguistic theories.

10.2 Limitations & directions for future research

For a wider applicability of the revised cline of difficulty (Figure 9.1), further investigation is necessary to examine whether a feature covertly realised in an identical way in the L1 and the L2 (*F_{covert} to F_{covert} no reassembly necessary*) is easier than when that feature is expressed overtly in both languages but some feature reassembly is required to achieve the target-like configuration (*F_{overt} to F_{overt} reassembly necessary*). Recall that the present study involved this contrast (L1-Korean learners' acquisition of the definiteness constraint on NQs vs. L1-English learners' acquisition of the definiteness marking by s*ono*) but did not gain evidence either supporting or falsifying the prediction. Thus, further investigation is required to test whether this prediction is born out. Additionally, another study would be appropriate with its focus on the contrast between the cases where the L1 and the L2 have corresponding covert properties vs. where the L1 does not have that covert property but an overt means to express it (e.g., acquisition of the definiteness constraint on NQs by L1-Korean learners vs. by L1-English learners) for the following reason. In the present study, the POS problem was implicated for one group (L1-English learners) but not for the other (L1-Korean learners). Although this learnability problem can seem to be overcome at least for some advanced learners, it could be an independent source of difficulty, which would then be a confounding factor in the investigation of the relation between the effects of overt/covert feature realisation and feature reassembly necessity. Therefore, it seems ideal to test the prediction with L1-L2 combinations that involve no POS problem for either L2 group.

As to the acquisition of *sono*, I tentatively concluded that the potential learnability problem regarding bridging definiteness proposed in the previous research (e.g., Cho, 2017) did not seem to affect the participants in the present study. However, this remains inconclusive, particularly because non-anaphoric bridging was not tested in this study (due to the apparently considerable challenge that acquisition of non-anaphoric bridging represents. See Chapter 4, Section 4.5). If the input problems surrounding the optional nature of *sono* in anaphoric contexts are real, the potential residual optionality with bridging definiteness in article-less L2s would be more reliably studied within non-anaphoric contexts, where anaphoric-marking demonstratives are not allowed and NPs must always be bare (hence more consistent input). Specifically, a comparison of non-anaphoric non-bridging definite contexts (i.e., out-of-the-blue definite contexts) with non-anaphoric bridging definite contexts would be informative for further investigation of the potentially challenging nature of bridging reference in L2.

There were several limitations regarding the samples in the present thesis. Firstly, as pointed out in Chapter 7, the sample size of the L1-Korean group (n = 12) in Main study 1 was smaller than planned after part of the data collection had to be cancelled due to COVID-19. This inevitably resulted in a lack of statistical power, particularly for the SPRT. Thus, increasing the numbers of participants and test items might make the target effects and between-groups differences more visible. Furthermore, additional AJT data from lower proficiency learners would be useful particularly to investigate the developmental stages for the acquisition of *sono* in terms of English-speaking learners' initial feature mapping. If English-speaking learners initially map the features of *that* ([+definite, +anaphoric, −bridging]) onto *sono* (as proposed in the previous chapter), they are expected to go through a phase where they accept *sono* in direct anaphoric contexts and reject it correctly in non-anaphoric contexts but infelicitously in bridging anaphoric contexts. In contrast, Korean-speaking counterparts with matched proficiency will accept *sono* in both anaphoric and non-bridging contexts while rejecting it in non-anaphoric contexts (target pattern), by virtue of positive transfer of the features of their L1 demonstrative *ku* ([+definite, +anaphoric, ±bridging]).

Finally, let me point out a limitation about the real-time/online data collection methodology. Although the self-paced reading method, the choice of the present study, failed to obtain evidence for the definiteness properties of *sono* and NQs, other types of real-time measure such as event-related potentials (ERPs) and eye tracking still seem to have a chance to do so. Indeed, as demonstrated by Jiang (2018), target-like online performance has been observed previously more often through ERP responses and eye-tracking than self-paced reading. Jiang explains that self-paced reading relies on a physical responses that is extraneous to language processing (i.e., key/button pressing), ERP data (i.e., involuntary changes in electrical voltage in the brain) and eye movements reflect participants' cognitive processing arguably more directly. Therefore, use of these alternative measures should be considered for future research into the definiteness properties investigated in this study, for their potentially keener online sensitivity. Nevertheless, the null results of the thesis regarding the Japanese definiteness phenomena makes an interesting contrast to recent findings from self-paced reading studies, in which both L1 and L2 English speakers were sensitive to the distinction between the English articles *a* and *the* (e.g., Cho, 2020; Ionin et al., 2019). It might be attributed to the difference that the phenomena are fairly categorical in English (i.e., articles are elements that affect grammaticality), whereas the phenomena that the present thesis investigates are more gradient (i.e., the presence/absence of *sono* and NQs does not affect grammaticality in Japanese). It warrants further research into why self-paced reading is sensitive to one definiteness property but not to another.

10.3 A final remark

I hope that the present thesis has shown that exploration of definiteness properties in article-less languages is informative not just because it enriches the scope of L2 acquisition research but also because it can provide us with a good testing ground for timely L2 theoretical proposals (e.g., the FRH and the cline of difficulty). My hope is that this research will invite further investigation of the L2 acquisition of definiteness in article-less languages. There still seem to be many missing pieces to the L2 acquisition puzzle, particularly concerning L1 transfer and learnability, which require us to consider not only linguistic factors, but also factors relating to input and methodological issues. The present thesis has aimed to offer a new piece to fit into the puzzle.

Appendices

Appendix 1: Consent form & information sheet

Note. Participants read the consent form and information sheet in their own native languages (i.e., Japanese, English, or Korean). Presented here are English versions of those documents given to L2 learners in Main study 1. This consent form and information sheet are representative of those used for all parts of the data collection.

Appendix 1A: Sample consent form (English ver.):

UNIVERSITY *of York*

DEPARTMENT OF LANGUAGE & LINGUISTIC SCIENCE
Heslington, York, YO10 5DD
www.york.ac.uk/language

Consent form

Title of study: reading comprehension of Japanese sentence by native speakers and learners of Japanese

This form is for you to state whether or not you agree to take part in the study. Once you have read and checked every item below, please sign your name along with the date and hand the form in to the researcher. If there is anything you do not understand, or if you want more information, please ask the researcher.

I have read and understood the information sheet about the study.	Yes ☐	No ☐
I have had an opportunity to ask questions about the study.	Yes ☐	No ☐
I understand that the information you provide will be held in confidence by the researcher.	Yes ☐	No ☐
I understand that I may withdraw from the study at any time before I leave the venue without giving any reason.	Yes ☐	No ☐
I understand that the information I provide may be used in future research on language.	Yes ☐	No ☐
Do you agree to take part in the study?	Yes ☐	No ☐

Your signature: ____________________

Researcher's name: Keisuke Kume

Date: ____________________

Appendix 1B: Sample information sheet (English ver.):

Reading comprehension of Japanese sentences by native speakers and learners of Japanese

My name is Keisuke Kume. I am a PhD student in the Department of Language and Linguistic Science, University of York. Thank you for your interest in participating in this research. This document is an information sheet about the research. Please take a moment to read through the following information about the study first. If you decide to participate, you will be asked to sign a consent form.

Information about the study

What is the study about?

This study examines how a range of Japanese sentences are comprehended by native Japanese speakers, and Japanese learners whose first language is English or Korean.

Who can participate?

You can participate in this study if you are (i) 18 years old or over and a native speaker of Japanese, or (ii) 18 years old or over and a learner of Japanese whose first language is English / Korean (but no other languages).

What does the study involve?

This study consists of four tasks. The content and estimated task completion time for each task is as follows:

(1) Self-paced reading task

You will read at your own natural pace a range of Japanese sentences displayed on the computer screen. Specifically, for each item, you read a short passage providing context, and then a continuation sentence. The continuation sentence is presented segment by segment. For some items, you will be asked to answer a question about the content of the sentences you have just read. This task will take about 40–50 minutes.

(2) Acceptability judgement task

You will read a series of contexts and sentences similar to those in the first task, and will judge the acceptability of the sentences. You will read a context and the underlined sentence that follows it. Then you will rate how natural the underlined sentence is as a continuation of the passage proving the context using a scale of 0(completely odd) – 6(completely natural). The estimated task completion time is about 35–45 minutes.

(3) Fill-in-the-blank task

You will read some Japanese passages with blanks and fill in each blank with an appropriate word by choosing from four options. This task will take 10–15 minutes.

(4) Participant questionnaire

This questionnaire asks you about your language background. It will take 5–10 minutes.

* Total estimated completion time

90–120 minutes

Do I have to take part?

No, you do not have to take part in the study. Even after you have agreed to participate, you will be free to withdraw from the study without giving a reason.

However, note that your data cannot be withdrawn after you have left the experiment venue as it will be impossible destroy your data. This is because your responses will be recorded anonymously to ensure you remain personally unidentifiable.

What are the possible risks of taking part?

There are no foreseeable risks to taking part.

Are there any benefits to participating?

As a token of gratitude for participation, 20 pounds will be given to each participant who has completed all the tasks. Furthermore, your responses will contribute to a better understanding of how Japanese sentences are.

What will happen to the data I provide?

The data you provide will be used together with the data of other participants to examine how a range of Japanese sentences are understood by native speakers of Japanese and learners of Japanese. The electronic data will be stored in a password-protected folder in the University of York secure server and the paper-form data will be stored in a locked storage cabinet in the Language & Linguistic Science department building in the University of York.

<u>What about confidentiality?</u>

Your responses to the survey will remain strictly confidential. No information will be collected that could be used to identify a participant personally. It will not, therefore, be possible to identify any individuals or their individual responses in publications or presentations.

※ This study has been reviewed and approved by the Departmental Ethics Committee of the Department of Language and Linguistic Science at the University of York. If you have any questions, please feel free to ask the researcher by email at the following address.

[Contact information of researcher]

Keisuke Kume (PhD student at Department of Language and Linguistics Science, University of York)

Email: kk958@york.ac.uk

Appendix 2: Language background questionnaire

Note. The English translation is provided for reference. The actual questionnaire was presented without it.

参加者アンケート Participant questionnaire

Q1. 年齢:(　　)歳

Age: (　　) years old

Q2. 性別: a. 男　　b. 女　　(答えに○をつけてください。)

Sex: a. Male　b. Female　(please circle your answer)

あなたの言語背景について教えてください。

Please tell me about your language background.

Q3. あなたの母語(生まれてから最初に自然に覚えた言語)は何ですか。次の a～d の中から選んで、○をつけてください。d を選んだ場合は、(　　)に答えを書いてください。(複数回答可)

What is your mother tongue (the first language you naturally learned)? Please chose your answer from a-d by circling it. If you chose d, please specify your answer in (　　).

a. 日本語　　b. 韓国語　　c. 英語　　d. その他(　　　　　　　　　　)

a. Japanese　b. Korean　c. English　d. Other (　　　　　　　　　　)

Q4. あなたのご両親の母語は何ですか。次の a～d の中から選んで、○をつけてください。d を選んだ場合は、(　　)に答えを書いてください。(複数回答可)

What languages do your parents speak as their mother tongue? Please chose your answer from a-d by circling it. If you chose d, please specify your answer in (　　). (Multiple answers possible)

父:a. 日本語 b. 韓国語 c. 英語 d. その他(　　　　　　　　　　)

Father: a. Japanese　b. Korean　c. English　d. Other (　　　　　　　　　　)'

母:a. 日本語 b. 韓国語 c. 英語 d. その他(　　　　　　　　　　)

Mother: a. Japanese　b. Korean　c. English　d. Other (　　　　　　　　　　)'

Go on to the next page

Q5. 子供のころに使っていた言語を教えてください。次の a～d の中から選んで、〇をつけてください。d を選んだ場合は、（　　）に答えを書いてください。（複数回答可）

What languages did you use as a child? Please chose your answer from a-d by circling it. When you chose d, please specify your answer in (　　). (Multiple answers possible)

家で：a. 日本語 b. 韓国語 c. 英語 d. その他（　　　　　　　　　　）
At home: a. Japanese　b. Korean　c. English　d. Other (　　　　　　　　　　)

学校で：a. 日本語 b. 韓国語 c. 英語 d. その他（　　　　　　　　　　）
At school: a. Japanese　b. Korean　c. English　d. Other (　　　　　　　　　　)

友達と：a. 日本語 b. 韓国語 c. 英語 d. その他（
With friends: a. Japanese　b. Korean　c. English　d. Other (　　　　　　　　　　)

あなたの現在までの日本語学習について教えてください（Q6～Q8 は日本語学習者の方のみお答えください）。

Please tell me about your Japanese learning history (Q6 ~ Q8 are only for Japanese learners).

Q6. 何歳で日本語を勉強し始めましたか。（　　　）歳

At what age did you started learning Japanese? (　　　) years old

Q7. どのくらいの期間日本語を勉強していますか。（　　）年（　　）カ月

How long have you been learning Japanese? (　　　) years (　　　) months

Q8. 日本に住んだことはありますか。a. はい　　b. いいえ

Have you ever lived in Japan?　a. Yes　b. No

「はい」と答えた場合は、日本での滞在期間について教えてください。
If you answered yes, please specify how long.

（　　　年　　　月）～（　　　年　　　月）
(Year:　　　Month:　　　) - (Year:　　　Month:　　　)

Go on to the next page

母語と日本語以外の言語を話しますか。もし話す場合は、自己評価でどの程度その言語を話せるかを教えてください。

Do you speak any other languages than your mother tongue and Japanese? If yes, please self-evaluate your skills in each language.

言語 Language　　　　自己評価 Self-evaluation

(　　　　　)語　　a. 初級　b. 中級　c. 上級　d. ネイティブ並み

Language:　　a. Beginner　b. Intermediate　c. Advanced　d. Near-native

(　　　　)

(　　　　　)語　　a. 初級　b. 中級　c. 上級　d. ネイティブ並み

Language:　　a. Beginner　b. Intermediate　c. Advanced　d. Near-native

(　　　　)

(　　　　　)語　　a. 初級　b. 中級　c. 上級　d. ネイティブ並み

Language:　　a. Beginner　b. Intermediate　c. Advanced　d. Near-native

(　　　　)

(　　　　　)語　　a. 初級　b. 中級　c. 上級　d. ネイティブ並み

Language:　　a. Beginner　b. Intermediate　c. Advanced　d. Near-native

(　　　　)

(　　　　　)語　　a. 初級　b. 中級　c. 上級　d. ネイティブ並み

Language:　　a. Beginner　b. Intermediate　c. Advanced　d. Near-native

(　　　　)

Q9. その他何でもあなたの言語背景について知っておくべきことがあれば、教えてください。

Please tell me anything else that I should know about your language background.

End of the questionnaire

Appendix 3: Proficiency task

Appendix 3A: Proficiency test & choices with correct answers

Note. See Appendix 3B for an English translation of the passage below.

穴埋めタスク

空白(______)に入る最も適切なことばを選択肢の中から選んでください。

「ルームシェア」

マンションやアパートの一部屋を家族や兄弟とではなく、友人同士などで借りて共同生活をすることを「ルームシェア」という。欧米では一般的だ(1)________、日本でも最近、都市部の(2)________の間で広まっている。「家賃が(3)________できる」「楽しい」「安心」など、理由は(4)________だが、海外留学などで経験(5)________ルームシェアの利点を知る人(6)________増え、他人と生活すること(7)________の抵抗感がなくなってきて(8)________という背景もあるよう(9)________。

友達と一緒に暮らす

F さん(27歳・女性)は中国で、香港(10)________の女性とルームシェアを経験(11)________。現在も都内で大学時代(12)________同級生の女性2 人と 2LDK(13)________部屋をルームシェアしている。家賃12万(14)________は 3 人で 4万円ずつ(15)________しているという。電気、水道、(16)________などの公共料金と食費(17)________、3 人共通の財布を用意(18)________、毎月1 人 3万円ずつ入れ、(19)________から支払う。住みはじめてからの(20)________を聞いてみると、一番の(21)________はやはり金銭面。都内で 1 人(22)________住むには 6万~8万円かかる(23)________、今はその半分。公共料金(24)________は 1 人で負担するより(25)________安くてすむ。一方、デメリットは、長電話(26)________しにくいことだという。

インターネットで「シェアメート」を探す

(27)________をインターネットを通して探す人(28)________増えている。また、「一緒に暮らして(29)________を覚えたい」という理由から、(30)________に外国人を希望する日本人(31)________多い。「国際交流協会」はシェアメート(32)________探す人たちの情報交換の(33)________を設けようと、昨年8 月に(34)________を立ち上げた(http://borderless-tokyo.com)。シェアメート募集の(35)________ には、日本人、外国人から多数(36)________書き込みがある。ホームページの管理者、近藤誠二(37)________によると、最近はアクセス(38)________が多いときは 1 日に 200(39)________になるという。だが、(40)________一般には、大家さんの(41)________が得られず、シェアを受け入れてくれる(42)________は少ないそうだ。

"✓" represents the correct answer.

(1) a. から
b. が ✓
c. って
d. と

(2) a. 若者(わかもの) ✓
b. 高齢者(こうれいしゃ)
c. 外国人(がいこくじん)
d. 富裕層(ふゆうそう)

(3) a. 貯金(ちょきん)
b. 改善(かいぜん)
c. 節約(せつやく) ✓
d. 検討(けんとう)

(4) a. 様々(さまざま) ✓
b. 単純(たんじゅん)
c. 不明(ふめい)
d. 調査中(ちょうさちゅう)

(5) a. せず
b. 中の
c. から
d. して ✓

(6) a. が ✓
b. は
c. を
d. に

(7) a. との
b. へ ✓
c. に
d. が

(8) a. しまう
b. いる ✓
c. ある
d. ほしい

(9) a. ね
b. だが
c. だ ✓
d. です

(10) a. 出身(しゅっしん) ✓
b. 製(せい)
c. 在住(ざいじゅう)
d. 滞在中(たいざいちゅう)

(11) a. しない
b. したい
c. 中だ
d. した ✓

(12) a. の ✓
b. は
c. が
d. に

(13) a. を
b. と
c. の ✓
d. な

(14) a. 元(げん)
b. ポンド
c. 円(えん) ✓
d. ドル

(15) a. 負担(ふたん) ✓
b. 節約(せつやく)
c. 貯金(ちょきん)
d. 計算(けいさん)

(16) a. 固定電話(こていでんわ) ✓
b. 食費(しょくひ)
c. 保険料(ほけんりょう)
d. 家賃(やちん)

(17) a. に
b. は ✓
c. が
d. と

(18) a. しつつ
b. しながら
c. すると
d. して ✓

(19) a. どこ
b. そこ ✓
c. あそこ
d. ここ

(20) a. 理由(りゆう)
b. 感想(かんそう) ✓
c. 不満(ふまん)
d. 経験(けいけん)

(21) a. メリット ✓
b. 利益(りえき)
c. デメリット
d. 心配事(しんぱいごと)

(22) a. に
b. と
c. で ✓
d. が

(23) a. が ✓
b. として
c. なら
d. ため

(24) a. のみ
b. だけ
c. など ✓
d. から

(25) a. きっと
b. ずっと ✓
c. そっと
d もっと

(26) a. は
b. が ✓
c. で
d. も

(27) a. シェアメート ✓
b. シェアハウス
c. アパート
d. マンション

(28) a. は
b. なら
c. も ✓
d. など

(29) a. 日本語（にほんご）
b. 外国語（がいこくご） ✓
c. 異文化（いぶんか）
d. 英語（えいご）

(30) a. 最初（さいしょ）
b. 顧客（こきゃく）
c. 家族（かぞく）
d. 相手（あいて） ✓

(31) a. で
b. も ✓
c. に
d. かなり

(32) a. を ✓
b. と
c. が
d. は

(33) a. 動機（どうき）
b. 理由（りゆう）
c. 手段（しゅだん）
d. 場（ば） ✓

(34) a. 不動産屋（ふどうさんや）
b. インターネット
c. マンション
d. ウェブサイト ✓

(35) a. 看板（かんばん）
b. 宣伝（せんでん）
c. 広告（こうこく）
d. 掲示板（けいじばん） ✓

(36) a. に
b. が
c. の ✓
d. から

(37) a. さん ✓
b. くん
c. 殿（どの）
d. 様（さま）

(38) a. 問（と）い合（あ）わせ
b. 中（ちゅう）
c. 状況（じょうきょう）
d. 数（すう） ✓

(39) a. 近（ちか）く ✓
b. 遠（とお）く
c. ごとく
d. 人（にん）

(40) a. たしかに
b. おそらく
c. まだ ✓
d. すでに

(41) a. 資金（しきん）
b. 理解（りかい） ✓
c. 権利（けんり）
d. 好感（こうかん）

(42) a. 雇用主（こようぬし）
b. 可能性（かのうせい）
c. 状況（じょうきょう）
d. ところ ✓

Appendix 3B: English translation of the proficiency test passage:

Note. This English translation of the cloze task is presented here for reference. It was taken from Marsden (2005, p. 324), which was originally adapted from *Nihongo Journal* (2000.6, pp. 46–47). It was not presented in the actual test.

