

# Definite and Indefinite Articles in Learner English: Identifying the Learning Problem and Addressing It with Processing Instruction Intervention

Kateryna Derkach

Fitzwilliam College

March 2022

This thesis is submitted for the degree of Doctor of Philosophy

**Declaration**

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee.

## **Definite and Indefinite Articles in Learner English: Identifying the Learning Problem and Addressing It with Processing Instruction Intervention**

### **Abstract**

Articles in L2 English (“a”/“the”) are notoriously difficult to acquire, especially for learners with article-less, [-art], L1s. The purpose of this dissertation is to identify the potential causes of learner difficulty with articles and to suggest an effective teaching intervention.

To identify the learner problem, I conducted a learner corpus-based study of English article use by L2 learners with four typologically distinct first languages (L1s): German and Brazilian Portuguese, both [+art], Chinese and Russian, [-art]. I coded and analysed 660 written scripts from the Education First Cambridge Open Database to investigate several semantic and morphosyntactic factors, such as specificity and prenominal modification. The key finding is the differential effect of specificity on definite and indefinite articles: learners tend to associate specificity with “a”, which results in article omission with non-specific indefinite singulars and overuse of “a” with specific indefinite mass nouns. Prenominal modifiers further contribute to perceived specificity leading to article overuse with modified indefinite mass nouns. However, in definite contexts, prenominal modifiers are associated with increased article omission.

Drawing on Input Processing theory, I argue that learner difficulties might stem from the inappropriate processing of articles in the input, which leads to learners making incorrect mappings between form (articles) and meaning (identifiability). In addition, learners at different levels are subject to cognitive constraints of varying strength, which can make it difficult or impossible to attend to morphosyntactic information. I designed a Processing Instruction Intervention containing activities that force correct article processing. Seventy pre-intermediate to upper-intermediate learners, half [-art] and half [+art], participated in a three-week-long online intervention experiment with randomised control-group pre-test post-test design. The pre-test contained a timed grammaticality judgment task (GJT) and an elicited imitation (EI) task. Mixed-effects regression modelling revealed a significant improvement in article accuracy on the timed GJT for the

Experimental group but not for the Control group for both [-art] and [+art] L1s, which was maintained by the delayed post-test (3 months later). The results of the EI were less robust but showed a similar trend of improvement for the Experimental group but not for the Control group, which was mediated by participants' education level and potentially linked to working memory effects in younger participants. Overall, Processing Instruction appears to be an effective and practical way of addressing common difficulties with articles in learner English.

### **Acknowledgements**

I would like to thank everyone who has contributed to this project in one way or another. I have been blessed with an incredible PhD supervisor Dr Dora Alexopoulou, who has shown unfailing support in all aspects of PhD life and beyond, genuine enthusiasm about this research project, and provided constructive and insightful feedback at all stages. I am grateful to colleagues at the Department of Theoretical and Applied Linguistics, Professor Ianthi Tsimpli, Professor Henriëtte Hendriks, and Dr Xiaobin Chen, for their valuable feedback at different stages of my PhD. I also thank Dr Akira Murakami for his particularly relevant feedback and advice, and the statisticians at the Cambridge Centre for Mathematical Sciences for their informed and accessible guidance.

I am indebted to the participants of my experiments, as this research would have been impossible without their interest and feedback. I would like to thank everyone who has helped in participant recruitment for the pilot and for the main data collection, to our partners at EF Education First and to the numerous language teachers around the world who expressed their interest in this project and promoted it among their students.

I would also like to thank the numerous native speakers of English who volunteered to trial and give feedback on the materials for my intervention study, as well as to those who recorded the audio test prompts for my experiments.

I thank the PhD community at the Department of Theoretical and Applied Linguistics both for their academic support and feedback and for their comradery, especially Athina Ntalli, James Algie, Carlos Balhana, Samantha Sie, and Itamar Shatz.

I am grateful to the community of Fitzwilliam College for being so supportive and inclusive both academically, mentally, and financially in the often-difficult times during which I have worked on my PhD. I am also grateful to the Cambridge Trust for funding my research and providing additional support when it was necessary.

As a large proportion of my time as a PhD student coincided with the Covid pandemic and the final stages before submission were severely affected by the war waged by Russia on my home

country Ukraine, I am grateful that my family, who have always been very supportive of all my academic endeavours, are alive and holding on despite the dangers they are facing daily back home.

Finally, I am eternally grateful to my husband Alex and my two-year-old daughter Sofia for their unconditional love, support and for the joy they bring to every aspect of my life.

## Table of Contents

|  |    |
|--|----|
| Abstract   | 3  |
| Acknowledgements   | 5  |
| Table of Contents  | 7  |
| List of Tables   | 10 |
| List of Figures  | 11 |
| Chapter 1: Introduction  | 13 |
| Chapter 2: Background  | 17 |
| Article Distribution in English  | 17 |
| Articles in Other Languages  | 17 |
| Key Semantic and Pragmatic Distinctions                                | 18 |
| Definiteness   | 18 |
| Specificity  | 19 |
| Discourse  | 21 |
| Chapter 3: Corpus-Based Study of Article Use in Learner English        | 23 |
| Factors Affecting Article Accuracy: Previous Research                  | 23 |
| The Role of L1   | 23 |
| The Role of Semantic Features: Specificity                             | 23 |
| The Role of Discourse  | 26 |
| The Role of Nominal Features   | 26 |
| Research Questions   | 27 |
| Methodology  | 28 |
| Learner Data   | 28 |
| Coding   | 29 |
| Mixed-Effects Logistic Regression                                      | 37 |
| Power Analysis and Sample  | 38 |
| Analysis and Results   | 43 |
| Data Distribution and Observed Accuracy Rates                          | 43 |
| Predictors of Accuracy in Definite and Indefinite Contexts             | 47 |
| The Effect of Discourse-Pragmatic Context on Definite Article Accuracy | 53 |
| Discussion   | 66 |
| Summary of Findings  | 66 |
| Specificity and Prenominal Modifier                                    | 67 |
| Syntactic Position   | 70 |
| Discourse-Pragmatic Context  | 71 |

|  |     |
|--|-----|
| Abstractness   | 72  |
| Limitations  | 72  |
| Chapter 4: Teaching and Learning Articles in L2 English                  | 74  |
| Teaching Articles to English Learners: Previous Research                 | 74  |
| Articles in L2 English: The Learning Problem                             | 80  |
| Interpretability Hypothesis and L2 English Articles                      | 82  |
| Cognitive Constraints and Processing                                     | 84  |
| Processing of L2 English Articles  | 86  |
| Sources of Different Types of Errors                                     | 89  |
| Processing Instruction Intervention for L2 English Articles              | 91  |
| Rationale and Research Questions   | 93  |
| Chapter 5: Processing Instruction Intervention Pilot Study               | 95  |
| Methodology  | 95  |
| Intervention Materials   | 95  |
| Assessment   | 101 |
| Procedure  | 104 |
| Coding   | 105 |
| Analysis and Results   | 106 |
| Comparing Groups at Pre-Test   | 106 |
| Descriptive Statistics   | 107 |
| The Effect of the Intervention on the GJT                                | 107 |
| The Effect of the Intervention on EI Results                             | 111 |
| The Effect of the Intervention on the Free Writing Task                  | 112 |
| Discussion and Issues Identified in the Pilot                            | 113 |
| Sample Size  | 113 |
| Participants' Proficiency Level  | 113 |
| Experimental Activities  | 114 |
| Chapter 6: Processing Instruction Intervention Main Study                | 116 |
| Methodology  | 116 |
| Power Analysis   | 116 |
| Changes to the Experimental and Testing Materials Based on Pilot Results | 120 |
| Procedure  | 127 |
| Coding and Scoring   | 128 |
| Analysis and Results   | 130 |
| Mixed-Effects Logistic Regression  | 130 |
| Comparing Groups at Pre-Test   | 130 |



|   |     |
|---|-----|
| GJT Results   | 132 |
| EI Results  | 144 |
| Written Production  | 151 |
| Discussion  | 152 |
| Key Findings  | 152 |
| The Role of Form-Meaning Mappings and of Cognitive Constraints                                      | 154 |
| Testing Practice Effects  | 155 |
| No Effect of Specificity on GJT or EI   | 157 |
| The Mitigating Effect of Education Level on EI  | 158 |
| Limitations   | 159 |
| Chapter 7: General Discussion   | 162 |
| Summary of the Findings   | 162 |
| Implications for Intervention Studies in SLA  | 163 |
| Acquisition of Articles in L2 English   | 163 |
| Implications for Learning and Teaching Practice   | 164 |
| Limitations and Further Research  | 165 |
| References  | 167 |
| Appendices  | 178 |
| Appendix A: Full List of Formulaic Sequences Identified in Corpus Data                              | 178 |
| Appendix B: Technical Details for Statistical Models  | 183 |
| Appendix C: Example of Explicit Information from the Pilot  | 184 |
| Appendix D: Structure of Pilot Experimental Group Activities  | 185 |
| Appendix E: Example Control Group Activity  | 187 |
| Appendix F: Pilot Tests Structure and Example Items   | 188 |
| Appendix G: Consent Form and Background Questionnaire for the Pilot                                 | 192 |
| Appendix H: Links to Example Video Lessons for Main Intervention Study                              | 204 |
| Appendix I: Example Activities from the Main Intervention   | 205 |
| Appendix J: Main Intervention Tests – Structure and Example Items                                   | 210 |
| Appendix K: Main Intervention Information Sheet, Consent Form, Background Questionnaire, and C-test | 227 |
| Appendix L: Main Intervention GJT Indefinite Mass and Plural Model Results                          | 240 |
| Appendix M: Main Intervention GJT Indefinite Singular Model Results                                 | 241 |