Room-sharing

'Room-sharing' means renting a condominium or apartment communally with friends and acquaintances instead of living with family members or siblings. This is a common practice in Europe and North America and now it is also starting to become popular among young people in Japanese cities. They like room-sharing for a variety of reasons: it lets you save money on rent, it's fun, and it provides a sense of security. More and more people are learning about these advantages by experiencing room-sharing while studying abroad, and this has lowered resistance to the idea.

Living with friends

Ms F (27) once shared a room with a woman from Hong Kong while living in China. Now, she shares a two-bedroom apartment in Tokyo with two friends from her university days. Each of the three friends pays 40,000 yen to cover the rent of 120,000 yen. They each also put 30,000 yen into a common purse every month to buy food and pay the electricity, water, telephone and other utility bills. After beginning this shared living arrangement, Ms F said that the cost-of-living savings were the biggest advantage. The rent on an apartment for a single person in Tokyo is 60,000 to 80,000 yen, and now she's only paying half of that. The utilities are also much lower than she would have to pay living alone. One disadvantage she mentioned was that she usually can't talk for a long time on the telephone.

Finding roommates on the Internet

An increasing number of people are finding roommates on the Internet. Many Japanese people want to find a foreign roommate so they can learn a foreign language while sharing accommodation. Last August, borderless-tokyo-corp (an international exchange association) established a website where people looking for roommates can exchange information (http://borderless-tokyo.com). The bulletin board for people seeking roommates is full of ads from both Japanese and foreigners. According to Seiji Kondo, who manages the website, the site sometimes gets as many as 200 hits a day. However, Kondo mentioned that landlords are resistant to the idea of room-sharing, and places that allow it can be few and far between.

Appendix 4: Pilot data

Appendix 4A: Pilot AJT ver. 1

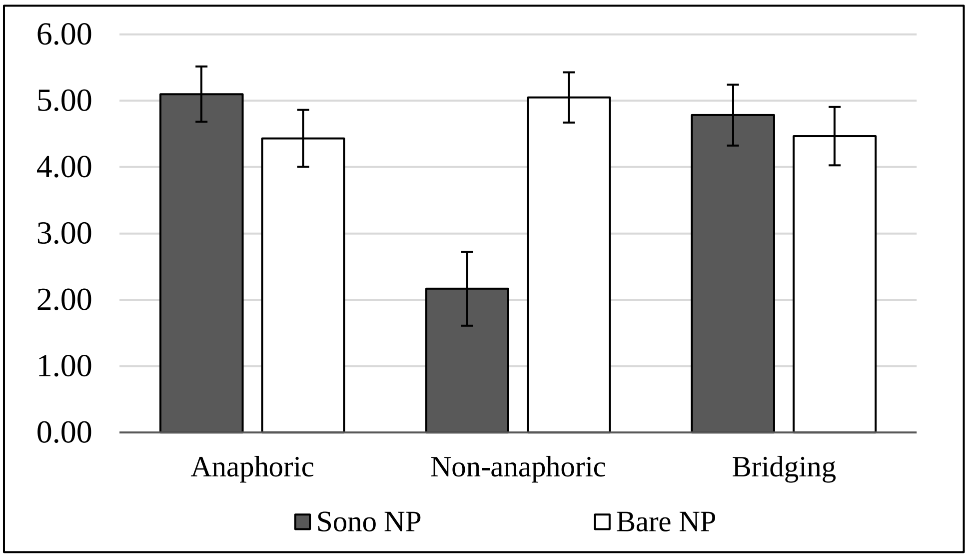


Figure Appx.1 Mean acceptability ratings for sono *items: Pilot AJT ver. 1 (error bars = SE)*

Appendix 4B: Pilot SPRT ver. 1

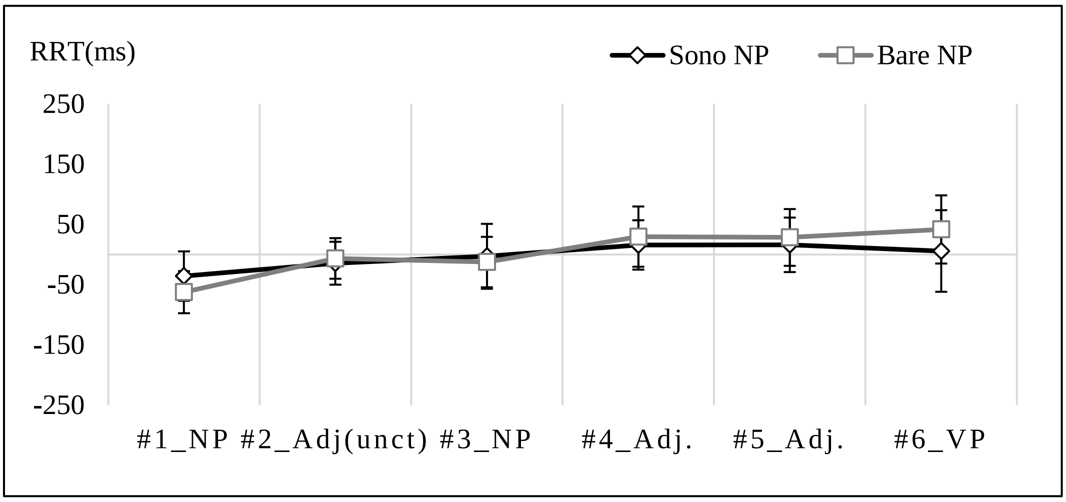


Figure Appx.2 Mean RRTs for sono *items in the anaphoric context: Pilot SPRT ver. 1 (error bars = SE)*

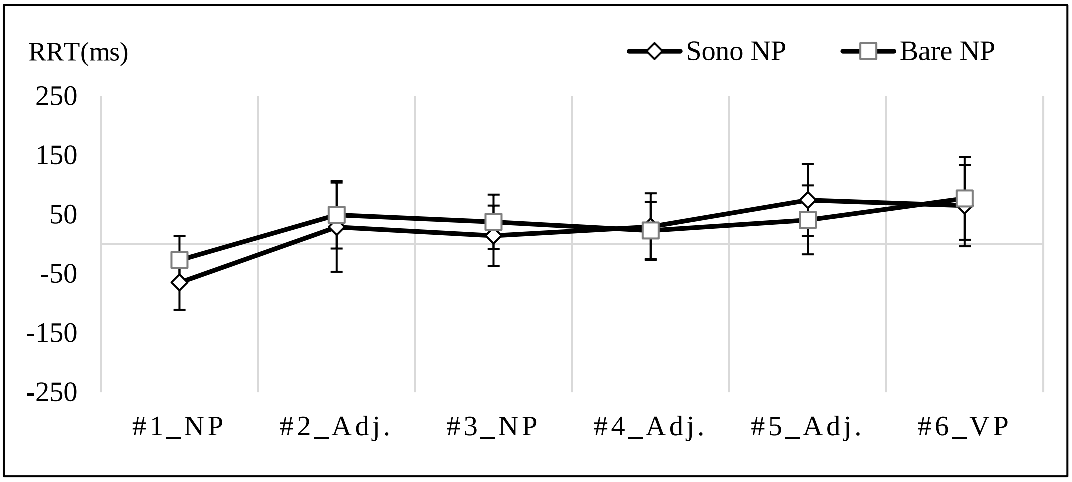


Figure Appx.3 Mean RRTs for sono *items in the non-anaphoric context: Pilot SPRT ver.1 (error bars = SE)*

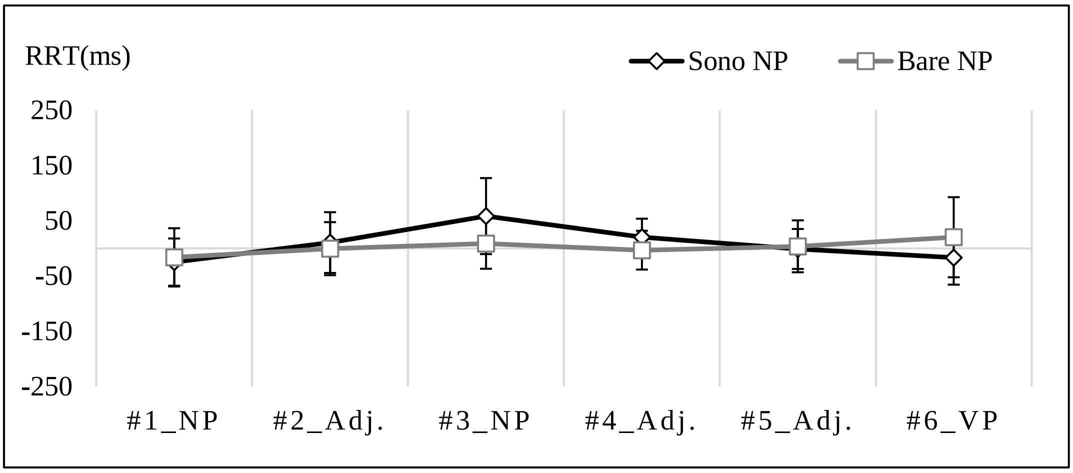


Figure Appx.4 Mean RRTs for sono *items in the bridging context: Pilot SPRT ver. 1 (error bars = SE)*

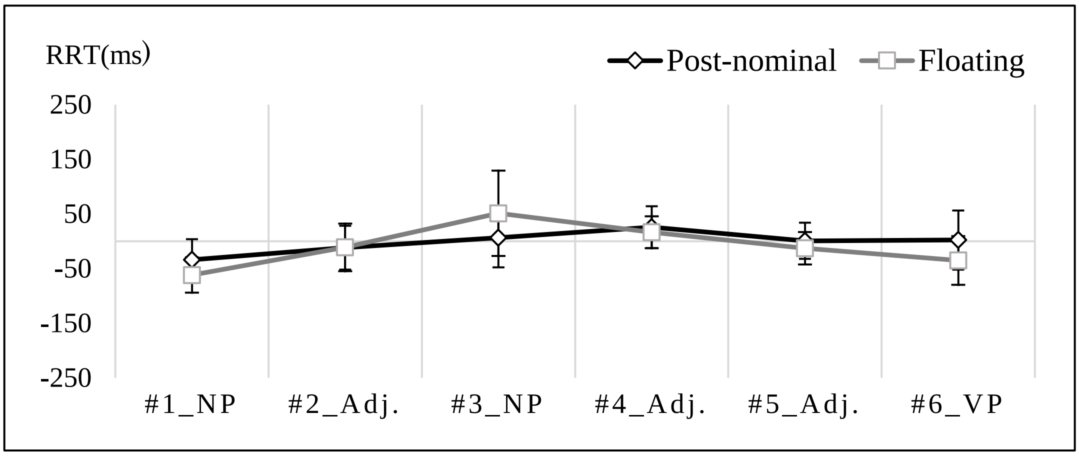


Figure Appx.5 Mean RRTs for NQ constructions items in the [+definite] context: Pilot SPRT ver. 1 (error bars = SE)

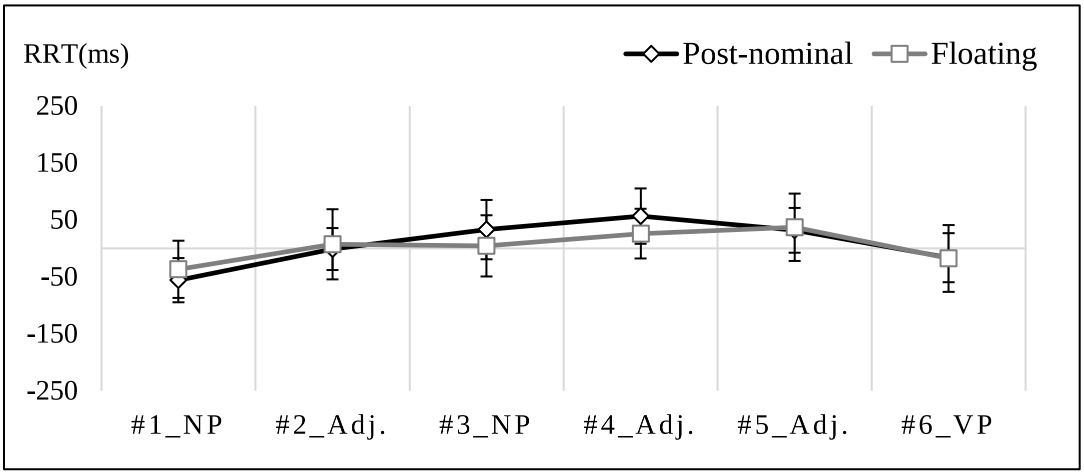


Figure Appx.6 Mean RRTs for NQ constructions items in the [−definite] context: Pilot SPRT ver. 1 (error bars = SE)

Appendix 4C: Pilot SPRT ver. 2

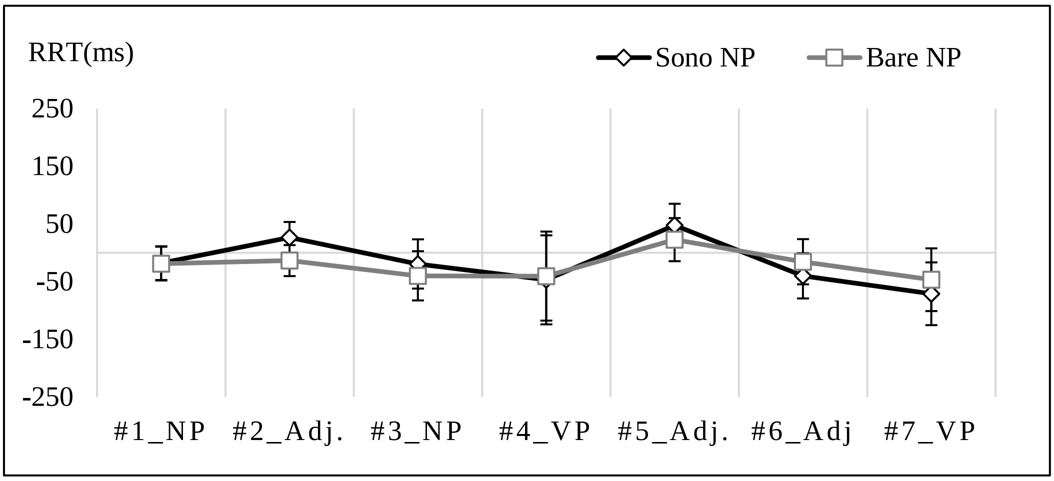


Figure Appx.7 Mean RRTs for sono *items in the anaphoric context: Pilot SPRT ver. 2 (error bars = SE)*

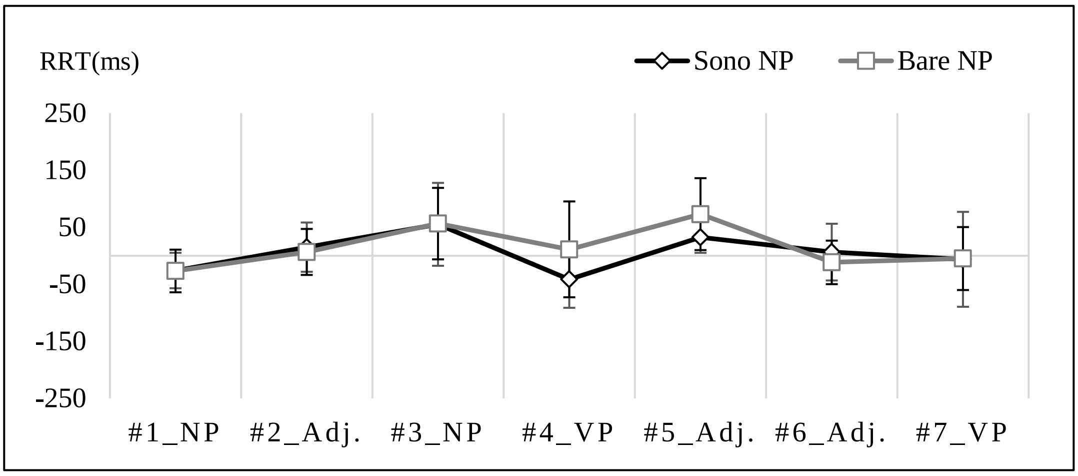


Figure Appx.8 Mean RRTs for sono *items in the non-anaphoric context: Pilot SPRT ver. 2 (error bars = SE)*

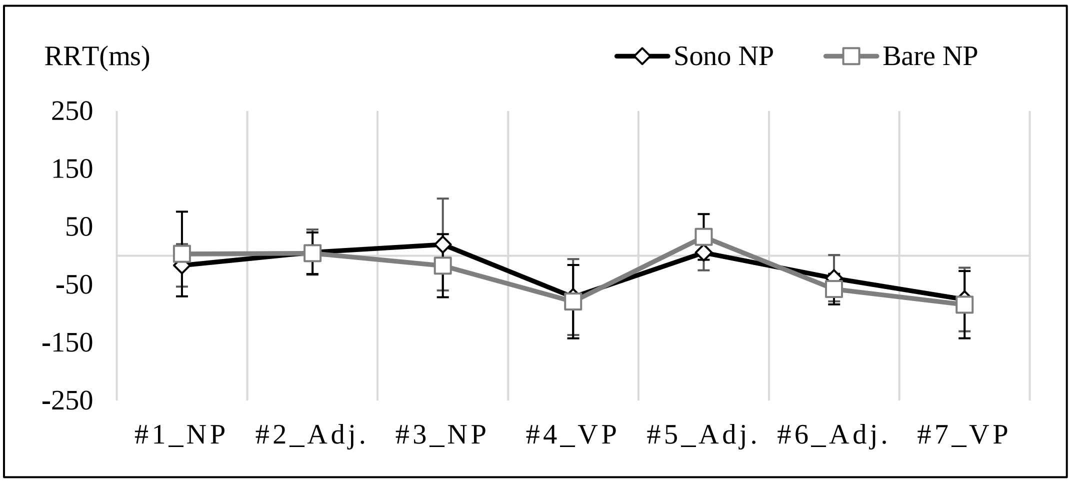


Figure Appx.9 Mean RRTs for sono *items in the bridging context: Pilot SPRT ver. 2 (error bars = SE)*

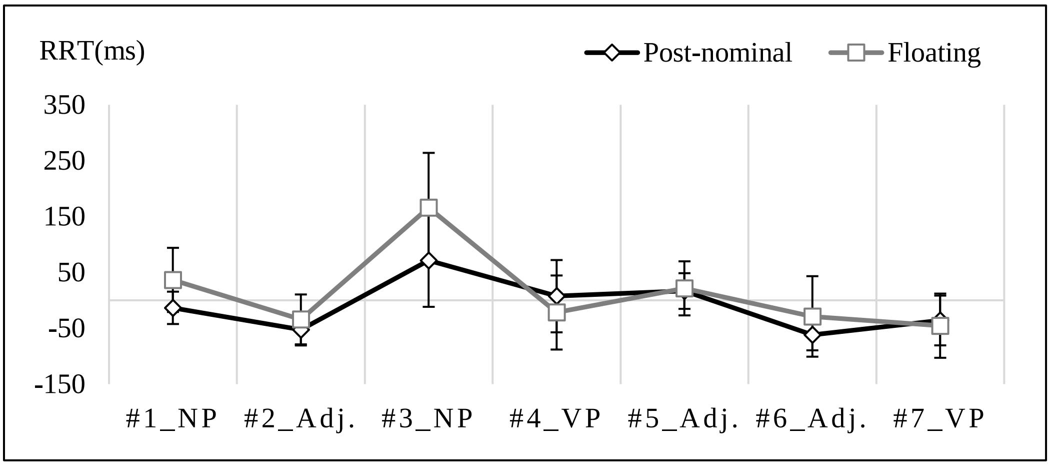


Figure Appx.10 Mean RRTs for NQ construction items in the [+definite] context: Pilot SPRT ver. 2 (error bars = SE)

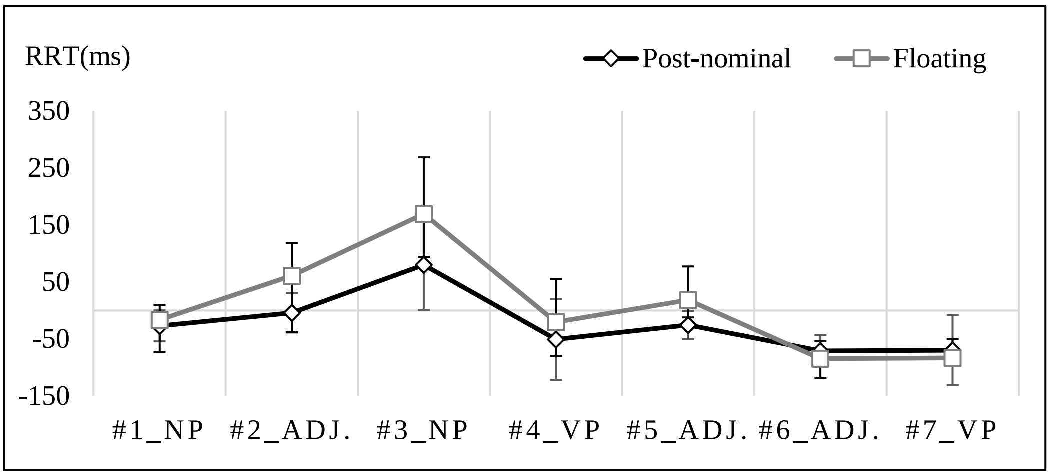


Figure Appx.11 Mean RRTs for NQ construction items in the [−definite] context: Pilot SPRT ver. 2 (error bars = SE)