### List of Tables

|   |     |
|---|-----|
| Table 1 Most Frequent Formulaic Sequences in the Subcorpus .....  | 32  |
| Table 2 Corpus Data Variable Coding .....   | 33  |
| Table 3 Example Coded Learner Script .....  | 36  |
| Table 4 Simulated Corpus Dataset Model Results .....  | 40  |
| Table 5 Distribution of Nominals in the Subcorpus Across EF Levels and NLs .....                                  | 42  |
| Table 6 Error Type Distribution .....   | 46  |
| Table 7 Full Accuracy Model Results and Performance Metrics .....   | 51  |
| Table 8 Definite Accuracy Model Results and Performance Metrics .....   | 55  |
| Table 9 Results of the Error Type Model for Indefinite Count Singular Nouns .....                                 | 59  |
| Table 10 Results of the Error Type Model for Definite Count Singular Nouns (L1-Chinese and L1-Russian only) ..... | 63  |
| Table 11 Results of the Error Type Model for Indefinite Mass Nouns .....  | 65  |
| Table 12 Participants of the Pilot Experiment .....   | 104 |
| Table 13 Pilot GJT Model Results .....  | 109 |
| Table 14 Numbers of Obligatory Contexts Produced in Pilot Free Writing Task .....                                 | 113 |
| Table 15 Simulated Intervention Dataset Model Results .....   | 118 |
| Table 16 Participant Distribution by L1, Proficiency Level, Education Level, Gender .....                         | 126 |
| Table 17 Results of the Model Comparing Control and Experimental Groups at Pre-Test .....                         | 131 |
| Table 18 GJT Model Results .....  | 137 |
| Table 19 Delayed Subgroup GJT Model Results .....   | 143 |
| Table 20 EI Model Results .....   | 147 |
| Table 21 Delayed Subgroup EI Model Results .....  | 148 |
| Table 22 Formulaic Sequences Identified in Corpus Data .....  | 178 |
| Table 23 Structure of Pilot GJT Test Items .....  | 188 |
| Table 24 Structure of Pilot EI Test Items .....   | 190 |
| Table 25 Pilot Free Writing Prompts .....   | 191 |
| Table 26 Structure of Main Intervention GJT Items .....   | 211 |
| Table 27 Main Intervention GJT Test A (Target Items Only) .....   | 211 |
| Table 28 Main Intervention GJT Test B (Target Items Only) .....   | 214 |
| Table 29 Main Intervention GJT Test C (Target Items Only) .....   | 216 |
| Table 30 Main Intervention EI Test A Structure and Items .....  | 220 |
| Table 31 Main Intervention EI Test B Structure and Items .....  | 222 |
| Table 32 Main Intervention EI Test C Structure and Items .....  | 223 |
| Table 33 Main Intervention Writing Prompts .....  | 226 |
| Table 34 Main Intervention GJT Indefinite Mass and Plural Model Results .....                                     | 240 |
| Table 35 Main Intervention GJT Indefinite Singular Model Results .....  | 241 |

## List of Figures

|  |     |
|--|-----|
| Figure 1 Distribution of Observations across Key Variables: Target Article, Specificity, Modifier, Abstractness, Syntactic Position, Discourse-Pragmatic Context.....                                | 44  |
| Figure 2 Development across EF Levels as Measured by Accuracy Rate (Left) and TLU (Right) .....  | 45  |
| Figure 3 Error Type Distribution across NL, Target Article, and Noun Type .....  | 46  |
| Figure 4 The Effect of Level across NLs (with 95% Confidence Intervals) .....  | 47  |
| Figure 5 The Effect of NL across Definiteness and Noun Type .....  | 48  |
| Figure 6 The Effect of Specificity (Left) and Modifier Presence (Right) across Definiteness and Noun Type .....  | 50  |
| Figure 7 The Effect of Syntactic Position .....  | 51  |
| Figure 8 The Effect of NL in Definite Contexts .....   | 54  |
| Figure 9 Interaction between Discourse-Pragmatic Context and Noun Type (Left), and Modifier and Noun Type (Right) .....  | 55  |
| Figure 10 The Effect of NL Alone (Top) and in Interaction with Level (Bottom) on Predicted Probabilities of Error Types in Count Singular Indefinites .....  | 57  |
| Figure 11 The Effect of Specificity in Interaction with Modifier on Predicted Probabilities of Error Types in Count Singular Indefinites .....   | 58  |
| Figure 12 The Effect of Syntactic Position on Predicted Probabilities of Error Types in Count Singular Indefinites .....   | 59  |
| Figure 13 The Effect of NL on Predicted Probabilities of Error Types in Singular Definites .....   | 61  |
| Figure 14 The Effects of Discourse-Pragmatic Context (Left) and Modifier Presence (Right) on Predicted Probabilities of Error Types in Singular Definites for L1-Russian and L1-Chinese Learners. 62 |     |
| Figure 15 The Effects of Specificity (Left) and Modifier Presence (Right) on Predicted Probabilities of Error Types in Mass Indefinites.....   | 65  |
| Figure 16 Target English Distribution of Articles on Definiteness-Specificity Plane .....  | 81  |
| Figure 17 Hypothesised Distribution of Articles in Learner Interlanguage.....  | 82  |
| Figure 18 Example Item from the Structured Input Activity (Indefinite Contexts) .....  | 96  |
| Figure 19 Example Item from the Structured Input Activity with an Ambiguous Noun .....   | 98  |
| Figure 20 Order of Item Presentation for Pilot Experimental Activities in the Indefinite Block (Sets 1-10).....  | 99  |
| Figure 21 Example Item from the Structured Input Activity (Definite vs. Indefinite Contexts).....  | 99  |
| Figure 22 Order of Item Presentation for Pilot Experimental Activities in the Definite/Indefinite Block (Sets 11-13) .....   | 100 |
| Figure 23 Pilot GJT Results .....  | 108 |
| Figure 24 Pilot EI Results.....  | 108 |
| Figure 25 Pilot Free Writing Results .....   | 109 |
| Figure 26 Pilot GJT Score Change as a Function of Pre-Test Score and Group .....   | 110 |
| Figure 27 Rate of Intended Target Production on the Pilot EI Task .....  | 112 |
| Figure 28 The Effect of the Intervention Predicted by the Simulated Dataset Model (with 95% Confidence Intervals).....   | 118 |
| Figure 29 Order of Item Presentation for Experimental Activities .....   | 121 |
| Figure 30 Participant Distribution by C-test Scores, Age, Age of Onset of English, Years in English-Speaking Countries, Hours Learning English.....  | 126 |
| Figure 31 Correlation between Self-Reported Proficiency Levels and C-test Scores .....   | 132 |
| Figure 32 The Overall Effect of the Intervention on GJT Target Items by L1-type .....  | 133 |
| Figure 33 Predicted Effect of Intervention on GJT Accuracy Rate by Definiteness, Noun Type, and Item Variant (with 95% Confidence Intervals) .....   | 134 |

|  |     |
|--|-----|
| Figure 34 Predicted Accuracy in Identifying Errors on IC Indefinite Mass and Plural GJT Items with or without Prenominal Modifiers.....                            | 136 |
| Figure 35 The Effect of Proficiency Level on GJT Accuracy Rate (with 95% Confidence Intervals)....   | 140 |
| Figure 36 The Effect of Noun Type in Interaction with Definiteness and L1 type on Learner Accuracy on Ungrammatical GJT Items (with 95% Confidence Intervals)..... | 141 |
| Figure 37 Predicted Accuracy Rates at Pre-, Immediate Post-, and Delayed Post-Tests for the Delayed Subgroup (n = 21), with 95% Confidence Intervals .....         | 143 |
| Figure 38 Production of Target Nominals on the EI Task .....   | 145 |
| Figure 39 The Overall Effect of the Intervention on Article Accuracy on the EI Task by L1-type .....   | 146 |
| Figure 40 Predicted Effect of Intervention on EI Accuracy Rate by Education Level (with 95% Confidence Intervals).....   | 147 |
| Figure 41 Error Types in Indefinite Contexts on the EI Task (Observed Data) .....  | 150 |
| Figure 42 The Effect of Proficiency Level on EI Accuracy Rate (with 95% Confidence Intervals) .....  | 151 |
| Figure 43 Correlation Between Age and Education Level of Participants .....  | 158 |
| Figure 44 Factorial Structure of Pilot Experimental Activities: Block 1 .....  | 185 |
| Figure 45 Item Structure Within 12-Item Sets .....   | 186 |
| Figure 46 Factorial Structure of Pilot Experimental Activities: Block 2 .....  | 186 |

## Chapter 1: Introduction

The acquisition of articles presents persistent challenges for learners of English as a second language (L2), even for those at near-native level. Despite the abundance of articles in the input and the detailed explanations and practice activities contained in most language courses and learning materials, article errors can reveal speakers' non-native identities even at higher proficiency levels (DeKeyser, 2005; Ionin et al., 2008; Master, 1987; Murakami & Alexopoulou, 2016a; Snape, 2008; Trenkic, 2007).