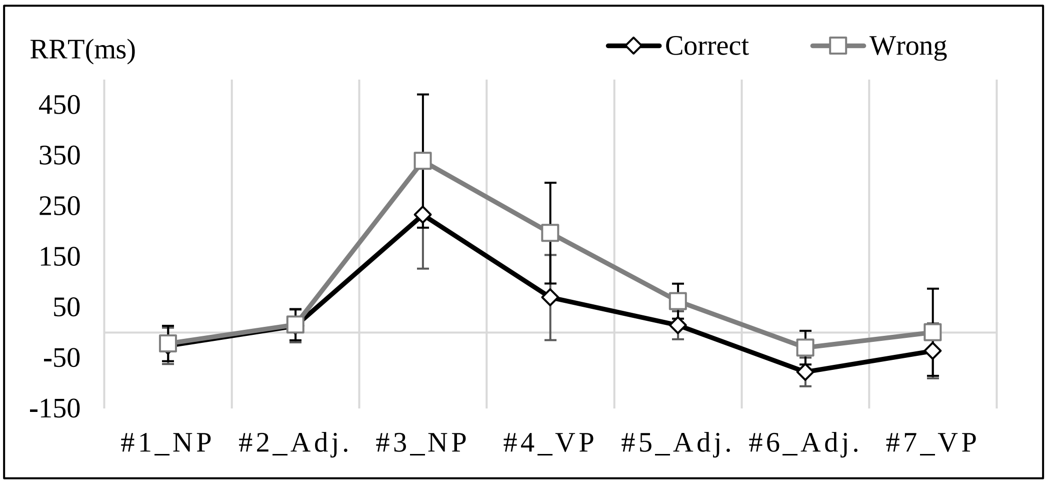


Figure Appx.12 Mean RRTs for noun-classifier agreement items: Pilot SPRT ver. 2 (error bars = SE)

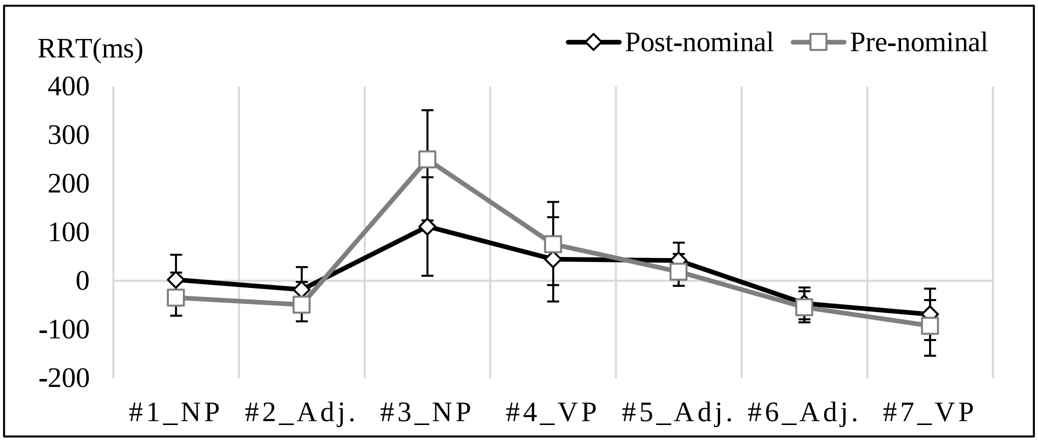


Figure Appx.13 Mean RRTs for items regarding collective reading of proper names with -tati*: Pilot SPRT ver. 2 (error bars = SE)*

Appendix 5: Test items for the main AJT and SPRT

Notes:

(i) The same sentences were used in the AJT and the SPRT.

(ii) "C" stands for context sentence, "T" target sentence and "Q" comprehension question (comprehension questions are only relevant to the SPRT).

(iii) Romanization of the Japanese text and glossing are provided only for the target sentences.

(iv) Furigana (reading aid) were provided for all the kanji (Chinese characters) in the actual test materials though omitted in this appendix for space reasons.

(v) "/" indicates a segment boundary for the SPRT.

(vi) The two alternatives are highlighted in green.

(vii) "*" means being predicted to be unacceptable/ungrammatical, and unacceptable/ungrammatical segments are in bold.

(viii) English translations given to unacceptable sentences are the intended meaning for which they are considered unacceptable/ungrammatical.

Target items (1.1–5.8)

*1. Numeral quantifier (NQ) constructions: [+definite] contexts (post-nominal vs. *floating)*

1.1 C 太郎には息子と娘が1人ずついて、朝はいつもひとりで世話をしています。…

‘Taroo has one son and daughter and always takes care of them all by himself in the morning. …’

T	太郎は/	今朝も/	{	子供2人を/	*OR*	***子供を2人/**	}	いつものように/	7時に/	起こしました。
	Taroo-wa	kesa-mo	{	kodomo huta-ri-o	*OR*	**kodomo-o huta-ri**	}	itumo-no yooni	siti-zi-ni	okosi-masita
	Taroo-TOP	this.morning-too	{	child-2-CL-ACC	*OR*	**child-ACC 2-CL**	}	as.always	7-o’clock-at	wake.up-POL.PST

‘Taroo, as always, woke the two children up at 7 this morning.’

Q この文章によると… *(asked only in the post-nominal condition)*

‘According to this passage…’

A 太郎は子供が2人います。

‘Taroo has two children.’

B 太郎は子供が3人います。

‘Taroo has three children.’

1.2 C 太郎はテニスがとても上手です。昨日も、友人のタカシ、ヒロシ、ゴロウと試合をしましたが、…

'Taroo is a good tennis player. He had tennis matches with his friends, Takesi, Hirosi and Goroo yesterday. '

T	太郎は/	いつものように/	{ 友人3人を/	OR	***友人を3人/**	}	簡単に/	すぐ/	負かしてしまいました。
	Taroo-wa	itumo-no yooni	{ yuuzin san-nin-o	OR	**yuuzin-o san-nin**	}	kantanni	sugu	makasite-simai-masita
	Taroo-TOP	as.always	{ friend-3-CL-ACC	OR	**friend-ACC 3-CL**	}	easily	quickly	beat-finish-POL.PST

'Taro, as always, beat the three friends easily and quickly.'

Q この文章によると… *(asked only in the post-nominal condition)*

'According to this passage...'

A 太郎は3人の友人にテニスで勝ちました。

'Taroo beat three friends at tennis.'

B 太郎は4人の友人にテニスで勝ちました。

'Taroo beat four friends at tennis.'

1.3 C 花子は「ポチ」、「シロ」、「ハチ」という犬を飼っていて、毎日がんばって世話をしています。昨日、…

'Hanako has dogs, named Poti, Siro, and Hati. She takes a good care of them every day. Yesterday, ... '

T	花子は/	久しぶりに/	{ 犬3匹を/	OR	***犬を3匹/**	}	お風呂場で/	きれいに/	洗ってあげました。
	Hanako-wa	hisasiburini	{ inu san-biki-o	OR	**inu-o san-biki**	}	ohuroba-de	kireini	aratte-age-masita
	Hanako-TOP	for.the.first.time.in.a.while	{ dog-3-CL-ACC	OR	**dog-ACC 3-CL**	}	bathroom-in	completely	wash-give-POL.PST

'Hanako washed the three dogs clean in the bathroom for the first time in a while.'

1.4 C 太郎は「タマ」と「ミケ」という名前の猫を飼っています。…

‘Taroo has cats named Tama and Mike. … ’

T	太郎は/	小学生の時に/	{ 猫2匹を/	*OR*	***猫を2匹/**	}	公園で/	偶然/	保護しました。
	Taroo-wa	syogakusei-no toki-ni	{ neko ni-hiki-o	*OR*	**neko-o ni-hiki**	}	kooen-de	guuzen	hogo-si-masita
	Taroo-TOP	as.a.primary.school.kid	{ cat-2-CL-ACC	*OR*	**cat-ACC 2-CL**	}	park-in	accidentally	protection-do-POL.PST

‘Taroo accidentally rescured the two cats abondoned in a park when he was in primary school.’

1.5 C 昨日、太郎は昼食に食べるのに、サケとツナとウメのおにぎりをコンビニで1つずつ買いました。しかし、…

‘Taroo bought rice balls with salmon, tuna, and ume fillings, one for each type, for lunch at a convenient store yesterday. But …’

T	太郎は/	通勤中/	{ おにぎり3つを/	*OR*	***おにぎりを3つ/**	}	電車で/	うっかり/	食べてしまいました。
	Taroo-wa	tuukin-tyuu	{ onigiri mit-tu-o	*OR*	**onigiri-o mit-tu**	}	densya-de	ukkari	tabete-simai-masita
	Taroo-TOP	during.commute	{ rice.ball 3-CL-ACC	*OR*	**rice.ball-ACC 3-CL**	}	train-on	thoughtlessly	eat-finish-POL.PST

‘Taroo thoughtlessly ate up the three rice balls on the train to work.’

Q この文章によると… *(asked only in the post-nominal condition)*

‘According to this passage…’

A 太郎は電車でおにぎりを全部食べてしまいました。

‘Taroo ended up eating all the rice balls on the train.’

B 太郎は電車でおにぎりを少し食べてしまいました。

‘Taroo ended up eating some of the rice balls on the train.’

1.6 C 太郎はいつも赤い手帳と黒い手帳を1冊ずつ持ち歩いています。今朝も、…

‘Taroo always carries with him red and black diaries, one for each colour. This morning, …’

T	太郎は/	いつものように/	{	手帳2冊を/	*OR*	***手帳を2冊/**	}	カバンに/	忘れずに/	入れました。
	Taroo-wa	itumo-no yooni	{	tetyoo ni-satu-o	*OR*	**tetyoo-o ni-satu**	}	kaban-ni	wasurezuni	ire-masita
	Taroo-TOP	as.always	{	diary 2-CL-ACC	*OR*	**diary-ACC 2-CL**	}	bag-in	without.forgetting	put-POL.PST

‘Taroo, as always, remembered to put the two diaries in his bag.’

Q この文章によると… *(asked only in the post-nominal condition)*

‘According to this passage…’

A 太郎は2冊の手帳をカバンに入れました。

‘Taroo put two diaries in his bag.’

B 太郎は3冊の手帳をカバンに入れました。

‘Taroo put three diaries in his bag.’

1.7 C 花子は誕生日に友人からクラシックとジャズとロックのCDを1枚ずつもらいました。…

‘Hanako got classic, Jazz, and rock CDs, one for each genre, from a friend for her birthday. … ’

T	花子は/	今朝/	{	CD3枚を/	*OR*	***CDを3枚/**	}	車で/	楽しく/	聞きました。
	Hanako-wa	kesa	{	siidjii san-mai-o	*OR*	**siidjii-o san-mai**	}	kuruma-de	tanosiku	kiki-masita
	Hanako-TOP	this.morning	{	CD 3-CL-ACC	*OR*	**CD-ACC 3-CL**	}	car-in	happily	listen-POL.PST

‘Hanako enjoyed listening to the three CDs in the car this morning.’

1.8 C 花子は自転車に乗るのが好きで、通勤用と、趣味用の自転車を1台ずつ持っています。…

‘Hanako likes biking and owns one bike for commuting and another for her pastime. … ’

T							
	花子は/	先週末/	{ 自転車2台を/	*OR* ***自転車を2台/** }	自分で/	ていねいに/	手入れしました。
	Hanako-wa	sensyu-matu	{ zitensya ni-dai-o	*OR* **zitensya-o ni-dai** }	zibun-de	teineini	teire-si-masita
	Hanako-TOP	last.week-end	{ bike 2-CL-ACC	*OR* **bike-ACC 2-CL** }	by.one’s.self	carefully	maintenance-do- POL.PST

‘Hanako gave a good maintenance to the two bikes by herself last weekend.’

2. NQ constructions: [−definite] contexts (post-nominal vs. floating)

2.1 C 太郎はとても優秀な警察官です。太郎は担当している地区のパトロールにとても熱心です。…

‘Taroo is a highly-skilled police officer. Taro is very enthusiastic about patrolling the area in his charge. … ’

T							
	太郎は/	今週だけで/	{ 泥棒3人を/	*OR* 泥棒を3人/ }	パトロール中に/	次々と/	捕まえました。
	Taroo-wa	konsyu dake de	{ doroboo san-nin-o	*OR* doroboo-o san-nin }	patorooru-tyuu ni	tugitugi-ni	tsukamae-masita
	Taroo-TOP	this.week only at	{ thief-3-CL-ACC	*OR* thief-ACC 3-CL }	while.patrolling	one.after.another	catch- POL.PST

‘This week alone, Taroo caught three thieves one after another while patrolling.’

2.2 C 花子は家の近くにできたばかりのレストランに行ってみたいと思っています。しかし、ひとりでは行きたくないので、…

‘A new restaurant just opened near Hanako's house and she wants to go there. But, since she does not want to go there alone, … ’

T							
	花子は/	昨日/	{ 友人2人を/	*OR* 友人を２人/ }	早速/	ランチに/	誘ってみました。
	Hanako-wa	kinoo	{ yuuzin huta-ri-o	*OR* yuuzin-o huta-ri }	sassoku	ranti-ni	sassote-mi-masita
	Hanako-TOP	yesterday	{ friend-2-CL-ACC	*OR* friend-ACC 2-CL }	immediately	lunch-for	ask.out-try.to-POL.PST

‘Hanako went ahead and asked two friends out for lunch yesterday.’

2.3 C 太郎は動物病院の医師です。犬を治療することがほとんどですが、…

‘Taroo works as a vet at an animal hospital. Although he mostly treats dogs, …’

T	太郎は/	先日/	{	猫3匹を/	*OR*	猫を3匹/	}	珍しく/	続けて/	治療しました。
	Taroo-wa	senzitu	{	neko san-biki-o	*OR*	neko-o san-biki	}	mezurasiku	tuzukete	tiryoo-si-masita
	Taroo-TOP	the.other.day	{	cat 3-CL-ACC	*OR*	cat-ACC 3-CL	}	unusually	in.a.row	treatment-do-POL.PST

‘Unusually for him, Taroo treated three cats in a row the other day.’

Q この文章によると… *(asked only in the floating condition)*

‘According to this passage…’

A 太郎は先日猫を治療しました。

‘Taroo treated cats the other day.’

B 太郎は先日うさぎを治療しました。

‘Taroo treated rabbits the other day.’

2.4 C 花子は犬が大好きで、昔から犬をずっと飼いたいと思っていました。…

‘Hanako loves dogs and has wanted to keep a dog for a long time. …’

T	花子は/	最近/	{	犬2匹を/	*OR*	犬を2匹/	}	ほとんど/	同時に/	飼い始めました。
	Hanako-wa	saikin	{	inu ni-hiki-o	*OR*	inu-o ni-hiki	}	hotondo	doozini	kai-hazime-masita
	Taroo-TOP	recently	{	dog 2-CL-ACC	*OR*	dog-ACC 2-CL	}	almost	at.the.same.time	keep-start-POL.PST

‘Hanako started keeping two dogs almost at the same time recently.’

Q この文章によると… *(asked only in the floating condition)*

‘According to this passage…’

A 花子は最近犬を飼い始めました。

‘Hanako started keeping some dogs recently.’

B 花子は最近猫を飼い始めました。

‘Hanako started keeping some cats recently.’

2.5 C 花子は今日の午後、久しぶりに遠くの親せきの家に遊びに行きました。その親せきは3人家族なので、…

‘Hanako visited a distant relative's house for the first time in a while this afternoon. Since they are a family of three, …’

T	花子は/	午前中/	{	チーズケーキ3つを/	*OR*	チーズケーキを3つ/	}	おみやげに/	ケーキ屋で/	買いました。
	Hanako-wa	gozen-tyuu	{	tiizukeeki mit-tu-o	*OR*	tiizukeeki-o mit-tu	}	omiyage-ni	keekiya-de	kai-masita
	Hanako-TOP	in.the.morning	{	cheese.cake 3-CL-ACC	*OR*	cheese.cake-ACC 3-CL	}	present-for	cake.shop-at	buy-POL.PST

‘Hanako bought three cheese cakes for them at a cakeshop in the morning.’

Q この文章によると… *(asked only in the floating condition)*

‘According to this passage…’

A 花子は今朝クッキーを買いました。

‘Hanako bought cookies this morning.’

B 花子は今朝ケーキを買いました。

‘Hanako bought cakes this morning.’

2.6 C 太郎はほぼ毎日インターネットで買い物をします。…

‘Taroo does online shopping almost every day. …’

T	太郎は/	昨日も/	{	本2冊を/	*OR*	本を2冊/	}	アマゾンで/	つい/	注文してしまいました。
	Taroo-wa	kinoo-mo	{	hon ni-satu-o	*OR*	hon-o ni-satu	}	Amazon-de	tui	tyuumon-site-simai-masita
	Taroo-TOP	yesterday-too	{	book 2-CL-ACC	*OR*	book-ACC 2-CL	}	Amazon-on	thoughtlessly	order-do-finish-POL.PST

‘Taroo could not resist ordering two books on Amazon yesterday, too.’

Q この文章によると… *(asked only in the floating condition)*

‘According to this passage…’

A 昨日太郎はCDを注文しました。

‘Taroo ordered some CDs yesterday.’

B 昨日太郎は本を注文しました。

‘Taroo ordered some books yesterday.’

2.7 C 先日、花子は海外旅行中、帰りの飛行機に乗る直前に、両親へのおみやげを買い忘れているのに気づきました。…

'The other day, while on a trip abroad, Hanako realised that she had not bought any gifts for her parents, right before boarding on the flight back home. ...'

T	花子は/	みやげ物屋で/	{ 絵はがき2枚を/	*OR*	絵はがきを2枚/	}	急いで/	両親に/	買いました。
	Hanako-wa	miyagemonoya-de	{ ehagaki ni-mai-o	*OR*	ehagaki-o ni-mai	}	isoide	ryoosin-ni	kai-masita
	Hanako-TOP	gift.shop-at	{ postcard 2-CL-ACC	*OR*	postcard-ACC 2-CL	}	in.a.hurry	parents-for	buy- POL.PST

'Hanako hurried up and bought two postcards for her parents at a gift shop.'

Q この文章によると… *(asked only in the floating condition)*

'According to this passage...'

A 花子は旅行のおみやげに絵はがきを買いました。

'Hanako bought some postcards as travel souvenirs.'

B 花子は旅行のおみやげにチョコレートを買いました。

'Hanako bought some chocolates as travel souvenirs.'

2.8 C 太郎は電気屋の優秀なセールスマンです。昨日は客が少ない日でしたが、…

'Taroo is a very competent salesperson at an electronics store. The store did not have many customers yesterday, but... '

T	太郎は/	午前中だけで/	{ パソコン3台を/	*OR*	パソコンを3台/	}	次々に/	上手く/	売ることができました。
	Taroo-wa	gozen-tyuu-dake-de	{ pasokon san-dai-o	*OR*	pasokon-o san-dai	}	tugitugini	umaku	uru koto-ga deki-masita
	Taroo-TOP	in.the.morning.alone	{ computer 3-CL-ACC	*OR*	computer-ACC 3-CL	}	one.after.another	well	was.able.to.sell

'In the morning alone, Taroo successfully sold three PCs one after another.'

Q この文章によると… *(asked only in the floating condition)*

'According to this passage...'

A 太郎は昨日、仕事でパソコンを売りました。

'Taroo sold some computers at work yesterday.'

B 太郎は昨日、仕事でテレビを売りました。

'Taroo sold some TVs at work yesterday.'

3. Sono*: Anaphoric contexts (*sono *NP vs. bare NP)*

3.1 C 太郎の家の近くにとても人気のある、安くて美味しいレストランがあります。…

‘There is a very popular restaurant which serves great food for reasonable prices near Taroo's house. …’

T	太郎は/	昔から/	{ そのレストランを/	*OR*	レストランを/	}	誰よりも/	よく/	利用しています。
	Taroo-wa	mukasi-kara	{ sono resutoran-o	*OR*	resutoran-o	}	dare-yori-mo	yoku	riyoo-site-i-masu
	Taroo-TOP	past-from	{ SONO restaurant-ACC	*OR*	restaurant-ACC	}	more.than.anyone	often	use-do-ASP-POL.NPST

‘Taroo has eaten at the restaurant more often than anyone since long ago.’

Q この文章によると… *(asked only in the* sono *NP condition)*

‘According to this passage…’

A 太郎にはよく利用するホテルがあります。

‘There is a hotel that Taroo often stays at.’

B 太郎は昨日、仕事でテレビを売りました。

‘There is a restaurant that Taroo often eats at.’

3.2 C 花子の家の近くには大きな公園があります。緑が豊かでとても落ち着く場所です。…

‘There is a big park near Hanako's house. It is a very calming place with rich green. …’

T	花子は/	長年/	{ その公園を/	*OR*	公園を/	}	毎朝/	必ず/	散歩しています。
	Hanako-wa	naganen	{ sono kooen-o	*OR*	kooen-o	}	maiasa	kanarazu	sanpo-site-i-masu
	Hanako-TOP	for.many.years	{ SONO park-ACC	*OR*	park-ACC	}	every.morning	always	walk-do-ASP-POL.NPST

‘Hanako has been taking a walk in the park every morning for many years.’

Q この文章によると… *(asked only in the* sono *NP condition)*

‘According to this passage…’

A 花子は海岸を毎朝散歩しています。

‘Hanako takes a walk along the coast every morning.’

B 花子は公園を毎朝散歩しています。

‘Hanako takes a walk in the park every morning.’

3.3 C 太郎にはドイツに住んでいる友人がいて、今年、ドイツまで遊びに行きました。…
‘Taroo has a friend living in Germany and visited him in Germany this year. …’

T										
	太郎は/	来年も/	{	その友人を/	*OR*	友人を/	}	絶対に/	また/	訪ねるつもりです。
	Taroo-wa	rainen-mo	{	sono yuuzin-o	*OR*	yuuzin-o	}	zettai-ni	mata	tazuneru-tumori-desu
	Taroo-TOP	next.year-too	{	SONO friend-ACC	*OR*	friend-ACC	}	surely	again	visit-intend.to-COP

‘Taroo is definitely going to visit the friend again next year.’

Q この文章によると… *(asked only in the* sono *NP condition)*
‘According to this passage…’

A 太郎は来年、フランスの友人を訪ねようと思っています。
‘Taroo is going to visit a friend in France next year.’

B 太郎は来年、ドイツの友人を訪ねようと思っています。
‘Taroo is going to visit a friend in Germany next year.’

3.4 C 花子は週末に今話題の映画を見に行きました。とても感動したので、…
‘Hanako went see a much-talked-about movie at the weekend. Since she was deeply moved by the movie, …’

T										
	花子は/	今日/	{	その映画を/	*OR*	映画を/	}	友人に/	早速/	すすめました。
	Hanako-wa	kyoo	{	sono eiga-o	*OR*	eiga-o	}	yuuzin-ni	sassoku	susume-masita
	Hanako-TOP	today	{	SONO movie-ACC	*OR*	movie-ACC	}	friend-to	immediately	recommend-POL.PST

‘She already recommended the movie to her friends today.’

Q この文章によると… *(asked only in the* sono *NP condition)*
‘According to this passage…’

A 花子は今日、小説を友人にすすめました。
‘Hanako recommended a novel to her friends today.’

B 花子は今日、映画を友人にすすめました。
‘Hanako recommended a movie to her friends today.’

3.5 C 太郎は小学生の時、学校の帰り道に捨てられていた子猫を拾い、家で飼うことにしました。それから15年経ちますが、…

‘When he was in primary school, Taroo found an abandoned kitten on the way home from school and took it home and decided to keep it. Fifteen years on, …’

T	太郎は/	今も/	{	その猫を/	*OR*	猫を/	}	毎日/	とても/	かわいがっています。
	Taroo-TOP	ima-mo	{	sono neko-o	*OR*	neko-o	}	mainiti	totemo	kawaigatte-i-masu
	Taroo-TOP	now-too	{	SONO cat-ACC	*OR*	cat-ACC	}	every.day	very.much	adore-ASP-POL.NPST

‘Taroo still adores the cat every day.’

3.6 C 高校生の花子は犬が苦手です。花子の家のとなりの家はもう10年以上、1匹の大きな犬を飼っていますが、…

‘Hanako is a high school student. She is not good with dogs. Hanako's next door neighbour has had a big dog for more than ten years. …’

T	花子は/	昔から/	{	その犬を/	*OR*	犬を/	}	いつも/	できるだけ/	避けています。
	Hanako-wa	mukasi-kara	{	sono inu-o	*OR*	inu-o	}	itumo	dekiru-dake	sakete-i-masu
	Hanako-TOP	past-from	{	SONO dog-ACC	*OR*	dog-ACC	}	always	as.much.as possible	avoid-ASP-POL.NPST

‘Hanako has always tried to avoid the dog as much as possible since a long time ago.’

3.7 C 太郎は最近、新しくてかっこいい車を買いました。…

‘Taroo has bought a new cool car recently. …’

T	太郎は/	毎日のように/	{	その車を/	*OR*	車を/	}	だれかに/	うれしそうに/	自慢しています。
	Taroo-wa	mainiti-no yooni	{	sono kuruma-o	*OR*	kuruma-o	}	dare-ka-ni	uresi sooni	ziman-site-i-masu
	Taroo-TOP	almost.every.day	{	SONO car-ACC	*OR*	car-ACC	}	someone-to	happily	pride-do-ASP-POL.NPST

‘Taroo has been happily showing off the car to someone almost everyday.’

Q この文章によると… *(asked only in the* sono *NP condition)*

‘According to this passage…’

A 太郎は毎日のようにバイクの自慢をしています。

‘Taroo is showing off his motor bike almost every day.’

B 太郎は毎日のように車の自慢をしています。

‘Taroo is showing off his car almost every day.’

3.8 C 昨日、花子は仕事の帰り道に迷子の男の子を見つけました。親が近くにいないようだったので、…

‘Hanako found a lost boy on her way home from work yesterday. Since his parent did not seem to be around, …’

T	花子は/	すぐ/	{	その男の子を/	*OR*	男の子を/	}	車で/	交番へ/	連れて行きました。
	Hanako-wa	sugu	{	sono otokonoko-o	*OR*	otokonoko-o	}	kuruma-de	kooban-e	turete-iki-masita
	Hanako-TOP	immediately	{	SONO boy-ACC	*OR*	boy-ACC	}	car-by	police.station-to	take-go-POL.PST

‘Hanako took the boy to the police station immediately by car.’

Q この文章によると… *(asked only in the* sono *NP condition)*

‘According to this passage…’

A 花子は迷子の女の子を交番に連れて行きました。

‘Hanako took a lost girl to the police station.’

B 花子は迷子の男の子を交番に連れて行きました。

‘Hanako took a lost boy to the police station.’

4. Sono*: Non-anaphoric contexts (**sono *NP vs. bare NP)*

4.1 C 太郎は夜に散歩するのが好きです。天気が良く、空に雲がない時は、…

‘Taroo likes taking a night walk. When the weather is nice and the sky is clear of clouds, …’

T	太郎は/	昔から/	{	***その月を/**	*OR*	月を/	}	ゆっくり/	楽しんで/	見ています。
	Taroo-wa	mukasi-kara	{	**sono tuki-o**	*OR*	tuki-o	}	yukkuri	tanosinde	mite-i-masu
	Taroo-TOP	past-from	{	**SONO moon-ACC**	*OR*	moon-ACC	}	slowly	gladly	watch-ASP-POL.NPST

‘Taroo relaxes and enjoys watching the moon as an old habit.’

4.2 C 花子は最近、楽しい夢を見ました。夢の中で…

‘Hanako had a pleasant dream recently. In that dream, …’

T	花子は/	鳥のように/	{	***その空を/**	*OR*	空を/	}	自由に/	気持ちよく/	飛びました。
	Hanako-wa	tori-no yooni	{	**sono sora-o**	*OR*	sora-o	}	jiyuuni	kimotiyoku	tobi-masita
	Hanako-TOP	like.a.bird	{	**SONO sky-ACC**	*OR*	sky-ACC	}	freely	pleasantly	fly-POL.NPST

‘Hanako enjoyed flying freely in the sky like a bird.’

4.3 C 太郎は昨夜テレビでニュースを見ていました。明日はハイキングに行くので、…

‘Taroo was watching news on TV last night. Since he was going on a hiking tomorrow, …’

T	太郎は/	明日のために/	{	***その天気を/**	*OR*	天気を/	}	忘れずに/	しっかりと/	確認しました。
	Taroo-wa	asita-no tame-ni	{	**sono tenki-o**	*OR*	tenki-o	}	wasurezuni	sikkarito	kakunin-si-masita
	Taroo-TOP	for.tomorrow	{	**SONO weather-ACC**	*OR*	weather-ACC	}	without.forgetting	firmly	check-do-POL.PST

‘Taroo did not forget to check the weather carefully for tomorrow.’

Q この文章によると… *(asked only in the bare NP condition)*

‘According to this passage…’

A 太郎は、昨夜、ハイキングのために天気を確認しました。

‘Taroo checked the weather for a hiking last night.’

B 太郎は、昨夜、ハイキングのために持ち物を確認しました。

‘Taroo checked the things to take to a hiking last night.’

4.4 C 花子は昨夜、変な夢を見ました。それは、花子がたった1人で宇宙人と戦うというものでした。…

‘Hanako had a strange dream last night. In that dream, Hanako had to fight aliens alone. …’

T	花子は/	宇宙人から/	{ ***その地球を/**	*OR* 地球を/ }	1人で/	なんとか/	守ることができました。
	Hanako-wa	utyuujin-kara	{ **sono tikyuu-o**	*OR* tikyuu-o }	hito-ri-de	nantoka	mamoru koto-ga deki-masita
	Hanako-TOP	alien-from	{ **SONO earth-ACC**	*OR* earth-ACC }	by.herself	somehow	was.able.to.protect

‘Hanako managed to protect the Earth from the aliens all by herself.’

Q この文章によると… *(asked only in the bare NP condition)*

‘According to this passage…’

A 花子は1人で宇宙人から地球を守りました。

‘Hanako protected the Earth from aliens by herself.’

B 花子は友人と一緒に宇宙人から地球を守りました。

‘Hanako protected the Earth from aliens with her friends.’

4.5 C ヒマワリは日が当たる方向に向く性質があります。…

‘Sunflowers turn towards the Sun. …’

T	ヒマワリの花は/	日中/	{ ***その太陽を/**	*OR* 太陽を/ }	必ず/	東から西に/	追いかけます。
	Himawari-no-hana-wa	nittyuu	{ **sono taiyoo-o**	*OR* taiyoo-o }	kanarazu	higasi-kara nisi-ni	oikake-masu
	sunflower-GEN-flower-TOP	during.the.daytime	{ **SONO Sun-ACC**	*OR* Sun-ACC }	always	East-from West-to	follow-POL.NPST

‘Sunflowers always follow the Sun from east to west during the daytime.’

4.6 C テクノロジーの進化は社会にいろんな変化を与えます。…
'Technological advances change our society in many ways. …'

T	インターネットは/	特に/	{ ***その世界を/**	*OR* 世界を/ }	短期間で/	大きく/	変えました。
	Intaanetto-wa	tokuni	{ **sono sekai-o**	*OR* sekai-o }	tankikan-de	ookiku	kae-masita
	Internet-TOP	particularly	{ **SONO world-ACC**	*OR* world-ACC }	short.period-in	greatly	change-POL.PST

'Particularly, the Internet has brought a major change to the world in such a short period of time.'

Q この文章によると… *(asked only in the bare NP condition)*
'According to this passage…'

A インターネットは世界を大きく変えました。
'The Internet has brought a major change to the world.'

B スマートフォンは世界を大きく変えました。
'Smart phones have brought a major change to the world.'

4.7 C 太郎は山奥の村で生まれました。太郎は生まれてから、あまり村の外に出たことがありません。…
'Taroo was born in a village deep in the mountains. Taroo has not been out of the village much since he was born. …'

T	太郎は/	まだ/	{ ***その海を/**	*OR* 海を/ }	自分の目で/	実際に/	見たことがありません。
	Taroo-wa	mada	{ **sono umi-o**	*OR* umi-o }	zibun-no me-de	zissaini	mi-ta koto-ga ari-masen
	Taroo-TOP	yet	{ **SONO ocean-ACC**	*OR* ocean-ACC }	self-GEN eye-with	in.reality	have.not.seen

'Taroo has never seen the ocean with his own eyes yet.'

Q この文章によると… *(asked only in the bare NP condition)*
'According to this passage…'

A 太郎は生まれてから一度も海を見たことがありません。
'Taroo has never seen the ocean in his life.'

B 太郎は生まれてから一度も雪を見たことがありません。
'Taroo has never seen snow in his life.'

4.8 C 電気自動車はガソリン車に比べて環境に優しいと考えられています。…

‘Electric cars are thought to be more environmentally-friendly than gasoline cars. …’

T	電気自動車は/	ガソリン車と違い/	{ ***その空気を/**	*OR*	空気を/ }	有毒なガスで/	ほとんど/	汚しません。
	denki-jidoosya-wa	gasorin-sya-to tigai	{ **sono kuuki-o**	*OR*	kuuki-o }	yuudokuna-gasu-de	hotondo	yogosi-masen
	electric.cars-TOP	unlike.gosoline.cars	{ **SONO air-ACC**	*OR*	air-ACC }	toxic-gas-with	mostly	pollute-POL.NEG

‘Electric cars are thought to be more environmentally-friendly than gasoline cars. …’

Q この文章によると… *(asked only in the bare NP condition)*

‘According to this passage…’

A 電気自動車は空気をほとんど汚しません。

‘Electric cars hardly pollute the air.’

B ハイブリッド車は空気をほとんど汚しません。

‘Hybrid cars hardly pollute the air.’

5. Sono*: Bridging contexts (*sono *NP vs. bare NP)*

5.1 C 花子は今、流行の小説に夢中になっていますが、…

Hanako is now into a popular novel. …’

T	花子は/	偶然/	{ その著者を/	*OR*	著者を/ }	東京で/	一瞬だけ/	見たことがあります。
	Hanako-wa	guuzen	{ sono tyosya-o	*OR*	tyosya-o }	Tokyo-de	issyun-dake	mi-ta koto-ga ari-masu
	Hanako-TOP	accidentally	{ SONO author-ACC	*OR*	author-ACC }	Tokyo-in	one.moment-only	have.seen

‘Hanako has caught sight of the author in Tokyo once.’

Q この文章によると… *(asked only in the* sono *NP condition)*

‘According to this passage…’

A 花子はある小説の著者を見かけたことがあると言います。

‘Hanako has seen the author of a novel.’

B 花子はある映画の監督を見かけたことがあると言います。

‘Hanako has seen the director of a movie.’

5.2 C 昨夜、太郎がラジオを聞いていると、大好きなクラシックの曲が流れてきましたが、…

Taroo heard his favourite classical song being played on the radio last night. …'

T	太郎は/	不思議なことに/	{	その作曲者を/	*OR* 作曲者を/	}	どうしても/	すぐには/	思い出せませんでした。
	Taroo-wa	husigina-koto-ni	{	sono sakkyokusya-o	*OR* sakkyokusya-o	}	doo-site-mo	suguni-wa	omoidas-e-masen-desita
	Taroo-TOP	strangely.enough	{	SONO composer-ACC	*OR* composer-ACC	}	by.any.means	quickly-TOP	could.not.recall

'Strangely enough, Taroo just could not recall who the composer was immediately.'

5.3 C 花子の町の駅前に最近、とても変なデザインのビルが建てられましたが、…

A building with a strange design has been constructed in front of the train station in the city where Hanako lives recently. …'

T	町の人は/	デザインのせいで/	{	その設計者を/	*OR* 設計者を/	}	あまり/	好ましく/	思っていません。
	mati-no hito-wa	dezain-no seide	{	sono sekkeisya-o	*OR* sekkeisya-o	}	amari	konomasiku	omotte-i-masen
	people.in.city-TOP	because.of.the.design	{	SONO designer-ACC	*OR* designer-ACC	}	much	pleasantly	think-ASP-POL.NEG

'Because of the design, people in the city do not have a positive image of the designer.'

5.4 C 花子には見たいと思っている公開されたばかりの映画がありますが、…

'Hanako wants to watch a flim that was just released. …'

T	花子は/	大学生の時から/	{	その監督を/	*OR* 監督を/	}	心から/	とても/	尊敬しています。
	Hanako-wa	daigakusei-no toki-kara	{	sono kantoku-o	*OR* kantoku-o	}	kokoro kara	totemo	sonkei-site-i-masu
	Hanako-TOP	since.university	{	SONO director-ACC	*OR* director-ACC	}	heart from	very much	respect-do-ASP-POL.PST

'Hanako has respected the director from the bottom of her heart since university.'

5.5 C 先月、太郎の町の美術館で写真コンテストが開かれました。優勝したのは、美しい海の写真で、…

‘A photo context was held at a museum in the city where Taroo lives. The first prize went to the photo of a beautiful ocean. …’

T	審査員は/	全員/	{ その撮影者を/	*OR* 撮影者を/	}	とても/	高く/	評価しました。
	Sinsain-wa	zenin	{ sono satueisya-o	*OR* satueisya-o	}	totemo	takaku	hyoka-si-masita
	judge-TOP	all	{ SONO photographer-ACC	*OR* photographer-ACC	}	very	highly	evaluation-do-POL.PST

‘All judges thought very highly of the photographer.’

5.6 C 太郎の町では100歳になった高齢者に金メダルを贈る決まりができましたが、…

‘Taroo's city has made a new rule that gold medals are given to centenarians. …’

T	太郎は/	実は/	{ その提案者を/	*OR* 提案者を/	}	昔から/	よく/	知っています。
	Taroo-wa	zitu-wa	{ sono teiansya-o	*OR* teiansya-o	}	mukasi-kara	yoku	sitte-i-masu
	Taroo-TOP	in.fact	{ SONO proposer-ACC	*OR* proposer-ACC	}	past-from	well	know-ASP-POL.NPST

‘Taroo has known the proposer very well for a long time.’

Q この文章によると… *(asked only in the* sono *NP condition)*

‘According to this passage…’

A 太郎は町のある決まりを提案した人を昔から知っています。

‘Taroo has known the proposer of a rule in the city from a long time ago.’

B 太郎は町のある決まりを提案した人を最近知り合いました。

‘Taroo got to know each other with the proposer of a rule in the city recently.’

5.7 C もうすぐ空を飛べる自転車が発売されることになりましたが、…

‘It has been announced that a flying bike will be on sale soon. However, …’

T	メーカーは/	まだ/	{ その発明者を/	*OR* 発明者を/	}	しばらくは/	公に/	発表しない予定です。
	meekaa-wa	mada	{ sono hatumeisya-o	*OR* hatumeisya-o	}	sibaraku-wa	ooyakeni	happyoo-si-nai yotei-desu
	manufacturer-TOP	yet	{ SONO inventor-ACC	*OR* inventor-ACC	}	for.the.time.being	publicly	plan.not.to.announce

‘The manufacturer plans not to publicise who the inventor is for the time being.’

5.8 C 花子は最近、好みの詩を見つけました。しかし、誰が書いたものかが分からないので、…

‘Hanako found a poem that is to her taste recently. But, since she does not know who wrote it, …’

T	花子は/	気になって/	{	その作者を/	*OR*	作者を/	}	必死に/	今/	調べているところです。
	Hanako-wa	ki-ni-natte	{	sono sakusya-o	*OR*	sakusya-o	}	hissini	ima	sagasite-i-ru tokoro-desu
	Hanako-TOP	anxiously	{	SONO poet-ACC	*OR*	poet-ACC	}	desparately	now	is.doing.research

‘Hanako is now trying hard to find out who the poet is.’

Control fillers (6.1–9.8)

*6. Control filler 1: Noun-numeral classifier agreement (correct vs. *wrong)*

6.1 C 太郎はたくさん友人がいて、毎日、だれかを家に呼んで遊んでいます。…

‘Taroo has a lot of friends and invites someone over to his house to hang out with every day. …’

T	太郎は/	今日も/	{	友人を2人/	*OR*	***友人を2匹**/	}	午後から/	家に/	呼びました。
	Taroo-wa	kyoo-mo	{	yuuzin-o huta-ri	*OR*	**yuuzin-o ni-hiki**	}	gogo-kara	ie-ni	yobi-masita
	Taroo-TOP	today-too	{	friend-ACC 2-CL	*OR*	**friend-ACC 2-CL**	}	afternoon-from	house-to	call-POL.PST

‘Taroo invited two friends over to his house this afternoon, too.’

6.2 C 田中先生は数学教師です。毎日、数学が苦手な学生を放課後、職員室に呼んで指導しています。…

‘Mr. Tanaka is a maths teacher. He tutors students who are not good at maths in the teacher's room after school every day. …’

T	田中先生は/	昨日も/	{ 学生を３人/	OR	***学生を3つ/**	}	放課後/	職員室で/	指導しました。
	Tanaka-sensei-wa	kinoo-mo	{ gakusei-o san-nin	OR	**gakusei-o mit-tu**	}	hooka-go	syokuinsitu-de	sidoo-si-masita
	Mr. Tanaka-TOP	yesterday-too	{ student-ACC 3-CL	OR	**student-ACC 3-CL**	}	after.school	teacher’s.room-in	instruction-do-POL.PST

‘Mr. Tanaka taught three students in the teacher's room after school yesterday.’

Q この文章によると… *(asked only in the correct condition)*

‘According to this passage…’

A 田中先生は放課後に数学を指導しました。

‘Mr. Tanaka taught maths after school.’

B 田中先生は昼休みに数学を指導しました。

‘Mr. Tanaka taught maths during the lunch break.’