Article accuracy is influenced by learner variables, such as their first language (L1) and proficiency in English. Inherent features of articles also play a role, e.g. nominal (morphosyntactic) features (e.g. number, countability), semantic features (e.g. definiteness, specificity), and article discourse functions. A large body of research has demonstrated the relevance of such features (Ionin et al., 2004; Liu & Gleason, 2002; Robertson, 2000; Snape, 2008; Trenkic, 2007), although no study has considered all these features together, so their relative importance and interactions are not yet known.

Articles in L2 English have also been the focus of multiple intervention studies testing the effectiveness of various teaching methods in improving learners' article accuracy (Kao, 2020; Muranoi, 2000; Snape et al., 2015; Umeda et al., 2017). Rather strikingly, there is little connection between the more linguistically driven and the more pedagogically driven studies in that the former usually do not address teaching implications in any detail, whereas the latter tend to gloss over the linguistic analysis and merely use articles as a test case for comparing different instructional approaches.

There are good reasons to think that informing pedagogical interventions with insights from Second Language Acquisition (SLA) developmental research on English articles can lead to more effective teaching. If, as suggested in previous SLA research, there are certain linguistic variables influencing learner article use and making it non-targetlike, then ignoring this influence in pedagogical interventions seems ill-advised. Yet, the intervention studies mentioned in the previous

paragraph either focus on a very limited aspect of article use (e.g. only first-mention and anaphoric uses of articles) or aggregate results across various article uses considering only overall accuracy rates.

As for teaching materials, much more attention is typically given to idiomatic article uses with various kinds of proper nouns (e.g. “Lake Ontario” but “the River Nile”, “Heathrow Airport” but “the Ritz Hotel”) than to practising article use in a wide variety of contexts. As a result, learners might feel relatively confident using articles in certain well-practised environments, for example with first- and second-mention referents in narratives (e.g. “I saw *a dog* in the garden. *The dog* was playing with a ball.”). However, they may find themselves completely unprepared to deal with real-world cases, as in the following examples taken from EFCAMDAT, a learner writing corpus<sup>1</sup>.

(1) I also like to read articles in different newspapers but often its not easy for me to find

~~the~~ *[a]* newspaper which matches my taste. (L1-German, B2 level)<sup>2</sup>

(2) When he came back his family home and attempted to live *[a]* normal human life he

immediately met lots of difficulties [...] (L1-Russian, B2 level)

The use of articles in the above examples is not easily explained by textbook rules. In (1) the indefinite article “a” is appropriate despite the fact that the noun is post-modified with a relative clause, which is frequently taught as a cue to using the definite article “the”. The noun “life” in (2) is sometimes used as a countable and sometimes as a mass noun in English, which will define the appropriateness of “a” in indefinite contexts, but this is not something typically covered in textbooks for English as a Foreign Language (EFL).

Incorporating linguistic variables into a teaching intervention would entail finding a way to help learners focus on relevant linguistic cues and ignore the irrelevant ones. This is a non-trivial challenge, and a few contrasting examples are nowhere near enough to explain the full paradigm to

---

<sup>1</sup> Further details on this corpus are provided in Chapter 3.

<sup>2</sup> Learners’ L1 and proficiency levels according to the Common European Framework of Reference, CEFR, are indicated in parentheses; original grammar and spelling are preserved; article errors are corrected in square brackets.

learners. Nevertheless, more efficient practical teaching solutions need to be both linguistically and developmentally well-grounded.

The present study aims to fill this gap by addressing both aspects of the issue and attempting to link linguistic analysis of articles and SLA developmental research with pedagogical practice. This is achieved by first conducting an empirical study involving corpus-based error analysis as a window into the representation of articles in learners' interlanguage (Corder, 1981). The study employs a corpus of learner writing in an EFL context aiming to document learner errors in real-life teaching environment as a better basis for developing a teaching intervention.

Based on the results of the learner corpus error analysis, I hypothesise that learners fail to process crucial structural cues, such as number and countability, and misanalyse the indefinite article as a marker of specificity. As shall be seen, this hypothesis aims to explain a number of observed patterns of learner article use, which can be traced back to the influence of multiple interacting factors, including countability, number, specificity of the referent, and the presence of a prenominal modifier.

This hypothesis forms the basis for designing an intervention that takes into account (i) the linguistic variables affecting learner article use, (ii) the developmental research on article processing and acquisition, and (iii) the pedagogical limitations on article teaching in English language classrooms.

This dissertation is structured as follows. Chapter 2 provides an overview of the article system in English and other languages, explaining the key semantic and pragmatic distinctions relevant for article use. Chapter 3 reviews existing SLA research on article acquisition and presents a corpus-based study of article use in learner English, which investigates the effect of multiple intertwined linguistic and learner variables on article accuracy. Building on the results of Chapter 3, Chapter 4 develops a hypothesis regarding the learning problem and suggests a teaching approach to address this problem. Chapter 5 presents the teaching intervention developed based on the suggestions in Chapter 4 and reports the pilot study of the intervention materials. Chapter 6

describes the changes made to the intervention materials based on the results of the pilot and reports the main intervention study. Chapter 7 concludes the thesis with a general discussion of the results and considers their implications.



## Chapter 2: Background

### Article Distribution in English

Articles are among the highest-frequency words in English, with “the” always at the top of frequency lists (Leech et al., 2001). However, bare nominals (i.e. those that require no article, henceforth labelled as “target  $\emptyset$ ” or “ $\emptyset$  contexts”) can account for approximately 50% of all nominals (Master, 1997). This creates a problem for learners, who notice the frequently occurring articles in English but may struggle to understand why they are often “omitted”.

The distribution of the three options (“the”, “a”, and  $\emptyset$ ), although not random, is determined by a combination of factors, some of which are meaning-related and some morphosyntactic. While the definite/indefinite distinction (discussed in more detail below) is semantic, the distinction between “a” and  $\emptyset$  is based on countability and number, namely “a” is required for count singular nouns, while  $\emptyset$  is appropriate for mass and plural nouns. Nevertheless, count singular nouns can appear bare in English, e.g. “go to school/work etc.”, “travel on foot/by car/plane etc.”, “stay in bed/at home/in town” etc. Such cases are limited to certain verb-preposition-noun combinations, but many are very frequently used in everyday communication. Thus, extracting patterns from the distribution of articles in target English can be challenging for learners, as it depends on various semantic, pragmatic, and morphosyntactic factors.

### Articles in Other Languages

Languages vary considerably regarding articles and their use (Lyons, 1999). Many of the major world languages lack articles (e.g. most Slavic languages, Hindi, Japanese), some only have the definite (Arabic, Hebrew) or the indefinite (Turkish) article (Dryer, 2013).

The article systems of languages which have them ([+art], e.g. Germanic and Romance languages) may deviate from the English pattern. For example, unlike Germanic languages, Romance languages do not allow bare mass nouns in argument position (“I eat *meat*” – “Je mange *de la viande*” in French).

Though definiteness is not grammaticalised in article-less, [-art], languages, they often have some notion akin to definiteness (e.g. familiarity, specificity), which can be expressed via lexical means, e.g. with demonstratives “this/that”, or syntactic means, e.g. through word order. Additionally, definiteness may be linked to other discourse and nominal features in distinct ways, e.g. in Chinese only definite and human referents can be marked as plural (Lardiere, 2009).

Given the variation not only in the presence of articles but in the meanings they grammaticalise across languages, the acquisition of articles is expected to present challenges for learners with [-art] L1s, as well as those with [+art] L1s where article systems function differently.