6.3 C 太郎は小さい生き物を飼うのが好きです。…

‘Taroo has a passion for keeping small creatures as pets. …’

T	太郎は/	最近/	{ カエルを2匹/	OR	***カエルを2人/**	}	家で/	こっそり/	飼っています。
	Taroo-wa	saikin	{ kaeru-o ni-hiki	OR	**kaeru-o hu-tari**	}	ie-de	kossori	katte-i-masu
	Taroo-TOP	recently	{ frog-ACC 2-CL	OR	**frog-ACC 2-CL**	}	home-at	secretly	keep-ASP-POL.PST

‘Taroo has been keeping two frogs secretly in his house recently.’

Q この文章によると… *(asked only in the correct condition)*

‘According to this passage…’

A 太郎は家でトカゲを飼っています。

‘Taroo keeps a lizard at home.’

B 太郎は家でカエルを飼っています。

‘Taroo keeps a frog at home.’

6.4 C 花子の家の猫のタマは小さな動物を捕まえるのが得意です。…

‘Tama, Hanako's cat, has a great skill for hunting small animals. …’

T	タマは/	昨日も/	{	ねずみを3匹/	*OR*	***ねずみを3台/**	}	庭で/	上手に/	捕まえました。
	Tama-wa	kinoo-mo	{	nezumi-o san-biki	*OR*	**nezumi-o san-dai**	}	niwa-de	jyoozuni	tukamae-masita
	Tama-TOP	yesterday-too	{	mouse-ACC 3-CL	*OR*	**mouse-ACC 3-CL**	}	garden-in	skillfully	catch-POL.PST

‘Tama skillfully caught three mice in the garden yesterday, too.’

Q この文章によると… *(asked only in the correct condition)*

‘According to this passage…’

A タマは昨日ハトを捕まえました。

‘Tama caught a pigeon yesterday.’

B タマは昨日ねずみを捕まえました。

‘Tama caught a mouse yesterday.’

6.5 C 花子は読書が趣味で、本を読むのがとても速いです。…

‘Hanako's favourite pastime is reading. She reads books really fast. …’

T	花子は/	毎日/	{	新しい本を2冊/	*OR*	***新しい本を2匹/**	}	とても/	楽しく/	読みます。
	Hanako-wa	mainiti	{	atarasii hon-o ni-satu	*OR*	**atarasii hon-o ni-hiki**	}	totemo	tanosiku	yomi-masu
	Hanako-TOP	every.day	{	new book-ACC 2-CL	*OR*	**new book-ACC 2-CL**	}	very	happily	read-POL.NPST

‘Hanako really enjoys reading two new books everyday.’

Q この文章によると… *(asked only in the correct condition)*

‘According to this passage…’

A 花子は毎日、映画を見ます。

‘Hanako watches movies every day.’

B 花子は毎日、本を読みます。

‘Hanako reads books every day.’

6.6 C 太郎は車が大好きで、すぐに新しい車が欲しくなってしまいます。…

‘Taroo has a passion for cars and cannot stop wanting new cars. …’

T	太郎は/	2年間で/	{	車を3台/	*OR*	***車を3枚**/	}	次々に/	新しく/	買いました。
	Taroo-wa	ni-nen-kan-de	{	kuruma-o san-dai	*OR*	**kuruma-o san-mai**	}	tugitugini	atarasiku	kai-masita
	Taroo-TOP	2-year-period-in	{	car-ACC 3-CL	*OR*	**car-ACC 3-CL**	}	one.after.another	newly	buy-POL.PST

‘Taroo has bought three new cars one after another in the last two years.’

6.7 C 太郎はいちごが大好きで、毎日必ず冷蔵庫に入れてあります。…

‘Taroo loves strawberries and always keeps them in the fridge. …’

T	太郎は/	今朝/	{	いちごを3つ/	*OR*	***いちごを3匹**/	}	デザートに/	美味しく/	食べました。
	Taroo-wa	kesa	{	itigo-o mit-tu	*OR*	**itigo-o san-biki**	}	dezaato-ni	oisiku	tabe-masita
	Taroo-TOP	this.morning	{	strawberry-ACC 3-CL	*OR*	**strawberry-ACC 3-CL**	}	dessert-for	deliciously	eat-POL.PST

‘Taroo enjoyed eating three strawberries for dessert this morning.’

Q この文章によると… *(asked only in the correct condition)*

‘According to this passage…’

A 太郎は今朝りんごを食べました。

‘Taroo ate some apples this morning.’

B 太郎は今朝いちごを食べました。

‘Taroo ate some strawberries this morning.’

6.8 C 花子は昨夜、仕事で帰るのがいつもより遅くなりました。夕ご飯を作る時間が無かったので、…

‘Hanako got off work later than usual last night. Since she did not have the time to cook dinner, …’

T	花子は/	夕食用に/	{ ピザを2枚/	*OR*	***ピザを2台/**	} 帰り道に/	スーパーで/	買いました。
	Hanako-wa	yuusyoku-yoo-ni	{ piza-o ni-mai	*OR*	**piza-o ni-dai**	} kaerimiti-ni	suupaa-de	kai-masita
	Hanako-TOP	for.dinner	{ pizza-ACC 2-CL	*OR*	**pizza-ACC 2-CL**	} on.the.way.home	supermarket-at	buy-POL.PST

‘Hanako bought two pizzas for dinner at a supermarket on her way home.’

Q この文章によると… *(asked only in the correct condition)*

‘According to this passage…’

A 花子は昨日、夕食にカレーを買いました。

‘Hanako bought some curry for dinner yesterday.’

B 花子は昨日、夕食にピザを買いました。

‘Hanako bought some pizzas for dinner yesterday.’

7. Control filler 2: Specificity constraint of NP-tati *(bare NP vs. *NP*-tati*)*

7.1 C 太郎は先日、会社を作りました。まだ社員がいないので、…

‘Taroo started a company the other day. He does not have any employees yet. …’

T									
太郎は/	昨日から/	{	社員を/	*OR*	***社員たちを/**	}	緊急で/	２人/	募集しています。
Taroo-wa	kinoo-kara	{	syain-o	*OR*	**syain-tati-o**	}	kinkyuu-de	huta-ri	bosyuu-site-i-masu
Taroo-TOP	yesterday-from	{	employee-ACC	*OR*	**employee-TATI-ACC**	}	urgently	2-CL	recruitment-do-ASP-POL.PST

‘Taroo has been urgently recruiting two employees since yesterday.’

Q この文章によると… *(asked only in the bare NP condition)*

‘According to this passage…’

A 太郎の会社はアルバイトを募集しています。

‘Taroo’s company is recruiting part-time workers.’

B 太郎の会社は社員を募集しています。

‘Taroo’s company is recruiting new employees.’

7.2 C 太郎の家の近くにもうすぐ大きな病院ができます。今人材を集めているところで、…

‘A big hospital is opening soon near Taroo's house. The hospital is recruiting staff now. …’

T									
その病院は/	特に/	{	事務員を/	*OR*	***事務員たちを/**	}	1人でも/	多く/	探しています。
sono byooin-wa	tokuni	{	zimuin-o	*OR*	**zimuin-tati-o**	}	hito-ri-demo	ooku	sagasite-i-masu
that hospital-TOP	especially	{	clerical.staff-ACC	*OR*	**clerical.staff-TATI-ACC**	}	1-CL-even	more	look.for-ASP-POL.NPST

‘In particular, the hospital wants as many clerical staff as possible.’

Q この文章によると… *(asked only in the bare NP condition)*

‘According to this passage…’

A 太郎の家の近くの病院は看護師を探しています。

‘The hospital near Taroo’s house is looking for nurses.’

B 太郎の家の近くの病院は事務員を探しています。

‘The hospital near Taroo’s house is looking for clerical staff.’

7.3 C 花子は付き合っている男性がいて、もうすぐ結婚する予定です。…

‘Hanako is dating with a man and they are going to marry soon. …’

T									
	花子は/	できれば/	{ 子供を/	OR	***子供たちを/**	}	30才までに/	2人/	産みたいと思っています。
	Hanako-wa	deki-reba	{ kodomo-o	OR	**kodomo-tati-o**	}	sanjyus-sai-made-ni	huta-ri	umi-tai to omotte-i-masu
	Hanako-TOP	if.possible	{ child-ACC	OR	**child-TATI-ACC**	}	30-year.old-until-by	2-CL	want.to.have

‘Hanako wants to have two children by the age of 30, if possible.’

7.4 C 太郎の大学の近所にあるコンビニでは、アルバイトがいません。…

‘The convenient store near Taro's university does not have any part-time workers. …’

T									
	そのコンビニは/	現在/	{ アルバイトを/	OR	***アルバイトたちを/**	}	少なくとも/	3人/	必要としています。
	sono konbini-wa	genzai	{ arubaito-o	OR	**arubaito-tati-o**	}	sukunaku-tomo	san-nin	hituyoo to site-i-ru
	that convenient.store-TOP	currently	{ part-time-ACC	OR	**part-time-TATI-ACC**	}	at.least	3-CL	in.need.of

‘The convenient store is currently in need of at least three part-time workers.’

7.5 C 山奥の田舎の村に住む花子にはいつかしてみたいことがあります。…

‘Hanako, who lives in a village deep in the mountains, has a thing that she wants to do some day. …’

T									
	花子は/	いつか/	{ 有名人を/	OR	***有名人たちを/**	}	一度は/	生で/	見てみたいと思っています。
	Hanako-wa	ituka	{ yuumeizin-o	OR	**yuumeizin-tati-o**	}	iti-do-wa	nama-de	mite-mi-tai to omotte-i-masu
	Hanako-TOP	some.day	{ celebrity-ACC	OR	**celebrity-TATI-ACC**	}	at.least.once	in.person	want.to.have.a.look

‘For once in her life, Hanako wants to see a celebrity in person.’

7.6 C 花子の家の近くの大学では今回の募集は定員に達しています。…

'The university near Hanako's house has already met the quota for the number of applicants this year. …'

T	その大学は/	もう/	{ 学生を/	OR	***学生たちを/**	}	今年は/	これ以上/	必要としていません。
	sono daigaku-wa	moo	{ gakusei-o	OR	***gakusei-tati-o**	}	kotosi-wa	kore-ijyoo	hituyoo to site-i-masen
	that-university-TOP	already	{ student-ACC	OR	**student-TATI-ACC**	}	this.year-TOP	more.than.this	not.in.need.of

'The university does not need any more students this year.'

Q この文章によると… *(asked only in the bare NP condition)*

'According to this passage…'

A 花子の家の近くの大学はまだまだ学生を募集しています。

'The university near Hanako's house is still accepting applications.'

B 花子の家の近くの大学はもう学生を募集していません。

'The university near Hanako's house is no longer accepting applications.'

7.7 C 最近花子の家の近くに学習塾ができました。しかし、まだ英語を教えられる講師がいないようで、…

'A new cram school has opened near Hanako's house recently. But they do not seem to have any English teachers yet. …'

T	その塾は/	現在/	{ 英語講師を/	OR	***英語講師たちを/**	}	緊急で/	たくさん/	採用しようとしています。
	sono jyuku-wa	genzai	{ eigo-koosi-o	OR	**eigo-koosi-tati-o**	}	kinkyuude	takusan	saiyoo-siyoo to site-imasu
	that cram-TOP	currently	{ English-teacher-ACC	OR	**English-teacher-TATI-ACC**	}	urgently	many	is.trying.to.recruit

'The cram school is urgently recruiting many English teachers.'

7.8 C 太郎は東京でホテルを経営しています。最近は国内からの客が減っているので、…

'Taroo runs a hotel in Tokyo. Since the number of domestic guests has been decreasing recently, …'

T	太郎は/	これからは/	{ 外国人客を/	OR	***外国人客たちを/**	}	どんどん/	積極的に/	獲得したいと考えています。
	Taroo-wa	kore-kara-wa	{ gaikokujinkyaku-o	OR	**gaikokujinkyaku-tati-o**	}	dondon	sekkyokutekini	kakutoku-si-tai to kangaete-imasu
	Taroo-TOP	from.now.on-TOP	{ overseas.guest-ACC	OR	**overseas.guest-TATI-ACC**	}	more.and.more	actively	want.to.get

'Taroo wants to actively attract overseas guests in future.'

*8. Control filler 3: Collective reading of common nouns with -*tati *(collective vs. *singular)*

8.1 C 昨日、田中先生が働く高校に { 1人の中学生の男の子とその子の両親が *OR* ***1人の中学生の男の子が** } 見学に来ました。…

'{ A junior high school boy and his parents *OR* **a junior high school boy** } visited and observed the high school that Mr. Tanaka works at yesterday. …'

T						
	田中先生は/	よろこんで/	男の子たちを/	30分ほど/	ていねいに/	案内しました。
	Tanaka-sensei-wa	yorokonde	otokonoko-tati-o	sanjyup-pun hodo	teineini	annai-si-masita
	Mr.Tanaka-TOP	happily	boy-TATI-ACC	30-minite about	politely	guide-do-POL.PST

'Mr. Tanaka happily showed the boy and the others(=his parents) around the school for about 30 minutes politely.' future.'

8.2 C 先日、花子の家に遠くに住んでいる { ある友人とその友人の母親が *OR* ***ある友人が** } 初めて遊びにきました。…

'The other day, { a friend of Hanako's and her mother *OR* **a friend of Hanako's** }, who live(s) afar, visited Hanako's house for the first time. …'

T						
	花子は/	すぐに/	友人たちを/	家族に/	簡単に/	紹介しました。
	Hanako-wa	suguni	yuuzin-tati-o	kazoku-ni	kantanni	syokai-si-masita
	Hanako-TOP	immediately	friend-TATI-ACC	family-to	briefly	introduction-do-POL.PST

'On their arrival, Hanako introduced her friend and the other(=her mother) briefly to her family.'

8.3 C 太郎はレストランでアルバイトをしていますが、 { 店長と先輩が *OR* ***店長が** } あまり仕事を教えてくれません。…

'Taroo works part-time at a restaurant. {The manager and his senior coworkers *OR* **the manager**} do(es) not teach him how to do the job much. …'

T						
	太郎は/	いつも/	店長たちを/	陰で/	悪く/	言っています。
	Taroo-wa	itumo	tentyoo-tati-o	kage-de	waruku	itte-i-masu
	Taroo-TOP	always	manager-TATI-ACC	behind.one's.back	badly	say-ASP-POL.NPST

'Taroo is always speaking ill of the manager and the others(=his senior coworkers) behind their back.'

8.4 C 花子は実家を離れて、ひとり暮らしをしていますが、最近、よく { 母親と姉と *OR* ***母親と** } 一緒に食事に行きます。…

'Hanako lives by herself away from her parents but she has been often eating out {with her mother and sister *OR* **with her mother**} recently. …'

T						
	花子は/	昨日も/	母親たちを/	自分から/	夕食に/	誘いました。
	Hanako-wa	kinoo-mo	hahaoya-tati-o	zibun-kara	yuusyoku-ni	sasoi-masita
	Hanako-TOP	yesterday-too	mother-PL-ACC	voluntarily	dinner for	ask.out-POL.PST

'Hanako asked her mother and the other(=her sister) out for dinner just yesterday.'

8.5 C 花子は { 3歳下の妹と妹の友人を *OR* ***3歳下の妹を** } 誘ってよく一緒に映画を見に行きます。…

'Hanako often asks out {her 3-year-younger sister and her sister's friend vs. **her 3-year-younger sister**} to watch movies together. …'

T						
	花子は/	先週末も/	妹たちを/	いつものように/	映画に/	誘いました。
	Hanako-wa	sensyuu-matu-mo	imooto-tati-o	itumo-no yooni	eiga-ni	sasoi-masita
	Hanako-TOP	last.week-end-too	sister-TATI-ACC	as.always	movie-for	ask.out-POL.PST

'Hanako ,as always, asked her sister and the other(=her friend) out to watch a movie last weekend too.'

8.6 C 今日、太郎の会社に、 { ある会社の社長と社員が *OR* ***ある会社の社長が** } 商談にやってきました。…

'{The president of a company and an employee vs. **the president of a company** } visited Taroo's company today for a business meeting. …'

T						
	太郎は/	商談の前に/	社長たちを/	応接室に/	ていねいに/	案内しました。
	Taroo-wa	syoodan-no mae-ni	syatyoo-tati-o	oosetusitu-ni	teineini	annai-si-masita
	Taroo-TOP	before.the.business.meeting	president-TATI-ACC	reception.office-to	politely	guide-do-POL.PST

'Before the meeting, Taroo showed them to the reception office politely.'

8.7 C 太郎には愛する { 妻と1人の娘が *OR* ***妻が** } います。…

'Taroo has a beloved {wife and daughter vs. **wife**}. …'

T	太郎は/	この夏/	妻たちを/	どこか/	旅行に/	連れて行くつもりです。
	Taroo-wa	kono natu	tuma-tati-o	doko-ka	ryokoo-ni	turete-iku-tumori-desu
	Taroo-TOP	this summer	wife-TATI-ACC	somewhere	trip-to	take-go-intend.to-COP

'Taroo is going to take his wife and the other(=his daughter) on a trip this summer.'

Q この文章によると… *(asked only in the associative condition)*

'According to this passage…'

A 太郎は妻を愛しています。

'Taroo loves his wife.'

B 太郎は娘を愛しています。

'Taroo loves his daughter.'

8.8 C 太郎は { 2歳下の弟と *OR* **.*弟の友人と** } 一緒にカラオケに行くのが好きです。…

'Taroo likes going to Karaoke {with his 2-year-younger brother and his brother's friend vs. **with his 2-year-younger brother**}. …'

T	太郎は/	先週末/	弟たちを/	久しぶりに/	カラオケに/	連れて行きました。
	Taroo-wa	sensyuu-matu	otooto-tati-o	hisasiburi-ni	karaoke-ni	turete-iki-masita
	Taroo-TOP	last.week-end	brother-TATI-ACC	for.the.first.time.in.a.while	karaoke-to	take-go-COP.PST

'Taroo took his brother and the other to karaoke last weekend for the first time in a while.' politely.' future.'

Q この文章によると… *(asked only in the associative condition)*

'According to this passage…'

A 太郎は弟とカラオケに行くのが好きです。

'Taroo loves going to karaoke with his brother.'

B 太郎は父親とカラオケに行くのが好きです。

'Taroo loves going to karaoke with his father.'

9. Control filler 4: Collective reading of proper names with -tati *(post-nominal vs. *pre-nominal)*

9.1 C 先日、太郎とタカシとヒロシは旅行中に道に迷いました。困っていると1人の男性が声をかけてきました。…

'Taroo, Takasi and Hirosi got lost while on a trip the other day. A man saw them in trouble and spoke to them. …'

T	その男性は/	親切に/	{ 太郎たち3人を/	*OR*	***3人の太郎たちを/**	}	車で/	行きたい場所まで/	案内してくれました。
	sono dansei-wa	sinsetuni	{ Taroo-tati san-nin-o	*OR*	**san-nin-no Taroo-tati-o**	}	kuruma-de	iki-tai basyo-made	annai-site-kure-masita
	that man-TOP	kindly	{ Taroo-TATI 3-CL-ACC	*OR*	**3-CL-GEN Taroo-TATI-ACC**	}	car-by	go-want.to. place-to	guide-do-give-POL.PST

'The man kindly drove the group of three represented by Taroo to the place where they wanted to go.'

Q この文章によると… *(asked only in the post-nominal condition)*

'According to this passage…'

A 先日、太郎とタカシとヒロシに話しかけてきたのは男性です。

'It was a man who spoke to Taroo,Takasi, and Hirosi the other day.'

B 先日、太郎とタカシとヒロシに話しかけてきたのは女性です。

' It was a woman who spoke to Taroo,Takasi, and Hirosi the other day.'

9.2 C 先週末、花子は両親と一緒に、電車で田舎の伯父に会いに行きました。…

'Hanako and her parents went to see her uncle living in the country by train last weekend. …'

T	伯父は/	夕方から/	{ 花子たち3人を/	*OR*	***3人の花子たちを/**	}	海に/	ドライブに/	連れて行ってくれました。
	ozi-wa	yuugata-kara	{ Hanako-tati san-nin-o	*OR*	**san-nin-no Hanako-tati-o**	}	umi-ni	doraibu-ni	turete-itte-kure-masita
	uncle-TOP	evening-from	{ Hanako-TATI 3-CL-ACC	*OR*	**3-CL-GEN Hanako-TATI-ACC**	}	sea-to	drive-for	take-go-give-POL.PST

'Her uncle took the group of three represented by Hanako.'

Q この文章によると… *(asked only in the post-nominal condition)*

'According to this passage…'

A 花子は先週末、両親と一緒に伯父に会いに行きました。

'Hanako went to see her uncle last weekend with her parents.'

B 花子は先週末、両親と一緒に伯母に会いに行きました。

'Hanako went to see her aunt last weekend with her parents.'

9.3 C 花子は今度の休日、ひろ子と他の友人3人と一緒に過ごす予定です。花子はハイキングに行きたがっていましたが、友人はみんな海に行きたがっていました。…
'Hanako is planning to spend the next holiday with Hiroko and three other friends. Hanako wanted to go on a hiking but her friends all wanted to go to the seaside. …'

T	花子は/	1週間かけて/	{ ひろ子たち4人を/	*OR*	***4人のひろ子たちを/**	}	必死に/	なんとか/	説得しました。
	Hanako-wa	i-syuu-kan kakete	{ Hiroko-tati yo-nin-o	*OR*	**yo-nin-no Hiroko-tati-o**	}	hissini	nantoka	settoku-si-masita
	Hanako-TOP	1-week-period spend	{ Hiroko-TATI 4 CL-ACC	*OR*	**4-CL-GEN Hiroko-TATI-ACC**	}	desperately	somehow	persuasion-do-POL.PST

'Hanako tried hard to persuade the group of four represented by Hiroko over a week and managed it.'

Q この文章によると… *(asked only in the post-nominal condition)*
'According to this passage…'

A 花子は今度の休日はだれかと一緒に過ごす予定です。
'Hanako is going to spend time together with someone on the upcoming holiday.'

B 花子は今度の休日は家で1人でいる予定です。
'Hanako is going to stay at home alone on the upcoming holiday.'

9.4 C 今日、たか子は会社の同僚3人と最近できたカフェにお茶をしに行きました。到着した時、カフェはとても混んでいました。…
'Takako and three colleagues of hers went to a recently opened café for tea. When they got to the café, it was very busy. …'

T	店員は/	約15分後/	{ たか子たち4人を/	*OR*	***4人のたか子たちを/**	}	ていねいに/	席まで/	案内しました。
	tenin-wa	yaku jyuugo-hun-go	{ Takako-tati yo-nin-o	*OR*	**yo-nin-no Takako-tati-o**	}	teineini	seki-made	annai-si-masita
	waiter-TOP	about 15-minite-after	{ Takako-TATI-4-CL-ACC	*OR*	**4-CL-GEN Takako-TATI-ACC**	}	politely	seat-to	guide-do-POL.PST

'About 15 minutes later, a waiter showed the group of four represented by Takako to their table politely.'

Q この文章によると… *(asked only in the post-nominal condition)*
'According to this passage…'

A たか子が今日行ったカフェは混んでいました。
'The café Takako went to today was crowded.'

B 今日、たか子が今日行ったカフェは空いていました。
'The café Takako went to today was not crowded.'

9.5 C 太郎はとても力持ちです。先週、友人のタカシとタカシの2人の弟に、自分がどれだけ重い物を持てるかを見せるために、…

‘Taroo is very strong. Last week, in order to show his friend Takasi and Takasi's two younger brothers how much weight he could lift, …’

T	太郎は/	片方の腕で/	{	タカシたち3人を/	*OR*	***3人のタカシたちを/**	}	一気に/	簡単に/	持ち上げました。
	Taroo-wa	katahoo-no ude-de	{	Takasi-tati san-nin-o	*OR*	***san-nin-no Takasi-tati-o**	}	ikkini	kantan-ni	motiage-masita
	Taroo-TOP	with.one.arm	{	Takasi-TATI 3-CL-ACC	*OR*	**3-CL-GEN Takasi-TATI-ACC**	}	at.one.time	easily	lift-POL.PST

‘Taroo lifted the group of three represented by Takashi with one arm at one easily.’

Q この文章によると… *(asked only in the post-nominal condition)*

‘According to this passage…’

A 太郎は力持ちです。

‘Taroo can lift heavy weights.’

B 太郎は足が速いです。

‘Taroo can run fast.’

9.6 C 昨日、太郎はせっかくの休みでしたが、特に予定がなく、退屈していました。ヒロシとタケシも特に予定がないと聞いていたので、…

‘It was a day off yesterday but Taroo had nothing special to do and was feeling bored. Since he knew that Hirosi and Takesi did not have any plan either, …’

T	太郎は/	午後から/	{	ヒロシたち2人を/	*OR*	***2人のヒロシたちを/**	}	電話で/	家に/	呼びました。
	Taroo-wa	gogo-kara	{	Hirosi-tati huta-ri-o	*OR*	**huta-ri-no Hirosi-tati-o**	}	denwa-de	ie-ni	yobi-masita
	Taroo-TOP	afternoon-from	{	Hirosi-TATI 2-CL-ACC	*OR*	**2-CL-GEN Hirosi-TATI-ACC**	}	phone-by	house-to	call- POL.PST

‘Taroo rang the two, Hirosi and the other (=Takesi), and invited them over to his house to hang out.’

Q この文章によると… *(asked only in the post-nominal condition)*

‘According to this passage…’

A 太郎は昨日、休みでした。

‘Taroo had a day off yesterday.’

B 太郎は昨日、仕事でした。

‘Taroo worked yesterday.’

9.7 C 昨日、小学生のタケシは小さなことで友人のゴロウとヒロシと学校でケンカをしました。…

'Takesi, a primary school kid, fought over a trivial thing with his friends, Goroo and Hirosi at school yesterday. …'

T	タケシは/	怒って/	{ ゴロウたち2人を/	OR	***2人のゴロウたちを/** }	つい/	軽く/	なぐってしまいました。
	Takesi-wa	okotte	{ Goroo-tati huta-ri-o	OR	**huta-ri-no Goroo-tati-o** }	tui	karuku	nagutte-simai-masita
	Takesi-TOP	in.a.rage	{ Goroo-TATI 2-CL-ACC	OR	**2-CL-GEN Goroo-TATI-ACC** }	carelessly	lightly	hit-finish-POL.PST

'Being in a rage, Takesi couldn't help but hit the two, Goroo and the other(=Hirosi) lightly.'

9.8 C 先週、花子の学校では定期テストがありました。友人のとも子と他の友人2人は英語のテストが全然できなかったと落ち込んでいました。…

'It was an exam week at Hanako's school last week. Her friends, Tomoko and two others were disappointed because they did very poorly on the English exam. …'

T	花子は/	気になって/	{ とも子たち3人を/	OR	***3人のとも子たち/** }	そっと/	優しく/	励ました。
	Hanako-wa	ki-ni-natte	{ Tomoko-tati san-nin-o	OR	**san-nin-no Tomoko-tati-o** }	sotto	yasasiku	hagemasi-masita
	Takesi-TOP	worried	{ Tomoko-TATI 3-CL-ACC	OR	**3-CL-GEN Tomoko-TATI-ACC** }	gently	kindly	cheer.up-POL.PST

'Hanako was worried about the three, Tomoko and the others, and cheered them up gently.'

Fillers (10.1–15.4)

10.Filler 1: Missing classifiers

10.1 C 昨日、太郎は何もすることが無かったので、家の近所を散歩することにしました。…

'Since he had nothing particular to do yesterday, he took a walk around his neighbourhood. …'

T	太郎は/	散歩中/	***近所の子どもを2/**	公園で/	偶然/	見ました。
	Taroo-wa	sanpo-tyuu	**kinjyo-no kodomo-o ni**	kooen-de	guuzen	mi-masita
	Taroo-TOP	walk-during	**neightbourhood-GEN child-ACC 2**	park-in	by.chance	see-POL.PST

'While taking a walk, he happened to see two neighbourhood children in the park.'

(*The CL, *nin* is missing.)

10.2 C 太郎は昔から猫が大好きです。 …

‘Taroo has loved cats since he was small. …’

T	太郎は/	小学生の時から/	***猫を3/**	家の中で/	10年以上/	飼っています。
	Taroo-wa	syoogakusei-no toki kara	***neko-o san**	ie-no naka-de	jyuu-nen-ijyoo	katte-i-masu
	Taroo-TOP	since.primary.school	**cat-ACC 3**	inside.the.house	10-year more.than	keep-ASP-POL.NPST

‘Taroo has kept three cats inside his house for more than ten years since he was in primary school.’

(*The CL, *hiki* is missing.)

10.3 C 花子は甘いものが大好きで、毎日、甘いものを少し食べるのが楽しみです。…

‘Hanako has a sweet tooth and takes a pleasure in eating sweets moderately.…’

T	花子は/	今日も/	***クッキーを2/**	おやつに/	味わって/	食べました。
	Hanako-wa	kyoo mo	**kukkii-o ni**	oyatu-ni	aziwatte	tabe-masita
	Hanako-TOP	today too	**cookie-ACC 2**	snack-for	taste	eat-POL.PST

‘Hanako enjoyed tasting two cookies for snack today, too.’

(*The CL, *hiki* is missing.)

10.4 C 花子は先日の休日に特にすることが無く暇でした。…

‘Hanako was bored having nothing particular to do on the last holiday. …’

T	花子は/	思いつきで/	***友達を3/**	メールで/	映画に/	誘いました。
	Hanako-wa	omoituki-de	**tomodati-o san**	meeru-de	eiga-ni	sasoi-masita
	Hanako-TOP	on.impulse	**friend-ACC 3**	mail-by	movie-for	ask.out-POL.PST

‘Hanako asked three friends out for a movie by email on a whim.’

(*The CL, *nin* is missing.)

11. Filler 2: -tati *suffixed to non-human nouns*

(*-*tati* can be suffixed to only human nouns.)

11.1 C 太郎は先週末、図書館で本をたくさん借りました。…

'Taroo borrowed a lot of books from the library last weekend. …'

T	太郎は/	寝る前に/	***本たちを**/	少しずつ/	じっくり/	読んでいます。
	Taroo-wa	neru mae-ni	***hon-tati-o**	sukosizutu	zikkuri	yonde-i-masu
	Taroo-TOP	before.going.to.sleep	**book-TATI-ACC**	little.by.little	slowly	read-ASP-POL.NPST

'Taroo is enjoying reading the books little by little before going to sleep.'

11.2 C 花子は毎日、朝ごはんにリンゴを食べていて、いつもスーパーに行く時は、必ず買うようにしています。しかし、…

'Hanako eats an apple every morning and makes it a habit to buy apples whenever she goes to a supermarket. But …'

T	花子は/	昨日/	***りんごたちを**/	スーパーで/	うっかり/	買い忘れました。
	Hanako-wa	kinoo	**ringo-tati-o**	suupaa-de	ukkuri	kai-wasure-masita
	Hanako-TOP	yesterday	**apple-TATI-ACC**	supermarket-at	carelessly	buy-forget-POL.PST

'Hanako carelessly forgot to buy apples at a supermarket yesterday.'

11.3 C　花子の家にはテレビがたくさんあります。しかし、…

'There are many TVs in Hanako's house. But …'

T	花子は/	昨夜/	***テレビたちを**/	寝る前に/	うっかり/	消し忘れました。
	Hanako-wa	sakuya	**terebi-tati-o**	neru mae-ni	ukkuri	kesi-wasure-masita
	Hanako-TOP	last.night	**TV-TATI-ACC**	before.going.to.sleep	carelessly	turn.off-forget-POL.PST

'Hanako carelessly forgot to switch off the TVs before going to bed.'

11.4 C　太郎は車を2台持っていて、毎週末、洗車しています。…

'Taroo owns two cars and wash them every weekend. …'

T	太郎は/	先週末も/	***車たちを**/	洗車場で/	ていねいに/	洗いました。
	Taroo-wa	sensyuu-matu-mo	**kuruma-tati-o**	sensyajyoo-de	teineini	arai-masita.
	Taroo-TOP	last.week.end too	**car-TATI-ACC**	car.wash-at	carelessly	wash- POL.PST

'Taroo washed the cars carefully at the car wash last weekend again.'

12. Filler 3: Wrong particles

12.1 C　太郎は勉強するのに、よく近所の図書館を利用します。…

'Taroo often studies at a library nearby his house. …'

T	太郎は/	今日も/	***その図書館に**/	長い時間/	ひとりで/	勉強をしました。
	Taroo-wa	kyoo-mo	**sono tosyokan-ni**	nagai jikan	hito-ri-de	benkyoo-o si-masita
	Taroo-TOP	today-too	**SONO library-to**	for.a.long.time	alone	study-ACC do-POL.PST

'Taroo studied at the library for a long time alone today again.'

(* The particle, *-de* must be used.)

12.2 C　花子はアメリカに住んでいる友人がいます。先日、仕事でアメリカを訪れた時、…

'Hanako has a friend living in the U.S. When she was on a bussiness trip to the U.S. the other day, …'

T	花子は/	久しぶりに/	***その友人を/**	数時間だけ/	空港で/	会いました。
	Hanako-wa	hisasiburi-ni	**sono yuuzin-o**	suu-zikan-dake	kuukoo-de	ai-masita
	Hanako-TOP	for.the.first.time.in.a.while	**SONO friend-ACC**	a.few-hour-only	airport-at	meet-POL.PST

'Hanako met up with the friend at an airport for just a few hours for the first time in a while.'

(* The particle, -*ni* must be used.)

12.3 C　太郎は昨日から始まったテレビドラマをとても楽しんで見ました。…

'Taroo really enjoyed a TV drama that started yesterday. …'

T	太郎は/	来週も/	***そのドラマに/**	必ず/	また/	見るつもりです。
	Taroo-wa	raisyuu-mo	**sono dorama-ni**	kanarazu	mata	miru-tumori-desu
	Taroo-TOP	next.week-too	**SONO drama-to**	always	again	watch-intention-COP

'Taroo is definitely going to watch the drama next week again.'

(*The particle, -*o* must be used.)

12.4 C 花子は自分の携帯電話の番号をなかなか覚えられません。昨日、友達に番号を聞かれましたが、…

'Hanako has trouble memorising her own mobile phone number. When a friend asked for her number yesterday, …'

T	花子は/	いつものように/	***番号に/**	どうしても/	全然/	思い出せませんでした。
	Hanako-wa	itumo-no yooni	**bangoo-ni**	doo-site-mo	zenzen	omoidas-e-masen-desita
	Hanako-TOP	as.always	**number-to**	by.any.means	at.all	could.not.remember

'Hanako, as always, just could not remember the number.'

(*The particle, *-o* must be used.)

13. Filler 4: Wrong tense marking

13.1 C 花子は昨日、近くのショッピングモールへ買い物に出かけました。ゆっくりいろんなお店を回っていると、…

'Hanako went to the shopping mall near her house yesterday. When she was taking her time to shop around, …'

T	花子は/	久しぶりに/	ひろ子を/	服屋で/	たまたま/	***見かけます。**
	Hanako-wa	hisasiburi-ni	Hiroko-o	hukuya-de	tamatama	**mikake-masu**
	Hanako-TOP	for.the.first.time.in.a.while	Hiroko-ACC	clothes.shop-at	accidentally	**catch.sight-POL.NPST**

'Hanako accidentally caught sight of Hiroko at a clothing shop for the first time in a while.'

(*The past tense form, *mikake-masita* must be used.)

13.2 C 太郎は動物を飼うのが大好きです。これまでいろんなペットを飼ってきました。例えば、…

'Taroo loves having pet animals. He has had many kinds of pets. For example, …'

T	太郎は/	小学生の時/	ウサギを/	庭で/	４年間/	***飼っています。**
	Taroo-wa	syoogakusei-no toki	usagi-o	niwa-de	yo-nen-kan	**katte-i-masu**
	Taroo-TOP	when.in.primary.school	rabbit-ACC	garden-in	for.4.years	**keep-ASP-NPST**

'Taroo had a rabbit in the garden for four years when he was in primary school.'

(*The past tense form, *katte-i-masita* must be used.)

13.3 C　花子は昨夜、近所に住んでいる小学生の女の子がひとりで歩いているのを見ました。心配だったので、…

'Hanako found a primary school girl who lives in her neighbourhood walking alone last night. Because she was worried about the girl, …'

T	花子は/	優しく/	その女の子を/	家まで/	車で/	***送っていきます。**
	Hanako-wa	yasiku	sono onnanoko-o	ie-made	kuruma-de	**okutte-iki-masu**
	Hanako-TOP	kindly	SONO girl-ACC	house-to	car-by	**see.off-go-NPST**

'Hanako kindly drove the girl home.'

(*The past tense form, *okutte-iki-masita* must be used.)

13.4 C　太郎は来月、友人と夕食を一緒に食べに行く約束をしました。行きたいレストランはとても人気なので、…

'Taroo made a promise with his friend to go out for dinner together next month. Because the restaurant they want to go to is very popular, …'

T	太郎は/	早速/	そのレストランを/	もう/	電話で/	***予約します。**
	Taroo-wa	sassoku	sono resutoran-o	moo	denwa-de	**yoyaku-si-masu**
	Taroo-TOP	immediately	SONO restaurant-ACC	already	phone-by	**reservation-do-NPST**

'Taroo already called the restaurant and reserved it.'

(*The past tense form, *yoyaku-si-masita* must be used.)

14. Filler 5: Infelicitous wh-floating quantifiers

14.1 C 花子が好きな俳優が2年後に公開される予定の映画の主役に選ばれました。…

'Hanako's favourite actor has been chosen to star in a movie to be released in two years. …'

T	花子は/	今から/	その映画を/	***何か**/	とても/	楽しみにしています。
	Hanako-wa	ima-kara	sono eiga-o	**nani-ka**	totemo	tamosimi-ni-site-i-masu
	Hanako-TOP	now-from	SONO movie-ACC	**something**	very much	looking.forward.to

'Hanako is already really looking forward to watching the movie.'

(* *Nanika* requires the associated noun, *sono eiga* to be nonspecific thus incompatible with the specific meaning forced by *sono*)

14.2 C 太郎は毎日、近所の喫茶店に行き、必ず同じコーヒーを注文します。…

'Taroo goes to a café near his house and orders the same coffee every day. …'

T	太郎は/	今日も/	そのコーヒーを/	***何か**/	いつものように/	飲みました。
	Taroo-wa	kyoo-mo	sono koohii-o	**nani-ka**	itumo-no yooni	nomi-masita
	Taroo-TOP	today-too	SONO coffee-ACC	**something**	as.always	drink-POL.NPST

'Taroo, as always, drank the coffee today.'

(* *Nanika* requires the associated noun, *sono koohii* to be nonspecific thus incompatible with the specific meaning forced by *sono*)

14.3 C 花子は昨日、仕事の帰り道に、電車の駅を探している中学生の男の子に出会いました。…

'Hanako came across a junior high school boy who was looking for a train station on the way home from work yesterday. …'

T	花子は/	親切に/	その男の子を/	***誰か/**	駅まで/	案内してあげました。
	Hanako-wa	sinsetuni	sono otokonoko-o	**dare-ka**	eki-made	annai-site-age-masita
	Hanako-TOP	kindly	SONO boy-ACC	**someone**	station-to	guide-do-give-POL.PST

'Hanako kindly guided the boy to the station.'

(* *Dare-ka* requires the associated noun, *sono otokonoko* to be nonspecific thus incompatible with the specific meaning forced by *sono*)

14.4 C 太郎は小学生の時の担任の先生にあこがれて、自分も教師になりました。…

'Taroo adored one of his classroom teachers from primary school so much that he became a teacher himself. …'

T	太郎は/	今でも/	その先生を/	***誰か/**	とても/	尊敬しています。
	Taroo-wa	ima-demo	sono sensei-o	**dare-ka**	totemo	sonkei-site-i-masu
	Taroo-TOP	now-even	SONO teacher-ACC	**someone**	very much	respect

'Taroo still respects the teacher very much.'

(* *Dare-ka* requires the associated noun, *sono sensei* to be nonspecific thus incompatible with the specific meaning forced by *sono*)

15. Filler 6: Infelicitous use of demonstrative ano

(**Sono* NPs are contextually more natural than *ano* NPs.)

15.1 C 太郎は先日、友人の家で、とても面白いテレビドラマを見ました。しかし、…

‘Taroo watched a very amusing TV drama at his friend's house the other day. But … ’

T	太郎は/	すでに/	***あのドラマの内容を/**	もう/	ほとんど/	忘れてしまいました。
	Taroo-wa	sudeni	**ano dorama-no naiyoo-o**	moo	hotondo	wasurete-simai-masita
	Taroo-TOP	already	**ANO drama-GEN content-ACC**	already	mostly	forget-finish-POL.PST

‘Taroo has already mostly forgot what the drama was like.’

15.2 C 花子は5歳の誕生日にもらったぬいぐるみがとても気に入っています。…

‘Hanako really loves a soft toy that she got for her 5th birthday. … ’

T	花子は/	どこへでも/	***あのぬいぐるみを/**	必ず/	一緒に/	連れて行きます。
	Hanako-wa	doko-e-demo	**ano nuigurumi-o**	kanarazu	issyoni	turete-iki-masu
	Taroo-TOP	anywhere	**ANO soft.toy-ACC**	always	toghether	take-go-POL.NPST

‘Hanako takes the soft toy with her wherever she goes.’

15.3 C 太郎の職場には山田というよくウソをつく人がいます。…

‘There is a man who often lies named Yamada at Taroo's workplace. ... ’

T	太郎は/	どんな時も/	***あの人を**/	絶対に/	簡単には/	信用しません。
	Taroo-wa	donna toki-mo	**ano hito-o**	zettaini	kantanni-wa	sinyoo-si-masen
	Taroo-TOP	at.any.time	**ANO person-ACC**	by.any.means	easily-TOP	trust-do- POL.NPST

‘Taroo never trusts the person easily.’