### **Key Semantic and Pragmatic Distinctions**

#### ***Definiteness***

The literature on definiteness from the point of view of language philosophy and semantics is vast and outside the scope of this work (Frege, 1960; J. A. Hawkins, 1978; Lyons, 1999; Russell, 1905, *inter alia*). For the purposes of this thesis, it suffices to draw the crucial semantic distinction, namely that the definite article makes presuppositions of uniqueness and existence, while the indefinite article may only presuppose existence (Heim, 2019). In (3) the speaker asserts that the book exists and is uniquely identifiable by both the speaker and the hearer, i.e. they both know which book is meant, and there are no other books to be considered. When the condition of uniqueness is not met, the indefinite article is used. For example, although the existence of the book is presupposed in (4), the speaker does not presume it is uniquely identifiable by the hearer. In (5) even the existence of the book is not necessarily presupposed.

(3) I bought *the book* about definiteness.

(4) I bought *a book* about definiteness.

(5) I need *a book* about definiteness.

Whether uniqueness can be presupposed depends on the knowledge shared by the speaker and the hearer, available from the relevant discourse and context (both immediate and larger, cultural context, or common knowledge). In (3) there may be only one book on definiteness in

existence or known within the speaker and hearer's social, cultural, or professional context (presupposition based on context). Alternatively, the speaker and the hearer may have already discussed a certain (and only one) book on definiteness (presupposition based on discourse), although other books on definiteness might exist and even be known to both the speaker and the hearer. By contrast, in (4) and (5) the speaker assumes that neither common knowledge nor discourse can help the hearer to uniquely identify the book the speaker has in mind. In (5) the speaker may not know which book on definiteness they want to buy or whether such a book even exists.

For mass and plural nouns, uniqueness applies to the entire mass or set, respectively, or the set of the noun's maximal elements (Heim, 2019, p. 36). In example (6), the nominal refers to the entire set of applicants who failed their exams, and this set is supposed to be known (e.g. all the applications have been received). By contrast, in (7) any applicant satisfying the description of having failed their exams has to be rejected, but the entire set is not known (e.g. applications are still coming in). Similarly, in (8) the speaker implies that the entire mass of expired meat was thrown away, while (9) makes no assumptions about the entire mass, i.e. there may be some expired meat that was not thrown away.

(6) We have to reject *the applicants* who failed their exams.

(7) We have to reject *applicants* who failed their exams.

(8) They threw away *the expired meat*.

(9) They threw away *expired meat*.

### ***Specificity***

A property cutting across the definiteness/indefiniteness distinction is specificity, which is not directly encoded in English. As with definiteness, there is extensive literature on the topic; however, one common feature across the different accounts is the "referential intention", or the communicative intention of the speaker to refer to something they have in mind (von Heusinger, 2019).

In this thesis, a specific referent is defined as one that refers to a certain entity which exists in the world and which the speaker has in mind. For example, “a book” in (4) is specific, since the speaker is referring to a certain existing book. This definition is similar to Bickerton’s “specific reference” (1981). Note the key difference from the definition adopted in some of the influential literature on L2 articles (most notably Ionin et al., 2004, and replications), which is that I do not consider noteworthiness (i.e. whether the speaker deems the referent noteworthy for the discourse) an essential feature. For this reason, (4) is considered specific when introducing a new referent into the discourse (first-mention indefinite) even if the identity of the book is irrelevant for further discourse and, thus, not noteworthy for the speaker.

The possibility of a non-specific reading arises in semantically opaque contexts, i.e. those involving opacity-creating operators, such as verbs of propositional attitude (e.g. “want”, “believe”), negation, questions, conditionals, modals, future, and intensional verbs (e.g. “look for”). In (5), the speaker may have a certain book in mind (specific reading) or may be satisfied with any book about definiteness (non-specific reading). The two possible readings are said to have wide scope (specific) or narrow scope (non-specific). Such ambiguity can occur in non-opaque contexts as well, as in (10), taken from Lyons (2009, p. 172), where “a student” has no fixed reference.

(10) I haven’t started the class yet; I’m missing *a student* – there should be fifteen, and I only count fourteen.

When the nominal refers to a certain entity, it may be termed “referential”, and when it has no fixed reference, as in (10) – non-referential (Lyons, 2009). This referentiality distinction also applies to definites. In (11) the winner is already known, unique, and specific, while in (12) there is clearly going to be one winner, but nobody knows who it is going to be because the competition has not finished yet, so “the winner” has no fixed reference. In (4) the speaker is referring to a specific book they bought, but in (5) any object satisfying the description will do.

(11) *The winner* got \$1000.

(12) *The winner* will get \$1000. (stated before the start of the competition)

The choice of the term “specificity” rather than “referentiality” is based on Lyons’s suggestion to use “specificity” as an “informal cover term” (1999, p. 173) to include both the wide/narrow scope and the referential/non-referential distinctions.

To summarise, the definition of specificity used in this thesis is looser than that in some of the current literature on L2 articles. The only essential criteria for specificity in this work are that the nominal refers to an existing (in a general sense) entity and that the reference is fixed.

### ***Discourse***

The definite article is an anaphoric device that can be connected with discourse and contextual antecedents in the following ways (based on J. A. Hawkins, 1978):

- Anaphoric: referring to something already mentioned in the discourse, e.g. “I bought *a book*. *The book* was interesting”.
- Situational: referring to something not explicitly mentioned in the discourse but implied by the situation or common knowledge, e.g. “Close *the window*” (immediate situation), “*The President* arrived” (larger situation), “I listened to his new *song*. *The tune* is catchy” (associative anaphoric use in Hawkins).
- Explanatory: referring to something unknown but identified by
  - an establishing relative clause, e.g. “Have you read *the book* which Diane recommended?”<sup>3</sup>;
  - an associative phrase, e.g. “*the roof* of the house”;
  - a noun-phrase complement, e.g. “I was shocked by *the fact* that they refused”;
  - a nominal modifier, e.g. “*the name* Nelly”.
- Reference to kind (generics), e.g. “*The lion* is a carnivorous animal” (only count singular nouns).

---

<sup>3</sup> A relative clause by itself does not make a nominal definite; the uniqueness presupposition still needs to be fulfilled, so there must be only one book recommended by Diane, cf. (6) and (7).

- Idiomatic: unexplainable/conventional use, e.g. “in *the morning*”.

In summary, the definite article functions to uniquely identify referents, while the indefinite article does not. Meanwhile, specificity is not encoded in the English article system, as both definites and indefinites can be specific or non-specific. Moreover, the indefinite article takes the form of “a/an” but only for countable singular nominals, while  $\emptyset$  is required for mass and plural nominals, whereas the definite article does not depend on these nominal features but performs a range of functions connecting discourse and contextual antecedents.

### Chapter 3: Corpus-Based Study of Article Use in Learner English

In this chapter, I will review the existing research on factors affecting article use in learner English and report a corpus-based study, which investigates the relative importance of these factors and how they work together in predicting article accuracy.

#### Factors Affecting Article Accuracy: Previous Research

##### *The Role of L1*

Studies directly comparing L2 English article use by [+art] and [-art] L1 learners unequivocally suggest that the latter are significantly disadvantaged. This has been observed in naturalistic spoken data (Master, 1987; Thomas, 1989), in data elicited by gap-fill tasks (R. Hawkins et al., 2006; Ionin et al., 2008; Reid et al., 2006; Snape, 2008), and in a large-scale corpus-based study of article accuracy in exam scripts by learners with various [+art] and [-art] L1s (Murakami & Alexopoulou, 2016a).

##### *The Role of Semantic Features: Specificity*

One of the most widely researched factors in the acquisition of articles is specificity. Early studies in both L1 and L2 acquisition documented increased use of “the” in [+specific, -definite] contexts (following Bickerton’s definition), such as first-mention indefinites (Bickerton, 1981; Brown, 1973; Cziko, 1986; Huebner, 1985; Maratsos, 1976; Thomas, 1989; Warden, 1976).

Butler (2002), also using Bickerton’s terminology, found evidence of using “the” with [+specific, -definite] referents in 80 L1-Japanese learners across four proficiency levels. The accuracy rate on the gap-fill task ranged from 55% at the lowest level to 75% at the highest. Further metalinguistic interviews confirmed that learners were confused about specificity. For example, in (13) the learner reported choosing “the” because they thought the woman was specific, disregarding the fact that this referent was being introduced for the first time. This problem persisted across proficiency levels. Additionally, lower-level learners often used “the” in the presence of a prenominal modifier, which they considered to indicate specificity (14).

- (13) Yesterday when I was sitting up in a bed looking at the rain [...], the woman appeared with long white box addressed to me, [...].

(14) School has just begun and I have already made *the terrible mistake*.

(from Butler, 2002)

Hua & Lee (2005) found that in a grammaticality judgment task L1-Chinese English learners were more likely to accept the ungrammatical bare count singulars in subject position when these were non-specific (as opposed to specific ones, following Bickerton's definition), regardless of whether the nouns were concrete (e.g. "computer") or abstract (e.g. "sentence"). In object position, however, learners generally did not accept bare count singular nouns, and no significant effect of specificity or abstractness was observed.

Ionin et al. (2004) hypothesised that learners from [-art] L1 backgrounds may interpret articles as markers of both definiteness and specificity or of specificity alone and demonstrate a fluctuation pattern<sup>4</sup>. Indeed, forced-choice elicitation (gap-fill task) demonstrated increased use of "the" with [+specific, -definite] (15) compared to [-specific, -definite] (16) referents, as well as increased use of "a" in [-specific, +definite] (17) compared to [+specific, +definite] (18) referents for both L1-Russian and L1-Korean learners.

(15) I am visiting ~~the~~ *[a] friend* from college – his name is Sam Brown, and he lives in Cambridge now<sup>5</sup>.

(16) I am hoping to find *a new good friend*!

(17) We are trying to find ~~a~~ *[the] murderer* of Mr. Peterson – but we still don't know who he is.

(18) Tomorrow, I'm having lunch with *the creator* of this comic strip – he is an old friend of mine.

---

<sup>4</sup> Ionin et al. defined specificity differently from the abovementioned studies, as speaker's intent to refer to an item which the speaker believes to possess a noteworthy property. For example, in "It [his birthday party] was great. He got lots of gifts – books, toys. And best of all – he got *a puppy*!", "a puppy" would be considered non-specific by Ionin et al. because the identity of the puppy is "irrelevant for the discourse" (2004, p. 23). Previously discussed authors (Butler, 2002; Huebner, 1985; Thomas, 1989 and others) would consider this example specific because "a puppy" refers to a certain puppy existing in the world.

<sup>5</sup> In all examples nominals of interest are italicised. Article corrections are given in square brackets. Erroneous articles are struck through.



(from Ionin et al., 2004)

Replications of Ionin et al. (2004) showed similar fluctuation patterns in [-art] L1-Japanese learners (R. Hawkins et al., 2006; Reid et al., 2006) and no fluctuation in [+art] L1 learners, namely Spanish (García Mayo, 2009; Ionin et al., 2008; Reid et al., 2006; Ting, 2005) and Greek (R. Hawkins et al., 2006). However, Ionin et al. (2008) observed the effect of specificity in L1-Russians only in indefinite contexts, i.e. (15) but not (17).

By contrast, Ting (2005) did not find evidence of fluctuation in L1-Mandarin-Chinese learners. Snape et al. (2006) suggest (based on Li & Thompson, 1981) that this may be because Mandarin Chinese is developing equivalents of the English definite (demonstrative “nei” – “that”) and indefinite (numeral “yi” – “one”) articles. However, in a further replication experiment Snape (2009) did find the expected fluctuation in a larger group of L1-Chinese intermediate learners.

To establish specificity, i.e. speaker’s intent to refer to a noteworthy item, Ionin et al. included explicit statements of speaker knowledge, which could be construed as speaker familiarity with the referent, e.g. “He is meeting with *the* director of his company. *I don’t know who that person is*” [emphasis added] (Trenkic, 2008, p. 13). Trenkic’s replication of Ionin et al. showed that the combined effect of specificity and explicit speaker knowledge statements (e.g. giving the name or personal details of the referent) was significantly stronger than the effect of specificity alone.

Trenkic (2007) offers an alternative explanation claiming that learners from [-art] L1 backgrounds may misanalyse articles as adjectives (therefore, optional elements) taking “the” and “a” to mean “definite/identifiable” and “indefinite/unidentifiable”, respectively. In her spoken and written data from L1-Serbian English learners, she observed increased omission of “a” and “the” with adjectives. Trenkic concluded that articles in [-art] L1 learners are not syntactically motivated but produced as lexical items motivated by the pragmatic need to express the meaning learners assign to them. Under this account, the effect of specificity is due to learners erroneously assigning the meaning “specific/non-specific” to “the/a” in addition to or instead of the meaning “definite/indefinite”. Trenkic further argues that articles are more likely to be dropped with

premodified nouns because producing additional lexical items drains working memory resources leading to omission of less (communicatively) important items. Butler's (2002) participants also commented that adjectives made nouns seem more "specific", making articles appear redundant.

To summarise, studies of naturalistic and elicited data from participants with various L1s have shown that learners with [-art] L1s might be uncertain about whether articles in English mark specificity or definiteness, although the effect is not always observable (Hua & Lee, 2005; Ting, 2005; White, 2003).

### ***The Role of Discourse***

Highly salient antecedents lead to article omission in subsequent discourse. For example, Robertson (2000) and Trenkic & Pongpaiboj (2013) found that article omission in [-art] L1 learners (Chinese and Thai, respectively) was more likely if the antecedent was mentioned in the immediately preceding sentence or visually cued.

Liu & Gleason (2002) found that learners with different L1s were more likely to omit "the" in anaphoric contexts than in other discourse-pragmatic uses of "the" (using Hawkins's classification, 1978), except for larger situation (cultural) and idiomatic uses, which were the most difficult according to the results of their gap-fill task.

### ***The Role of Nominal Features***

Many of the abovementioned studies report that "the" is acquired earlier than "a". Based on a feature-assembly approach, Lardiere (2004) suggests that the indefinite article is more complex, as it includes number and countability features, while the definite article does not (also R. Hawkins et al., 2006).

Snape (2008) combines the investigation of the role of discourse functions (using Hawkins's classification, 1978) and nominal features (countability and number). He argues that learners with L1s without a count/mass distinction (e.g. Japanese) may associate definiteness with number features, e.g. use articles only with count singulars, since bare mass and plural (but not count singular) nouns are allowed in English. Using a forced-choice elicitation task, Snape found that

advanced L1-Japanese learners continued omitting “the” with plural and mass nouns in most discourse contexts, while advanced L1-Spanish learners tended to omit “the” (but to a lesser extent) only with plurals in larger situation (cultural) contexts.

There is indirect evidence that learners may be less accurate in their article choice with abstract (e.g. “environment”) vs. concrete (e.g. “book”) nouns, as it might be harder to identify an abstract noun as count or mass. The effect of abstractness on countability judgments was found in Hua & Lee’s (2005) study, whose L1-Chinese participants were more likely to accept ungrammatical combinations of abstract count nouns with “much” as opposed to concrete count nouns, i.e. they were more likely to accept “much sentence” than “much computer”. In Butler (2002), up to 20% of L1-Japanese learners’ errors were due to misdetection of noun countability, especially where it depended on context or where the nouns were abstract.

In sum, number and countability may affect article accuracy of learners from [-art] L1 backgrounds, who may be misled by the absence of articles before mass and plural indefinites in English and generalise this pattern to all mass and plural nouns. The acquisition of the count/mass distinction, especially for abstract nouns, further impacts on the article accuracy of learners lacking grammatical number in their L1.

### **Research Questions**

Based on the previous research and the cross-linguistic differences, multiple factors may influence how learners interpret and use articles in English, including learners’ [+/-art] L1 and proficiency level in English, semantic features, which are not encoded by English articles (specificity, familiarity, abstractness), morphosyntactic and syntactic features (number, countability, syntactic position, premodification), and the articles’ discourse functions.

The goal of the study reported in this chapter is to investigate how these factors work together and interact to predict article accuracy and article error types (omission, substitution, overuse) in learners of L2 English. The aim is to assess the relative strength of influential factors and,

thus, gain insight into how grammatical (morphosyntactic) factors interact with semantic and discourse-pragmatic factors.

## **Methodology**

### ***Learner Data***

**Corpus.** For this analysis, I used a large open-access learner corpus, EFCAMDAT, the EF Education First Cambridge Open Language Database (Geertzen et al., 2013). EFCAMDAT contains writings submitted in response to communicative tasks, such as writing a holiday postcard, a film review, describing a terrifying experience etc., to EF's online language school. EFCAMDAT is organised across 16 proficiency levels aligned with the Common European Framework of Reference for languages (CEFR). The corpus is pseudo-longitudinal, as most learners do not complete all the levels. New learners are placed in levels 1/4/7/10/13 based on their result on a placement test. Each level comprises 8 modules each ending with a writing task. The corpus contains 1,180,310 writings (scripts) from learners around the world (198 nationalities). National Language (NL), crossing nationality with country of access to the online school, is used as a proxy for L1 in EFCAMDAT, since direct information on L1 background is not collected, which inevitably introduces noise in the data (e.g. multilingualism is not captured). Nevertheless, NL has been shown to be a reliable proxy for L1 (Alexopoulou et al., 2015; Murakami, 2013). For more information on the corpus see Alexopoulou et al. (2015), Alexopoulou et al. (2017), Michel et al. (2019).