15.4 C 花子は先日、道に迷って困っている大学生に出会いました。特に急いでいなかったので、…

‘Hanako found a university student being lost on the street the other day. Because she was not in a hurry, ... ’

T	花子は/	親切に/	***あの学生を**/	行きたい場所まで/	歩いて/	案内してあげました。
	Hanako-wa	sinsetuni	***ano gakusei-o**	iki-tai basyo-made	aruite	annai-site-age-masita
	Hanako-TOP	kindly	**ANO student-ACC**	go-want place-to	walk	guide-do-give-POL.PST

‘Taroo never trusts the person easily.’

Abbreviations

∧	Conjunction
∩	The nominalisation operator
≤	Part-of relation
∅	Zero, null/overt form
ACC	Accusative
ACP	The Article Choice Parameter
adj.	Adjunct
ADJ/adj	Adjective
ADV	Adverb
AJT	Acceptability judgement task
ANOVA	Analysis of variance
Appdx	Appendix
ASP	Aspect
C	Context sentence
CASE	Case marker
CaseP	Case phrase
CEFR	The Common European Frame of Reference for Languages
CI	Confidence Interval
CL	Classifier
CLP	Classifier phrase
COMP	Complementiser
COP	Copula
D	Determiner
D’	D-bar (an intermediate projection of D)

DAT	Dative
DEC(L)	Declarative
def	definite
DP	Determiner phrase
dr	domain restriction
E	English
e	Empty category
ERP	Event-related potential
et	Exception tolerance
F	Feature
F_{covert}	Covertly realised feature
FH	The Fluctuation Hypothesis
F_{overt}	Overtly realised feature
FRH	The Feature Reassembly Hypothesis
GEN	Genitive
HK	Hearer's knowledge
HSD	Tukey's honestly significant difference
I-level	Individual-level
IMPERATIVE	Imperative marker
INAN.OBJ	Inanimate object
J	Japanese
JLPT	the Japanese Language Proficiency Test
K	Korean
L1	First language
L2	Second language
LVL	Level

M	Mean
M age	Mean age
M-Diff	Mean difference
ms	Millisecond
N	Noun
n	Number of cases
N'	N-bar (an intermediate projection of N)
N/A	Not applicable
NEG	Negative
NOM	Nominative
NP	Noun phrase
NPST	Non-past tense
NQ	Numeral quantifier
Num	Number
NumP	Number phrase
p	Probability
PA(S)T	Past tense
PIC	The Phase Impenetrability Condition
PL	Plural
POL	Polite marker
POS	Poverty of the stimulus
PRE	Present tense
Q	Quantifier/Comprehension question
QP	Quantifier phrase
RQ	Research question
RRT	Residualised reading time

RT	Reading time
SD	Standard deviation
SE	Standard error
SLA	Second language acquisition
S-level	Stage-level
Spec	Specifier
spec	specific
SPRT	Self-paced reading task
SR	Specific referent in the mind of the speaker’s
t	Trace/Student’s *t* distribution
T	Tense/Target sentence
the_{weak}	Weak definite article
$\text{the}_{\text{strong}}$	Strong definite article
TOP	Topic
TVJT	Truth-value judgement task
UG	Universal Grammar
V	Verb
v	Little/Small verb
V’	V-bar (an intermediate projection of V)
VP	Verb phrase
X	Minimal projection
XP	Maximal projection
yrs	Years
z	Standardised score
$\exists$	The existential operator
β	Beta coefficient

ι	The iota operator
χ^2	The Chi-square distribution

References

3A Network. (1998). *Minna no Nihongo* series. Tokyo: 3A Corporation.

Ahn, D. (2017). Definite and demonstrative descriptions: A micro-typology. In M. Y. Erlewine (Ed.), *Proceedings of GLOW in Asia XI–Volume 1* (pp. 33–47). Cambridge, MA: MIT Working Papers in Linguistics.

Anderssen, M. (2007). The acquisition of compositional definiteness in Norwegian. *Papers from the Language Acquisition Workshop*, *34*(3), 252–275.

Bader, M., & Lasser, I. (1994). German verb-final clauses and sentence processing: Evidence for immediate attachment. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 225–242). New York, NY: Lawrence Erlbaum Associates.

Banno, E., Ohno, Y., Sakane, Y., Shinagawa, C., & Takashiki, K. (1999). *Genki: An integrated elementary Japanese course*. Tokyo: The Japan Times.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Belletti, A., Bennati, E., & Sorace, A. (2007). Theoretical and developmental issues in the syntax of subjects: Evidence from near-native Italian. *Natural Language & Linguistic Theory*, *25*(4), 657–689

Bickerton, D. (1981). *Roots of language*. Ann Arbor, MI: Karoma.

Boyd, J. K., & Goldberg, A. E. (2011). Learning what not to say: The role of statistical preemption and categorization in a-adjective production. *Language*, *87*, 55–83.

Campbell, R. (1996). Specificity operators in Spec DP. *Studia Linguistica*, *50*(2), 161–188.

Chang, S. J. (2009). *Nominal structure and interpretation: On the syntax of the Korean* *determiner phrase* (Unpublished doctoral dissertation). University of Georgia.

Charolles, M. (1999). Associative anaphora and its interpretation. *Journal of Pragmatics*, *31*(3), 311–326.

Chierchia, G. (1998). Reference to kinds across language. *Natural Language Semantics*, *6*(4), 339-405.

Choi, M. H., & Lardiere, D. (2005). *The interpretation of wh-in-situ in Korean second language acquisition.* In A. Belletti et al. (Eds.), Language Acquisition and Development: *Proceedings of GALA 2005,* (pp. 125–135). Cambridge: Cambridge Scholars Press.

Cho, J. (2012). *Remapping nominal features in the second language* (Unpublished doctoral dissertation). University of Iowa.

Cho, J. (2017). The acquisition of different types of definite noun phrases in L2-English. *International Journal of Bilingualism*, *21*(3), 367–382.

Cho, J. (2020). Online processing and offline judgments of L2-English articles. *Linguistic Approaches to Bilingualism*. Advance online publication. https://doi.org/10.1075/lab.18053.cho

Cho, J., & Slabakova, R. (2014). Interpreting definiteness in a second language without articles: The case of L2 Russian. *Second Language Research*, *30*(2), 159–190.

Cho, J., & Slabakova, R. (2015). A feature-based contrastive approach to the L2 Acquisition of Specificity. *Applied Linguistics*, *38*(3), 318–339.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. (2000). Minimalist inquiries: The framework. In R. Martin, et al. (Eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik,* (pp. 89–155). Cambridge, MA: MIT Press.

Chomsky, N. (2004). Beyond explanatory adequacy. In A. Belletti (Ed.), *Structures and beyond* (pp. 104–131): Oxford: Oxford University Press.

Christensen, R. (2018). *Ordinal—regression models for ordinal data*. R package version 2018.8-25, http://www.cran.r- project.org/ package=ordinal/.

Clark, H. H. (1975). Bridging. In R. C. Schank & B. L. Nash-Webber (Eds.), *Proceedings of the 1975 workshop on theoretical issues in natural language processing* (pp.169–174). New York, NY: Association for Computing Machinery.

Council of Europe (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

Cowart, W. (1997). *Experimental syntax*. Thousand Oaks, CA: SAGE Publishing.

Crosthwaite, P. R. (2014). Definite discourse–new reference in L1 and L2: A study of bridging in Mandarin, Korean, and English. *Language Learning*, *64*(3), 456–492.

Crosthwaite, P. R., Yeung, Y., Bai, X., Lu, L., & Bae, Y. (2018). Definite discourse-new reference in L1 and L2: The case of L2 Mandarin. *Studies in Second Language Acquisition*, *40*(3), 625–649.

Dekydtspotter, L., & Sprouse, R. A. (2001). Mental design and (second) language epistemology: Adjectival restrictions of wh-quantifiers and tense in English-French interlanguage. *Second Language Research, 17*(1), 1–35.

Dekydtspotter, L., Sprouse, R. A., & Anderson, B. (1997). The interpretive interface in L2 acquisition: The process-result distinction in English-French interlanguage grammars. *Language Acquisition, 6*(4), 297–332.

Diesing, M. (1992). *Indefinites*. Cambridge, MA: MIT press.

Dillion, B., & Wagers, M. (forthcoming). Approaching gradience in acceptability with the tools of signal detection theory. In G. Goodall (Ed.), *Cambridge Handbook of Experimental Syntax*.

Domínguez, L., Arche, M. J., & Myles, F. (2017). Spanish imperfect revisited: Exploring L1 influence in the reassembly of imperfective features onto new L2 forms. *Second Language Research, 33*(4), 431–457.

Downing, P. A. (1996). *Numeral classifier systems: The case of Japanese*. Amsterdam: John Benjamins.

Ebert, K. (1971). Zwei formen des bestimmten artikels. In D. Wunderlich (Ed.), *Probleme und fortschritte der transformationsgrammatik* (pp. 159–74). Munich: Max Hueber.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, *27*(2), 141–172.

Ellis, R. (2006). Modelling learning difficulty and second language proficiency: The differential contributions of implicit and explicit knowledge. *Applied Linguistics, 27*(3), 431–463.

Enç, M. (1991). The semantics of specificity. *Linguistic inquiry, 22*(1), 1–25.

Feng, S. (2019). The acquisition of English definite noun phrases by Mandarin Chinese speakers. *Studies in Second Language Acquisition, 41*(4), 881–896.

Ferreira, F., & Clifton Jr, C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, *25*(3), 348–368.

Foote, R. (2011). Integrated knowledge of agreement in early and late English-Spanish bilinguals. *Applied Psycholinguistics, 32*(1), 187–220.

Frege, G. (1960). On sense and reference. In P. Geach & M. Black (Eds.), *Translations from the philosophical writings of Gottlob Frege* (pp. 56–78). Oxford: Basil Blackwell.

Fukushima, K. (1991). *Generalized floating quantifiers* (Unpublished doctoral dissertation). University of Arizona.

Furuya, K. (2012). Specificity effects for Japanese, an articleless language. *Studies in the Linguistic Sciences: Illinois Working Papers 2012*, 32–43.

Gabriele, A. (2005). *The acquisition of aspect in a second language: A bidirectional study of learners of English and Japanese* (Unpublished doctoral dissertation). City University of New York.

Gabriele, A. (2008). *Calculating telicity in native and non-native English.* In R. Slabakova, J.

Rothman, O. Kempchinsky, & E. Gavruseva (Eds.), *Proceedings of the 9th Generative Approaches to Second Language Acquisition (GASLA 9)* (pp. 37–46). Somerville, MA: Cascadilla Proceedings Project.

Gabriele, A. (2009). Transfer and transition in the SLA of aspect: A bidirectional study of learners of English and Japanese. *Studies in Second Language Acquisition, 31*(3), 371–402.

Gabriele, A. (2010). Deriving meaning through context: Interpreting bare nominals in second language Japanese. *Second Language Research, 26*(3), 379–405.

Gil, K.-H., & Marsden, H. (2013). Existential quantifiers in second language acquisition: A feature reassembly account. *Linguistic Approaches to Bilingualism, 3*(2), 117–149.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Studies in syntax and semantics III: Speech acts* (pp. 183–198). New York, NY: Academic Press.

Havik, E., Roberts, L., Van Hout, R., Schreuder, R., & Haverkort, M. (2009). Processing subject-object ambiguities in the L2: A self-paced reading study with German L2 learners of Dutch. *Language Learning, 59*(1), 73–112.

Hawkins, J. A. (1978). *Definiteness and indefiniteness*. Atlantic Highlands: Humanities Press.

Hawkins, R. (2019). *How second languages are learned: an introduction*. Cambridge: Cambridge University Press.

Hawkins, R., Al-Eid, S., Almahboob, I., Athanasopoulos, P., Chaengchenkit, R., Hu, J., Rezai, M., Jaensch, C., Jeon, Y., Jiang, A., Yan-kit, L.I., Matsunaga, K., Ortega, M., Sarko, G., Snape, N., & Velasco-Zárate, K. (2006). Accounting for English article interpretation by L2 speakers. *EUROSLA Yearbook, 6*, 7–25.

Hawkins, R., & Hattori, H. (2006). Interpretation of English multiple wh-questions by Japanese speakers: A missing uninterpretable feature account. *Second Language Research, 22*(3), 269–301.

Heim, I. (1982). *The semantics of definite and indefinite noun phrases* (Doctoral dissertation).

Retrieved from ProQuest Dissertations and Theses database. AAI8229562

Heim, I. (1991). Articles and definiteness, published in German as ‘Artikel und definitheit’. In A.V. Stechow & D. Wunderlich (Eds.), *Semantics: An international handbook of contemporary research* (pp. 487–535). Berlin: De Gruyter.

Heim, I., & Wexler, K. (2000). *Class notes for "Topics in semantics and acquisition"*. Unpublished manuscript, Massachusetts Institute of Technology.

Heycock, C. (2008). Japanese *-wa*, *-ga*, and information structure. In S. Miyagawa, & M. Saito (Eds.), *Handbook of Japanese Linguistics* (pp. 54–83). Oxford: Oxford University Press.

Hoji, H., Kinsui, S., Takubo, Y., & Ueyama, A. (2003). The demonstratives in modern Japanese. In Y-A. H. Li & Y. Simpson (Eds.), *Functional structure(s), form and interpretation: Perspectives from east Asian languages* (pp. 97–128). New York, NY: Routledge.

Hopp, H. (2009). The syntax-discourse interface in near-native L2 acquisition: Off-line and on-line performance. *Bilingualism: Language and Cognition*, *12*(4), 463–483.

Hopp, H. (2010). Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua, 120*(4), 901–931.

Huang, C. T. J., & Ochi, M. (2014). Remarks on classifiers and nominal structure in East Asian. In *Peaches and plums* (pp. 53–74). Taipei: Academia Sinica.

Huebner, T. (1983). *A longitudinal analysis of the acquisition of English*. Ann Arbor, MI: Karoma.

Hwang, S. H., & Lardiere, D. (2013). Plural-marking in L2 Korean: A feature-based approach. *Second Language Research, 29*(1), 57–86.

Hyönä, J., & Hujanen, H. (1997). Effects of case marking and word order on sentence parsing in Finnish: An eye fixation analysis. *The Quarterly Journal of Experimental Psychology: Section A*, *50*(4), 841–858.

Inagaki, S. (2001). Motion verbs with goal PPs in the L2 acquisition of English and Japanese. *Studies in Second Language Acquisition*, *23*(2), 153–170.

Inoue, K. (1978). *Nihongo-no bunpoo kisoku* [Grammar rules in Japanese]. Tokyo: Taisyukan.

Ionin, T. (2003). *Article semantics in second language acquisition* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.

Ionin, T., Baek, S., Kim, E., Ko, H., & Wexler, K. (2012). That's not so different from the: Definite and demonstrative descriptions in second language acquisition. *Second Language Research*, *28*(1), 69–101.

Ionin, T., Choi, S. H., & Liu, Q. (2019). Knowledge of indefinite articles in L2-English: Online vs. offline performance. *Second Language Research*. Advance online publication. https://doi.org/10.1177%2 F0267658319857466

Ionin, T., Choi, S. H., & Liu, Q. (2020). What transfers (or doesn't) in the second language acquisition of English articles by learners from article-less native languages? *Linguistic Approaches to Bilingualism*. Advance online publication. https://doi.org/10.1075/lab.19057.ion

Ionin, T., Ko, H., & Wexler, K. (2004). Article semantics in L2 acquisition: The role of specificity. *Language Acquisition*, *12*(1), 3–69.

Ionin, T., Montrul, S., & Crivos, M. (2013). A bidirectional study on the acquisition of plural NP interpretation in English and Spanish. *Applied Psycholinguistics, 34*(3), 483–518.

Ionin, T., Zubizarreta, M. L., & Bautista-Maldonado, S. (2008). Sources of linguistic knowledge in the second language acquisition of English articles. *Lingua*, *118*(4), 554–576.

Irani, A. (2019). On (in)definite expressions in American Sign Language. In A.-G. Ana, J. P. Loyo, & V. V.-R. Maldonado (Eds.), *Definiteness across languages* (pp. 113–151). Berlin: Language Science Press.

Ishii, Y. (1999). A note on floating quantifiers in Japanese. In M. Masatake & I. Enoch (Eds.), *Linguistics: In search of the human mind: A festschrift for Kazuko Inoue* (pp. 236–267). Tokyo: Kaitakusha.

Ivanov, I. (2009). *Second language acquisition of Bulgarian object clitics: A test case for the Interface Hypothesis* (Unpublished doctoral dissertation). University of Iowa.

Jackendoff, R. (2002). *Foundations of language*. Oxford: Oxford University Press.

Jegerski, J. (2014). Self-paced reading. In J. Jegerski & B. van Pattern (Eds.), *Research methods in second language psycholinguistics* (pp. 20–49). New York, NY: Routledge.

Jenks, P. (2015). Two kinds of definites in numeral classifier languages. *Semantics and Linguistic Theory*, *25*, 103–124.

Jenks, P. (2018). Articulated definiteness without articles. *Linguistic Inquiry, 49*(3), 501–536.

Jiang, N. (2012). *Conducting reaction time research in second language studies*: New York, NY: Routledge.

Jiang, N. (2018). *Second language processing: An introduction*: New York, NY: Routledge.

Jiang, N., Novokshanova, E., Masuda, K., & Wang, X. (2011). Morphological congruency and the acquisition of L2 morphemes. *Language Learning, 61*(3), 940–967.

Jonz, J. (1990). Another turn in the conversation: What does cloze measure? *Tesol Quarterly, 24*(1), 61–83.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, *111*(2), 228–238.

Juzek, T. S. (2015). *Acceptability Judgement Tasks and Grammatical Theory* (Unpublished doctoral dissertation). University of Oxford.

Kahraman, B., Sato, A., Ono, H., & Sakai, H. (2010). Relative clauses processing before the head noun: Evidence for strong forward prediction in Turkish. In H. Maezawa & A. Yokogoshi (Eds.), *Proceedings of the 6th Workshop on Altaic Formal Linguistics*

(WAFL6). MIT Working Papers in Linguistics 61 (pp. 155–170). Cambridge, MA: MIT Press.

Kamide, Y., & Mitchell, D. C. (1999). Incremental pre-head attachment in Japanese parsing. *Language and Cognitive Processes, 14*(5/6), 631–662.

Kamp, H. (1981). A theory of truth and semantic representation. In P. Portner & B. Partee (Eds.), *Formal semantics: The essential readings* (pp. 189–222). Oxford: Blackwell.

Kaneko, M. (2014). The semantics and syntax of Japanese adnominal demonstratives. In P. Cabredo Hofherr & A. Zribi-Hertz (Eds.), *Crosslinguistic studies of noun phrase structure and reference* (pp. 239–268). Leiden: Brill.

Kang, A. (2021). Marking definiteness in articleless language: The role of the domain restrictor KU in Korean. *Language and Linguistics, 22*(2), 302–337.

Katz, J. J. (1976). A hypothesis about the uniqueness of natural language. *Annals of the New York Academy of Sciences, 280*(1), 33–41.

Keating, G. D., & Jegerski, J. (2015). Experimental designs in sentence processing research: A methodological review and user's guide. *Studies in Second Language Acquisition, 37*(1), 1–32.

Kim, A. H. O. (1995). Word order at the noun phrase level in Japanese. In P. A. Downing & M. Noonan (Eds.), *Word order in discourse* (pp. 199–246). Amsterdam: John Benjamins.

Kim, C. (2005). Order and meaning: Numeral classifiers and specificity in Korean. In J. Alderte, C-H Han, & A. Kochetov (Eds.), *Proceedings of the 24th West Coast Conference on Formal Linguistics* (pp. 218–226). Somerville, MA: Cascadilla Press.

Kim J. B., & Yang, J. (2006a). Deep processing of Korean floating quantifier constructions. In Y. Matsumoto, R.W. Sproat, K.F. Wong, & M. Zhang (Eds.), *Computer Processing of Oriental Languages. Beyond the Orient: The Research Challenges Ahead. ICCPOL 2006. Lecture Notes in Computer Science, vol 4285* (pp. 387–394).

Kim, J. B., & Yang, J. (2006b). Processing Korean numeral classifier constructions in a typed feature structure grammar. In P. Sojka, I. Kopeček, & K. Pala (Eds.), *Text, Speech and Dialogue. TSD 2006. Lecture Notes in Computer Science, vol 4188* (pp. 103–110).

Kim, K. L., & Lakshmanan, U. (2009). The processing role of the Article Choice Parameter evidence from L2 learners of English. In M. del P. G. Mayo & R. Hawkins (Eds.), *Second language acquisition of articles: Empirical findings and theoretical implications* (p.p. 87– 113). Amsterdam: John Benjamins.

Kitagawa, Y., & Kuroda, S.-Y. (1992). *Passive in Japanese*. Unpublished manuscript, University of Rochester and University of California. San Diego, CA.

Kobayashi, M., & Yoshimoto, K. (2001). *A Parallel Interpretation of floated quantifiers and adverbials.* Paper presented at the Proceedings of the 15th Pacific Asia Conference on Language, Information and Computation, Hong Kong.

Kobuchi-Philip, M. (2007). Floating numerals and floating quantifiers. *Lingua*, *117*(5), 814–831.