One reason for choosing this corpus, aside from its large size and availability of data from learners with a variety of NLs, is that it can potentially reveal issues with article use that typically arise in EFL settings. This is useful in view of the overarching goal of this dissertation, which is to make a link between linguistic analysis and the practical needs of learners and teachers.

**Sampled Subcorpus.** I focused the analysis on four learner subgroups, Brazilian Portuguese (henceforth, Brazilian), German, Chinese, and Russian, since the typological differences between them allow for a comparison of [+art] (Brazilian and German) and [-art] (Chinese and Russian) NLs. I included two example languages from each language type to tease apart the typological effect of the

presence/absence of articles from potentially L1-independent effects (Murakami & Alexopoulou, 2016a).

In this study, I used a total of 660 scripts (165 per NL, 72-76 scripts per EF level<sup>6</sup>), summing to 67,120 words. I will describe the sample in more detail after I introduce the coding scheme.

### **Coding**

To measure the accuracy with which learners use English articles, I first manually retrieved all nominals from the scripts. I treat as a nominal any phrase consisting of a noun, an optional article, and a prenominal modifier, e.g. “book”, “an interesting book”, “the books”, but excluding demonstratives and quantifier items, such as “this book”, “many interesting books”.

**Exclusions.** In the retrieval process the following types of instances were not included.

- Proper nouns, since the use of articles with these is highly idiomatic and often acquired on an item-by-item basis (e.g. “Bill”, “Hillary”, but “the Clintons”).
- Second/third etc. nouns in sequences of coordinate nouns where the article is not obligatory. For example, in (19) only the noun “roads” was included because in such cases it is unclear whether the learner has mastered noun phrase coordination or made an omission error.

(19) The main justification for this car tax is the need for building and maintaining *the roads* and highways. (L1-German, B2, ID 966727)<sup>7</sup>

- Gerunds in examples like “do the shopping” as it is not clear if learners conceive these as nouns or verb forms in their writing (however, some well-established nouns derived from verbs were included, e.g. “warning”, “understanding”).
- Cases where the noun is omitted (e.g. “I’m a happy.”).

---

<sup>6</sup> A slightly larger number of scripts was drawn from lower levels to account for the fact that scripts tend to be shorter at lower levels.

<sup>7</sup> For all the corpus examples I provide the learner’s L1, CEFR level, and writing ID. The original spelling and grammar are preserved. Nominals of interest are italicised. Article corrections are given in square brackets. Overused articles are struck through.

- Nominals which are not part of a full clause or appear to be written in telegraphic style, e.g. bullet points, notes (20).

(20) TWO different aspects of security • technical-security - offered options; • user-security - private responsibility. (L1-German, B2, ID 576446)

- Incomprehensible instances.

Because I decided to focus on the use of articles specifically rather than determiners more generally, I also excluded the following types of instances.

- Nominals introduced by the demonstratives “this”, “that”, “these”, “those”, possessive pronouns (e.g. “my”, “her” etc.) or nouns (e.g. “John’s”, “teacher’s” etc.), all of which are in complementary distribution with articles (although there were instances of article overuse in such cases (e.g. “the my book”), these were extremely rare).
- Nominals introduced by other determiners (“some”, “any”, “no”, “many”, “much”, “more”, “less”, “most”, “least”, “such”, “each”, “every”, “both”, “enough”, “several”, “all”, “another”, “other”, “what”, “which”). A rough estimate shows that there is up to a thousand of such instances in the subcorpus out of a total of approximately 13,000 nominals (this total number excludes proper nouns), around 7.7% of the total. Such instances were excluded regardless of whether an article was supplied in addition to the determiner (21), although in some cases this may represent correct usage, e.g. “all the work”, “the other option”, “such a reporter”.

(21) The most entrees are \$15. (L1-Russian; A2, ID 973419)

- Nominals introduced by quantifiers “a lot of”, “a little”, “a few”, “a (little) bit of”, which include articles in them, since these are likely to be used as chunks.
- Nominals introduced by cardinal numbers; although the definite article is possible in such cases, I did not encounter a single instance in levels A2 or B1 where it would be required (I then stopped keeping track of these for level B2).

- Cases where an article is incorrectly used but where the target is not an article, since such cases do not represent obligatory article contexts (e.g. in “she gave me *a son’s* clothes” “her” should be used instead of “a”).

**Formulaic Sequences.** When analysing accuracy in learner production data from an EFL context, it is important to recognise the role of formulaic sequences (FSs), which are commonly used by learners to bridge the gap between their limited productive grammars and their communicative needs (Myles, 2012). A FS can be defined as “a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar” (Wray, 2002, p. 9). FSs play an important role in language acquisition (L. W. Fillmore, 1979; Myles, 2012), however they cannot be taken as evidence of acquisition of the individual elements appearing in them. I, thus, decided to mark FSs in the subsample and exclude them from this analysis. I expect that the accuracy of article use will be significantly higher in FSs as opposed to language generated by productive grammars.

Although the definition of a FS is clear, the identification of FSs in corpus data is less straightforward, since in a corpus one only sees the result of learner production but not the process whereby it was generated or retrieved. A more practical operationalisation is offered by Alexopoulou et al. (2015), who (building on O’Donnell et al., 2013) used multiple measurable indicators to identify FSs, namely ranking word n-grams by frequency, number of learners who used them, length, and word association strength, as well as comparing the complexity of a given learner’s productive language with the complexity of their FSs, and considering the variation within FSs across learners. This approach is suitable for the analysis of relative clauses, which are the focus of Alexopoulou et al.’s paper, but it is hard to see how it could be applied to the case of nominals. It may be applicable to some examples of nominals embedded in larger FSs, such as “to make a long story short”. However, there are numerous instances of nominals which appear to be retrieved

whole from memory rather than generated but are too short for the implementation of the word association strength approach, e.g. “in the morning” or “at night”.

Taking into account that the distinction between FSs and productive language is not dichotomous (Alexopoulou et al., 2015; Myles, 2012) and that I am manually coding a relatively small dataset, I decided to identify FSs manually on a case-by-case basis. In this process, I took the following factors into consideration:

- How frequent the sequence is across learners
- Whether the sequence is a relatively fixed expression in English
- Whether the sequence is typically taught as a set phrase in EFL contexts
- Whether the use (or non-use) of the article in the sequence is largely conventional and independent of context

The 11 most frequent formulaic sequences are listed in Table 1.

**Table 1**

*Most Frequent Formulaic Sequences in the Subcorpus*

| Formulaic sequence                                     | Number of occurrences |
|--|-----------------------|
| next/last + temporal modifier (e.g. year, month, time) | 52                    |
| for example  | 28                    |
| at/to/after etc. school                                | 23                    |
| at/to/after etc. work                                  | 22                    |
| kind/type/sort of N (e.g. type of film, kind of car)   | 20                    |
| a/an + temporal modifier (e.g. twice a week)           | 18                    |
| at/to/after etc. university                            | 16                    |
| at/to the beach  | 14                    |
| at home  | 12                    |
| in addition  | 10                    |
| all over/around the world                              | 10                    |

The full list of FSs identified in the data is provided in Appendix A.

**Copies from Task Prompts and Model Answers.** In addition to separating FSs, I also considered sequences provided in task prompts and model answers, which were often copied by



learners in their writing. If a nominal was copied from the prompt or model answer within a larger copied sequence or within the same context, it was coded as a copy and excluded from the analysis. For example, one task prompt contained a job advert, which included the following: “Benefits: [...] regular trips to Florida to observe wild crocodiles”. In this case, all of the following instances of the nouns “trips” and “crocodiles” were coded as copies:

(22)The salary and the benefits are absolutely gorgeous: [...] and the most amazing among them: *regular trips* to Florida to observe *wild crocodiles*.

(23)Flossington zoo is offering [...] *trips* to Florida - to observe *crocs*.

(24)you will have [...] *regular trips* to Florida to observe ~~the~~ *wild crocodiles*.

(25)The benefits is [...] *regular trips* to Florida to observe *wild crocodiles*.

(26)Furthermore, you are going to have *regular trips* to Florida to observe *crocodiles*!

Note that example (24) contains an article error (overuse of “the”) but is still coded as a copy because the whole sequence appears to be lifted from the task prompt. I do not expect copies to be completely error free, but, as with FSs, they are likely to have a significantly higher accuracy rate.

**Coding Variables.** All the coded variables are listed in Table 2 along with examples and notes.

**Table 2**

*Corpus Data Variable Coding*

| Variable             | Levels   | Note/Examples  |
|----------------------|--|--|
| NL                   | de – German, br – Brazilian,<br>cn – Chinese, ru – Russian |  |
| Level                | 25-96 (72 total)   | Each number on the scale represents one module. Eight modules comprise one EF level, e.g. modules 25-32 comprise EF level 4. Three EF levels cover one CEFR level, e.g. EF levels 4-6 cover CEFR level A2. |
| Topic ID<br>(prompt) | 1-123  | Each module ends in a writing prompt, referred to as “topic” in the corpus. There  |

|                                 |  |   |
|---------------------------------|--|---|
|                                 |  | are more topic IDs (123) than the total number of modules (72) because some modules had two prompt options (see Shatz, 2020).   |
| Definiteness                    | definite/indefinite  |   |
| Target article                  | "a"/"the"/ $\emptyset$   |   |
| Response                        | correct/incorrect  | Correct if "a"/"the"/ $\emptyset$ is used appropriately   |
| Error type                      | omission of: "a", "the";<br>overuse of: "a", "the";<br>substitution of "a" instead of "the",<br>substitution of "the" instead of "a" | Omission: failure to supply "a"/"the" in obligatory "a"/"the" contexts;<br>Overuse: use of "a"/"the" in target $\emptyset$ contexts;<br>Substitution: use of "a" in obligatory "the" contexts or use of "the" in obligatory "a" contexts. |
| Noun type                       | count singular<br>count plural<br>mass   | Considered context: "cake" count in "they ate a cake", mass in "they ate cake"  |
| Abstractness                    | abstract/concrete  | Coded as binary, but I acknowledge that it is a gradient characteristic (Scott et al., 2019).<br>Considered context: "things" concrete in "I saw interesting things", abstract in "I learned interesting things".                         |
| Syntactic position <sup>8</sup> | a. subject<br>b. object<br>c. predicate<br><br>d. existential  | including objects of verbs and of prepositions<br>nominals following copula "be" (and "become").<br>Included comparative constructions with "as" and "like" (e.g. "as tall as <i>a giraffe</i> ", "work like <i>a robot</i> ").           |

<sup>8</sup> Other syntactic positions were excluded as they constituted less than 10% of the data (temporal modifiers, e.g. "last week", appositives, e.g. "Tom, the leading man", genitives, e.g. "people" in "people's attitudes").

|                             |  |   |
|-----------------------------|--|---|
|                             |  | nominals following “there” + “be” constructions, e.g. “ <i>there is a book/there are books/there is dust on the table</i> ” |
| Specificity                 | specific/non-specific                                | Definition discussed in Chapter 2   |
| Discourse-pragmatic context | anaphoric, situational, explanatory, kind, idiomatic | Classification described in Chapter 2   |
| Prenominal modifier         | present/absent                                       | e.g. “a big book”, “a science book”   |

To determine coding reliability, 100 randomly selected items were coded by another doctoral student of Applied Linguistics (English native speaker). The level of agreement was strong for most variables,  $\kappa > 0.85$ , and moderate for discourse-pragmatic context,  $\kappa = 0.78$  (McHugh, 2012).

Below I provide an example of a coded script from an L1-Brazilian learner at B1 level, EF level 8, task 1 “Reviewing a song for a website”, writing ID 1017525. All the nominals are numbered and italicised. The codes for each nominal are given in the table below. Excluded instances show “n/a” codes, and the reason for exclusion is provided in the “Comment” column.

I just listened (1) *home Improvement*, (2) *the song* by (3) *Josh Woodward*. This is (4) *a good song*, but without (5) *cool resources* like (6) *guitars and Drums solo*. I believe that (7) *a good music* to hear around (8) *the beach*, with (9) *a something friends* during (10) *a barbecue*, for me (11) *this music* looks like (12) *a reggae* with (13) *mix* (14) *country music*. (15) *The letter* talk about (16) *the age* when (17) *the people* lived in (18) *the countryside* working like (19) *a farmer* and how is different (20) *the nowadays life style*. On (21) *short words*, regular but (22) *good music* to listened and reflect about what we will need for us and (23) *our family* on (24) *the future*.

**Table 3***Example Coded Learner Script*

| No. | Correct / Error type        | Noun type    | Specif.      | Abstr.   | Syntactic position   | Modif. | Discourse-pragmatic context | Comment  |
|-----|-----------------------------|--------------|--------------|----------|----------------------|--------|-----------------------------|--|
| 1   | n/a                         | n/a          | n/a          | n/a      | n/a                  | n/a    | n/a                         | proper name                                    |
| 2   | error: "the" instead of "a" | count sing.  | specific     | concrete | appositive           | no     | n/a                         |  |
| 3   | n/a                         | n/a          | n/a          | n/a      | n/a                  | n/a    | n/a                         | proper name                                    |
| 4   | correct: "a"                | count sing.  | specific     | concrete | predicate (property) | yes    | n/a                         |  |
| 5   | correct: ∅                  | count plural | non-specific | abstract | object               | yes    | n/a                         |  |
| 6   | error: omitted "a"          | count sing.  | non-specific | abstract | object               | yes    | n/a                         | assuming learner meant "a guitar or drum solo" |
| 7   | error: overused "a"         | mass         | specific     | concrete | predicate (property) | yes    | n/a                         |  |
| 8   | correct: "the"              | count sing.  | non-specific | concrete | object               | no     | idiomatic                   | formulaic                                      |
| 9   | n/a                         | n/a          | n/a          | n/a      | n/a                  | n/a    | n/a                         | possibly meant "some friends"                  |
| 10  | correct: "a"                | count sing.  | non-specific | concrete | object               | no     | n/a                         |  |
| 11  | n/a                         | n/a          | n/a          | n/a      | n/a                  | n/a    | n/a                         | demonstr. "this"                               |
| 12  | error: overused "a"         | mass         | non-specific | abstract | predicate (property) | no     | n/a                         |  |
| 13  | error: omitted "a"          | count sing.  | non-specific | abstract | object               | no     | n/a                         | assuming learner meant "with a mix of"         |
| 14  | correct: ∅                  | mass         | non-specific | concrete | object               | yes    | n/a                         |  |
| 15  | correct: "the"              | count plural | specific     | abstract | subject              | no     | situational                 | assuming learner meant "the lyrics"            |
| 16  | correct: "the"              | count sing.  | specific     | abstract | object               | no     | explanatory                 |  |

| No. | Correct / Error type  | Noun type    | Specif.      | Abstr.   | Syntactic position   | Modif. | Discourse-pragmatic context | Comment  |
|-----|-----------------------|--------------|--------------|----------|----------------------|--------|-----------------------------|--|
| 17  | error: overused “the” | count plural | non-specific | concrete | subject              | no     | n/a                         |  |
| 18  | correct: “the”        | mass         | non-specific | concrete | object               | no     | kind                        |  |
| 19  | correct: “a”          | count sing.  | specific     | concrete | predicate (property) | no     | n/a                         |  |
| 20  | correct: “the”        | count sing.  | specific     | abstract | subject              | yes    | situational                 | assuming learner meant “how the modern lifestyle is different” |
| 21  | correct: ∅            | count plural | non-specific | abstract | object               | yes    | n/a                         |  |
| 22  | correct: ∅            | mass         | specific     | concrete | predicate (property) | yes    | n/a                         | assuming learner meant “this is regular but good music”        |
| 23  | n/a                   | n/a          | n/a          | n/a      | n/a                  | n/a    | n/a                         | possessive “our”   |
| 24  | correct: “the”        | count sing.  | specific     | abstract | object               | no     | situational                 |  |

Note. Specif. – specificity, abstr. – abstractness, modif. – modifier presence, sing. – singular

### ***Mixed-Effects Logistic Regression***

To investigate the effect of the independent variables on article accuracy, I used generalised linear mixed-effects logistic regression models where the dependent variable is binary (correct/incorrect article use), using the `lme4` package in R (Bates, Mächler, et al., 2015; R Core Team, 2021). I created one such model based on the entire subcorpus including all the variables (Full Accuracy Model), except discourse-pragmatic context, and another model on definites only to assess the effect of discourse-pragmatic context, which is only relevant for definite nominals (Definite Accuracy Model).

For further analysis of error types, I used multinomial logistic regression models, which allow for more than two outcomes, using the `mcllogit` R package (Elff, 2021). This type of model estimates predictor variable effects on the change in the odds of the different outcomes (omission,

substitution, overuse) compared to a chosen baseline outcome (no error). In other words, one could explore whether a mass noun, as opposed to a count singular noun, increases the odds of article omission vs. no error.

For error type analysis the data was split into three subsets according to target article<sup>9</sup>, as each has different error possibilities.

- Definite (obligatory “the”, 2039 observations): no error, omission of “the”, substitution (“a” instead of “the”)
- Indefinite count singular (obligatory “a”, 1679 observations): no error, omission of “a”, substitution (“the” instead of “a”)
- Indefinite plural and mass (obligatory  $\emptyset$ , 2060 observations): no error, overuse of “the”, overuse of “a”

Random effects were included in all models because each script contained multiple observations. I also included the topic (prompt) ID as a random effect to capture potential prompt effects where possible.