Ko, H., Ionin, T., & Wexler, K. (2010). The role of presuppositionality in the second language acquisition of English articles. *Linguistic inquiry, 41*(2), 213–254.

Ko, H., Perovic, A., Ionin, T., & Wexler, K. (2008). Semantic universals and variation in L2 article choice. In R. Slabakova, J. Rothman, O. Kempchinsky, & E. Gavruseva (Eds.), *Proceedings of the 9th Generative Approaches to Second Language Acquisition (GASLA 9)* (pp. 118–129). Somerville, MA: Cascadilla Proceedings Project.

Köylü, Y. (2019). *Representation and acquisition of kind reference in L2 English* (Unpublished doctoral dissertation). Indiana University.

Kratzer, A. (1995). Stage-level predicates and individual-level predicates. In G. N. Carlson & F. J. Pelletier (Eds.), *The generic book* (pp. 125–G175). Chicago, IL: University of Chicago Press.

Krifka, M. (1999). At least some determiners aren't determiners. In K. Turner (Ed.), *The* *semantics/pragmatics interface from different points of view* (pp. 257–291). Amsterdam: Elsevier.

Kuno, S. (1973). *The structure of the Japanese language*. Cambridge, MA: MIT press.

Kuroda, S.-Y. (1972). The categorical and thetic judgement. *Foundations of Language, 9,* 153–185.

Kuroda, S.-Y. (2005). Focusing on the matter of topic: A study of *wa* and *ga* in Japanese. *Journal of East Asian Linguistics, 14*(1), 1–58.

Lago, S., Shalom, D. E., Sigman, M., Lau, E. F., & Phillips, C. (2015). Agreement processes in Spanish comprehension. *Journal of Memory and Language, 82*, 133–149.

Lardiere, D. (2008). Feature assembly in second language acquisition. In J. M. Liceras, H. Zobl, & H. Goodluck, (Eds.), *The role of formal features in second language acquisition* (pp. 106–140). New York, NY: Lawrence Erlbaum Associates.

Lardiere, D. (2009). Some thoughts on the contrastive analysis of features in second language acquisition. *Second Language Research*, *25*(2), 173–227.

Lee, C. (2001). Contrastive topic and polarity in discourse. In A. M. Di Sciullo (Ed.), *Asymmetry in grammar* (pp. 345–371). Amsterdam: John Benjamins.

Lee, E., & Lardiere, D. (2019). Feature reassembly in the acquisition of plural marking by Korean and Indonesian bilinguals. *Linguistic Approaches to Billingualism*, *9*(1), 73–119.

Lee, E., & Shimojo, M. (2016). Mismatch of topic between Japanese and Korean. *Journal of East Asian Linguistics*, *25*(1), 81–112.

Lee, H. Y. (2013). Definiteness effects in Korean numeral classifier constructions. *Revista da ABRALIN, 12*(1), 253–267

Liddell, T. M., & Kruschke, J. K. (2018). Analyzing ordinal data with metric models: What could possibly go wrong? *Journal of Experimental Social Psychology*, *79*, 328–348.

Li, Y.-H. A. (1999). Plurality in a classifier language. *Journal of East Asian Linguistics*, *8*(1), 75–99.

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York, NY: Oxford University Press.

Luka, B. J., & Barsalou, L. W. (2005). Structural facilitation: Mere exposure effects for grammatical acceptability as evidence for syntactic priming in comprehension. *Journal of Memory and Language*, *52*(3), 436–459.

Lyons, C. (1999). *Definiteness*. Cambridge: Cambridge University Press.

Maekawa, K., Yamazaki, M., Ogiso, T., Maruyama, T., Ogura, H., Kashino, W., & Den, Y. (2014). Balanced corpus of contemporary written Japanese. *Language Resources and Evaluation*, *48*(2), 345–371.

Mai, Z., & Yuan, B. (2016). Uneven reassembly of tense, telicity and discourse features in L2 acquisition of the Chinese *shì … de* cleft construction by adult English speakers. *Second Language Research*, *32*(2), 247–276.

Maratsos, M. P. (1976). *The Use of definite and indefinite reference in young children*. Cambridge: Cambridge University Press.

Marsden, H. (2005). *Quantifier scope in non-native Japanese: A comparative interlanguage study of Chinese-, English- and Korean-speaking learners* (Unpublished doctoral dissertation). Durham University.

Marsden, H. (2008). Pair-list readings in Korean-Japanese, Chinese-Japanese and English-Japanese interlanguage. *Second Language Research*, *24*, 189–226.

Marsden, H. (2009). Distributive quantifier scope in English-Japanese and Korean-Japanese Interlanguage. *Language Acquisition*, *16*(3), 135–177.

Marsden, H., Whong, M., & Gil, K.-H. (2018). What’s in the textbook and what’s in the mind: polarity item “any” in learner English. *Studies in Second Language Acquisition*, *40*(1), 91–118.

Master, P. A. (1987). *A cross-linguistic interlanguage analysis of the acquisition of the*

English article system (Unpublished doctoral dissertation). UCLA.

Matsui, T. (2000). *Bridging and relevance*. Amsterdam: John Benjamins.

Milin, P., Divjak, D., & Baayen, R. H. (2017). A learning perspective on individual differences in skilled reading: Exploring and exploiting orthographic and semantic discrimination cues. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(11), 1730–1751.

Miyamoto, E. T. (2002). Case markers as clause boundary inducers in Japanese. *Journal of Psycholinguistic Research, 31*(4), 307–347.

Mohri, F., & Isse, Y. (2018). Dainigengo syutoku ni okeru bogo no yakuwari ni tuite [The role of native language in second language acquisition]. *The Bulletin of Central Research Institute Fukuoka University*, *6*, 1–6.

Montrul, S., & Slabakova, R. (2003). Competence similarities between native and near-native speakers: An investigation of the preterite-imperfect contrast in Spanish. *Studies in Second Language Acquisition*, *25*(3), 351–398.

Mosel, U., & Hovdhaugen, E. (1992). *Samoan reference grammar*. Oslo: Scandinavian University Press.

Nakanishi, K. (2007). Formal properties of measurement constructions. Berlin: De Gruyter.

Nakanishi, K. (2008). The syntax and semantics of floating numeral quantifiers. In S. Miyagawa & M. Saito (Eds.), *Handbook of Japanese linguistics* (pp. 287–319). Oxford: Oxford University Press.

Nakanishi, K., & Tomioka, S. (2004). Japanese plurals are exceptional. *Journal of East Asian Linguistics*, *13*(2), 113–140.

Nemoto, N. (2005). On mass denotations of bare nouns in Japanese and Korean. *Linguistics, 43*(2), 383–413.

Nemoto, N. (2015). On the occurrences of anaphoric bare NPs in Japanese. *Linguistics, 53*(5), 1203–1231.

Nihongo Journal (2002). Room sharing. *Nihongo Journal*, 2000.6, 46–47.

Ochi, M. (2012). Universal numeric quantifiers in Japanese. *IBERIA: An International Journal of Theoretical Linguistics*, *4*(2), 40–77.

Okuma, T. (2019). Acquisition of floating quantifiers by L1 English speakers of L2 Japanese. In M. M. Brown & B. Dailey (Eds.), *Proceedings of the 43rd Annual Boston University Conference on Language Development* (pp. 496–508). Somerville, MA: Cascadilla Press.

Orfitelli, R., & Polinsky. M. (2017). When performance masquerades as comprehension: Grammaticality judgments in non-native speakers. In M. Kopotev, A. Mustajoki, & O. Lyashevskaya (Eds.), *Quantitative approaches to the Russian language* (pp. 197–211). New York, NY: Routledge.

Parrish, B. (1987). A new look at methodologies in the study of article acquisition for learners of ESL. *Language learning, 37*(3), 361–384.

Pearlmutter, N. J., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language*, *41*(3), 427–456.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203.

Prévost, P., & White, L. (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research, 16*(2), 103–133.

Ramchand, G., & Svenonius, P. (2008). Mapping a parochial lexicon onto a universal semantics. In M. Biberauer (Ed.), *Limits of syntactic variation* (pp. 219–245). Amsterdam: John Benjamins.

R Core Team. (2018). *R: A language and environment for statistical computing.* R Foundation for statistical computing, Vienna, Austria. https://www.R-project.org/.

Roberts, C. (2002). Demonstratives as definites. In K. van Deemter & R. Kiibble (Eds.), *Information sharing* (pp. 89–196). Stanford, CA: CSLI.

Roberts, L., & Felser, C. (2011). Plausibility and recovery from garden paths in second language sentence processing. *Applied Psycholinguistics*, *32*(2), 299–331.

Roberts, L., Gullberg, M., & Indefrey, P. (2008). Online pronoun resolution in L2 discourse: L1 influence and general learner effects. *Studies in Second Language Acquisition*, *30*(3), 333–357.

Roberts, L., & Liszka, S. A. (2013). Processing tense/aspect-agreement violations on-line in the second language: A self-paced reading study with French and German L2 learners of English. *Second Language Research, 29*(4), 413–439.

Robertson, D. (2000). Variability in the use of the English article system by Chinese learners of English. *Second Language Research*, *16*(2), 135–172.

Rule, S., & Marsden, E. (2006). The acquisition of functional categories in early French second language grammars: The use of finite and non-finite verbs in negative contexts. *Second Language Research*, *22*(2), 188–218.

Sahkai, H. (2015). Demonstratives and Information Structure in Spoken Estonian. In M. M. Jocelyne Fernandez-Vest, J. Van Valin, & D. Robert. (Eds.), *Information structuring of spoken language from a cross-linguistic perspective* (pp. 211–235). Berlin: De Gruyter.

Sawasaki, K. (2007). *L2 reading by learners of Japanese: A comparison of different L1s* (Unpublished doctoral dissertation). Ohio State University.

Schaeffer, J., & Matthewson, L. (2005). Grammar and pragmatics in the acquisition of article systems. *Natural Language & Linguistic Theory, 23*(1), 53–101.

Schütze, C. T. (1996). *The empirical base of linguistics: Grammaticality judgments and linguistics methodology*. Chicago, IL: The University of Chicago.

Schütze, C. T., & Sprouse, J. (2013). Judgement data. In R. J. Podesva & S. Devyani (Eds.), *Research methods in linguistics* (pp. 27–50). Cambridge: Cambridge University Press.

Schwartz, B. D., & Sprouse, R. A. (1996). L2 cognitive states and the full transfer/full access model. *Second Language Research*, *12*(1), 40–72.

Schwartz, F. (2009). *Two types of definites in natural language* (Unpublished doctoral dissertation). University of Massachusetts at Amherst.

Schwartz, F. (2013). Two kinds of definites cross-linguistically. *Language and Linguistics Compass*, *7*(10), 534–559.

Schwartz, F. (2019). Weak vs. strong definite articles: Meaning and form across languages. In A.-G. Ana, J. P. Loyo, & V. V.-R. Maldonado (Eds.), *Definiteness across languages* (pp.1–37). Berlin: Language Science Press.

Schwarzschild, R. (2002). The grammar of measurement. In *Semantics and linguistic theory* (Vol. 12, pp. 225–245).

Shibatani, M. (1990). *The languages of Japan*. Cambridge: Cambridge University Press.

Shimanskaya, E., & Slabakova, R. (2015). Re-assembling objects: A new look at the L2 acquisition of pronominal clitics. *Bilingualism: Language and Cognition*, 1–18.

Shin, K. Y. (2009). Numeral Quantifiers: NP modifiers and relational quantity nominals. *Language Research*, *45*(1), 131–156.

Shin, K. Y. (2017). Partitive descriptions in Korean. *Glossa: a journal of general linguistics*, *2*(1), 1–21.

Slabakova, R. (2003). Semantic evidence for functional categories in interlanguage grammars. *Second Language Research*, *19*(1), 42–75.

Slabakova, R. (2006). Learnability in the second language acquisition of semantics: A bidirectional study of a semantic parameter. *Second Language Research*, *22*(4), 498–523.

Slabakova, R. (2008). *Meaning in the second language*. Berlin: Mouton de Gruyter.

Slabakova, R. (2009). Features or parameters: Which one makes second language acquisition easier, and more interesting to study? *Second Language Research, 25*(2), 313–324.

Slabakova, R. (2016). *Second language acquisition*, Oxford: Oxford University Press.

Slabakova, R., & Ivanov, I. (2011). A more careful look at the syntax-discourse interface. *Lingua*, *121*(4), 637–651.

Smith, M. (2016). *Testing the shallow structure hypothesis in L2 Japanese* (Unpublished doctoral dissertation). Michigan State University.

Snape, N. (2008). *The acquisition of the English determiner phrase by L2 learners: Japanese and Spanish*, Saarbrücken: VDM Verlag Dr. Müller.

Snyder, W. (2000). An experimental investigation of syntactic satiation effects. *Linguistic Inquiry, 31*(3), 575–582.

Sorace, A. (2006). Gradedness and optionality in mature and developing grammars. In G. Fanselow, C. Féry, R. Vogel, & M. Schlesewsky (Eds.), *Gradience in Grammar* (pp. 106–123). Oxford: Oxford University Press.

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. *Linguistic Approaches to Bilingualism*, *1*(1), 1–33.

Sorace, A., & Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research*, *22*(3), 339–368.

Sorace, A., & Serratrice, L. (2009). Internal and external interfaces in bilingual language development: Beyond structural overlap. *International Journal of Bilingualism*, *13*(2), 195–210.

Spinner, P. (2013). The second language acquisition of number and gender in Swahili: A feature reassembly approach. *Second Language Research, 29*(4), 455–479.

Spinner, P., & Gass, S. M. (2019). *Using judgments in second language acquisition research*. New York, NY: Routledge.

Sprouse, J. (2009). Revisiting satiation: Evidence for an equalization response strategy. *Linguistic Inquiry*, *40*(2), 329–341.

Sprouse, J. (2011). A validation of Amazon Mechanical Turk for the collection of acceptability judgments in linguistic theory. *Behavior Research Methods*, *43*(1), 155–167.

Sprouse, J. (2015). Three open questions in experimental syntax. *Linguistics Vanguard, 1*(1), 89–100.

Sprouse, J., Schütze, C. T., & Almeida, D. (2013). A comparison of informal and formal acceptability judgments using a random sample from Linguistic Inquiry 2001–2010. *Lingua*, *134*, 219–248.

Sprouse, R. A. (2006). Full transfer and relexification. In C. Lefebvre, L. White, & C. Jourdan (Eds.), *L2 acquisition and Creole genesis: Dialogues* (pp. 169–181). Philadelphia, PA: John Benjamins.

Strawson, P. F. (1950). On referring. *Mind*, *59* (235), 320–344.

Stringer, D. (2010). The gloss trap. In Linguistic relativity in Z.-H. Han & T. Cadierno (Eds.), *SLA: Thinking for speaking* (pp. 102–124). Clevedon: Multilingual Matters.

Su, J. (2019). Reassembly of plural and human features in the L2 acquisition of Chinese by adult Korean speakers. *Second Language Research*, *35*(4), 529–555.

The Japan foundation (2017). *JF nihongo kyooiku sutandaado ni motozuita hyoka to nihongo nooryoku siken no goohi hantei to no kankei* [Relation between JFS-based evaluation and JLPT pass or fail determination]. Retrieved from https://jfstandard.jp/pdf /jfs_jlpt_report2017.pdf

Thomas, M. (1989). The acquisition of English articles by first- and second-language learners. *Applied Psycholinguistics*, *10*(3), 335–355.

Trahey, M., & White, L. (1993). Positive evidence and preemption in the second language classroom. *Studies in Second Language Acquisition*, *15*(2), 181–204.

Trenkic, D. (2007). Variability in second language article production: Beyond the representational deficit vs. processing constraints debate. *Second Language Research*, *23*(3), 289–327.

Trenkic, D. (2008). The representation of English articles in second language grammars: Determiners or adjectives? *Bilingualism: Language and Cognition*, *11*(1), 1–18.

Trenkic, D., Mirkovic, J., & Altmann, G. T. (2014). Real-time grammar processing by native and non-native speakers: Constructions unique to the second language. *Bilingualism: Language and Cognition*, *17*(2), 237–257.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, *33*(3), 285–318.

Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(3), 528–553.

Tryzna, M. (2009). Questioning the validity of the Article Choice Parameter and the Fluctuation Hypothesis: Evidence from L2 English article use by L1 Polish and L1 Mandarin Chinese speakers. In M. del P. G. Mayo & R. Hawkins (Eds.), *Second language acquisition of articles: Empirical findings and theoretical implications* (pp. 67–86). Amsterdam: John Benjamins.

Tsimpli, I., & Dimitrakopoulou, M. (2007). The interpretability hypothesis: Evidence from wh-interrogatives in second language acquisition. *Second Language Research*, *23*(2), 215–242.

Tsimpli, I., & Sorace, A. (2006). Differentiating interfaces: L2 performance in syntax-semantics and syntax-discourse phenomena. In D. Bamman, T. Magnitskaia, & C. Zaller (Eds.), *Proceedings of the 30th Annual Boston University Conference on Language Development* (pp. 653–664). Somerville, MA: Cascadilla Press.

Tsutsumi, R. (2012). *Gendai Nihongo Shijishi no Soogooteki Kenkyu* [An overall study of the demonstratives of modern Japanese]. Tokyo: Coco Publishing Company.

Tucker, M. A., Idrissi, A., & Almeida, D. (2015). Representing number in the real-time processing of agreement: Self-paced reading evidence from Arabic. *Frontiers in Psychology*, *6*, Article 347. doi:10.3389/fpsyg.2015.00347

Tuniyan, E. (2018). *Second language acquisition of definiteness: A feature-based contrastive approach to second language learnability* (Unpublished doctoral dissertation). University of Southampton.

Unsworth, S. (2005). *Child L2, adult L2, child L1: Differences and similarities. A study on the acquisition of direct object scrambling in Dutch*. Utrecht, LOT.

Verkuyl, H. J. (1981). Numerals and quantifiers in X-bar-syntax and their semantic interpretation. In T. M. V. Janssen, M. J. B. Stokhof, & J. A. G. Groenendijk (Eds.), *Formal methods in the study of language* (pp. 567–599). Amsterdam: Mathematical Centre Tracts.

Vermeulen, R. (2012). Word order variation and information structure in Japanese and Korean. In A. Neeleman & R. Vermeulen (Eds.), *The Syntax of Topic, Focus, and Contrast: An Interface-based Approach* (pp. 77–118). Berlin: De Gruyter Mouton.

von Heusinger, K. (2011). Specificity. In K. von Heusinger, C. Maienborn, & P. Portner (Eds.), *Semantics: An international handbook of natural language Meaning* (Vol. 2, pp. 1025-1058). Berlin: De Gruyter.

Wakabayashi, S. (1998). Systematicity in the use of the definite article by Japanese learners of English. *Bulletin of Gunma prefectural women's university*, *19*, 91–107.

Watanabe, A. (2006). Functional projections of nominals in Japanese: Syntax of classifiers. *Natural Language & Linguistic Theory*, *24*(1), 241–306.

White, L. (2003). *Second language acquisition and Universal Grammar*. Cambridge: Cambridge University Press.

White, L. (2009). Some questions about feature re-assembly. *Second Language Research*, *25*(2), 343–348.

White, L. (2011). Second language acquisition at the interfaces. *Lingua*, *121*(4), 577–590.

White, L. (2018), What is easy and what is hard: Lessons from linguistic theory and SLA research. In J. Cho, M. Iverson, T. Judy, T. Leal, & E. Shimanskaya (Eds.) *Meaning and structure in second language acquisition* (pp. 263–282). Amsterdam: John Benjamins.

Winter, B. (2020). *Statistics for linguists: an introduction using R*. New York, NY: Routledge.

Wolter, L. (2006). *That's that: The semantics and pragmatics of demonstrative noun phrases* (Unpublished doctoral dissertation). University of California at Santa Cruz.

Yang, M., & Ionin, T. (2009). L2 English articles and the computation of uniqueness. In J. Crawford, K. Otaki, & M. Takahashi (Eds.), *Proceedings of the 3rd Conference on Generative Approaches to Language Acquisition North America (GALANA)* (pp. 325–335). Somerville, MA: Cascadilla Press.

Yoshida, E. (2011). *Referring Expressions in English and Japanese: patterns of use in dialogue processing*. Amsterdam: John Benjamins.

Yuan, B. (2010). Domain-wide or variable-dependent vulnerability of the semantics—syntax interface in L2 acquisition? Evidence from wh-words used as existential polarity words in L2 Chinese grammars. *Second Language Research*, *26*(2), 219–260.

Yuan, B. (2014). 'Wh-on-earth' in Chinese speakers' L2 English: Evidence of dormant features. *Second Language Research*, *30*(4), 515–554.

Zhang, L. (2020). The influence of first language at the semantics-pragmatics interface: Evidence from definite and demonstrative determiners in L2 English. *Linguistic Approaches to Bilingualism*. Advance online publication. https://doi.org/10.1075/lab.18096.zha

Zufferey, S., Mak, W., Degand, L., & Sanders, T. (2015). Advanced learners' comprehension of discourse connectives: The role of L1 transfer across on-line and off-line tasks. *Second Language Research, 31*(3), 389–411.