An explanation of the general approach to and the technical details of model selection along with the details of the model performance metrics used in the corpus study and throughout this dissertation is provided in Appendix B.

### ***Power Analysis and Sample***

To determine the appropriate sample size, I conducted a power analysis in R, for which I simulated a dataset with the following specifications:

- 300 scripts (1 script per learner)
- Two NL-types: [+art] and [-art]
- Three CEFR levels (A2, B1, B2, corresponding to levels 4-6, 7-9, and 10-12, respectively, in EFCAMDAT)

---

<sup>9</sup> The split created difficulties with the random effects structure. In fact, I was only able to include random intercepts by writing ID, since the addition of any random slopes resulted in singular fits.

To assess the approximate distribution of the variables I was interested in, I examined 160 randomly selected scripts from the L1-Brazilian and L1-Russian subcorpora (80 scripts each) between EF levels 4-9 (A2-B1 according to CEFR). This enabled me to determine the average number of scorable nominals per script ( $M = 11.5$ ,  $SD = 3$ ), the average accuracy rate (80%), and the rough distribution and the effects of the features that I wanted to include as independent variables in the regression model, i.e. definiteness, specificity, countability and number, abstractness, syntactic position, and presence of a prenominal modifier. These findings were used to further specify the simulated dataset.

To introduce random variation for the mixed model, I randomly generated 300 normally distributed numbers (one for each learner),  $M = 0$ ,  $SD = 0.15$  log odds, which were added to the means of each respective script ID, which was equivalent to learner ID in this set (as there was one script per learner).

I then fitted a generalised linear mixed-effects regression model onto the simulated dataset. Accuracy of article use in obligatory contexts was the independent variable. Based on the accuracy rates set in the simulated dataset, each item had a generated score of 1 (correct) or 0 (incorrect), resulting in a binary variable. The model included 7 fixed effects:

- NL-type: [+art] or [-art]
- Proficiency level (A2/B1/B2)
- Definiteness (definite or indefinite)
- Specificity (specific or non-specific, as defined in this study)
- Noun type, combining countability and number (count singular, count plural, or mass)
- Modifier (present or absent)
- Abstractness (abstract or concrete)

I also included a two-way interaction between the NL-type and all the independent variables, except for modifier presence, resulting in 5 two-way interactions.

The writing (script) ID was included as a random effect to control for individual variation. The syntactic position was also initially included in the model but did not yield a statistically significant effect for the simulated dataset and was thus not included in the final model for power analysis. The results of the model are presented in Table 4.

**Table 4**

*Simulated Corpus Dataset Model Results*

| <i>Predictors</i>                                    | <b>Accuracy Rate</b> |                   |               |          |
|--|----------------------|-------------------|---------------|----------|
|  | <i>Log-Odds</i>      | <i>std. Error</i> | <i>95% CI</i> | <i>p</i> |
| (Intercept)  | 1.88 ***             | 0.20              | 1.49   2.27   | <0.001   |
| L1: [-art]   | -1.30 ***            | 0.25              | -1.78   -0.81 | <0.001   |
| level: B1  | 0.95 ***             | 0.20              | 0.55   1.34   | <0.001   |
| level: B2  | 1.08 ***             | 0.20              | 0.68   1.47   | <0.001   |
| def: indefinite                                      | -0.11                | 0.23              | -0.56   0.35  | 0.641    |
| spec: non-specific                                   | -0.66 **             | 0.21              | -1.07   -0.24 | 0.002    |
| Ntype: pl  | 0.52 *               | 0.24              | 0.06   0.98   | 0.028    |
| Ntype: mass  | -0.86 ***            | 0.18              | -1.22   -0.51 | <0.001   |
| abstr: concrete                                      | 0.53 **              | 0.17              | 0.20   0.85   | 0.002    |
| mod: mod   | -0.30 **             | 0.11              | -0.52   -0.08 | 0.007    |
| L1: [-art] * level: B1                               | -0.80 **             | 0.25              | -1.29   -0.30 | 0.002    |
| L1: [-art] * level: B2                               | -0.43                | 0.26              | -0.94   0.08  | 0.095    |
| L1: [-art] * def: indefinite                         | 0.65 *               | 0.28              | 0.10   1.20   | 0.021    |
| L1: [-art] * spec: non-specific                      | 0.89 **              | 0.28              | 0.35   1.43   | 0.001    |
| L1: [-art] * Ntype: pl                               | 1.14 ***             | 0.32              | 0.52   1.76   | <0.001   |
| L1: [-art] * Ntype: mass                             | 1.58 ***             | 0.24              | 1.11   2.05   | <0.001   |
| L1: [-art] * abstr: concrete                         | -0.70 ***            | 0.21              | -1.12   -0.29 | 0.001    |
| <b>Random Effects by Writing ID</b>                  |                      |                   |               |          |
| SD (Intercept)                                       | 0.06                 |                   |               |          |
| N <sub>lrm_id</sub>                                  | 300                  |                   |               |          |
| Observations   | 3332                 |                   |               |          |
| Marginal R <sup>2</sup> / Conditional R <sup>2</sup> | 0.189 / 0.190        |                   |               |          |

*Note.* The omitted levels of multi-level variables in this model are “A2” for level, “sing” for Ntype.

I chose the effect with the smallest  $\beta$ -coefficient, which was the estimate for modifier presence (-0.30), as the basis for the power analysis, and reduced it even further to -0.25 to make sure that the sample size is big enough to identify even a smaller effect. I then used the `simr` package (Green & Macleod, 2016) to determine the optimal sample size for the model to have enough power to detect this effect. The package works in three steps. First, it uses the given model to generate a new dataset with the same parameters and the same effects estimates. Second, it



refits the model on the new dataset. Finally, it tests the new model for statistical significance. This process is repeated many times for each given sample size, e.g. I ran at least 200 such simulations per sample size. The resulting statistical power estimate is equal to the percentage of statistically significant models for a given sample size. For example, if out of 100 simulations for a sample size of 300 scripts, only 50 simulations resulted in a statistically significant model, this means that the statistical power for such a sample size is only 50%. A statistical power of 80% is usually regarded as the golden standard in social sciences (Green & Macleod, 2016).

The result of the power analysis showed that a sample of 6664 observations would be needed to have a statistical power of 84.5%, 95% CI [78.73, 89.22]. Given that the average number of scorable nominals per script was estimated at 11.5 (based on preliminary analysis), one would require 579 scripts for the subcorpus. To be on the safe side, I decided to draw a sample of 660 scripts (165 per NL, 72-76 scripts per EF level).

The scripts were randomly selected across proficiency levels from CEFR level A2 to B2, inclusive, which corresponds to EF levels 4-12<sup>10</sup>. I also ensured that scripts were equally distributed across topics, by which I mean the specific writing prompt at the end of each module (e.g. “Write a short autobiography”), and that no more than one script was contributed by the same learner.

The resulting subcorpus contained 5772 nominals, after all the exclusions detailed above (Table 5).

---

<sup>10</sup> I excluded lower levels, where the writings are mostly formulaic, and higher levels, where there is generally much less data in the corpus.

**Table 5***Distribution of Nominals in the Subcorpus Across EF Levels and NLs*

| EF levels    | NL          |             |             |             | Total       |
|--------------|-------------|-------------|-------------|-------------|-------------|
|              | de          | br          | cn          | ru          |             |
| <b>4</b>     | 65          | 85          | 95          | 86          | 331         |
| <b>5</b>     | 107         | 119         | 71          | 96          | 393         |
| <b>6</b>     | 95          | 128         | 148         | 125         | 496         |
| <b>7</b>     | 110         | 153         | 110         | 127         | 500         |
| <b>8</b>     | 177         | 121         | 128         | 158         | 584         |
| <b>9</b>     | 161         | 188         | 131         | 194         | 674         |
| <b>10</b>    | 237         | 254         | 191         | 206         | 888         |
| <b>11</b>    | 246         | 218         | 244         | 247         | 955         |
| <b>12</b>    | 263         | 287         | 138         | 263         | 951         |
| <b>Total</b> | <b>1461</b> | <b>1553</b> | <b>1256</b> | <b>1502</b> | <b>5772</b> |

## Analysis and Results

In this section, I first give an overview of data distribution, observed accuracy rates, and error types.

Then I report the results of regression modelling. First, I regressed article accuracy (as a binary outcome: correct/incorrect) against the independent variables on the entire dataset (the Full Accuracy Model). Second, I fitted the same type of model on definite contexts only to assess the effect of discourse-pragmatic context, which is only relevant for the definite article (the Definite Accuracy Model).

Finally, I report three separate models of error types in (i) indefinite singular contexts (target “a”), (ii) definite contexts (target “the”), and (iii) indefinite mass and plural contexts (target  $\emptyset$ ).

### ***Data Distribution and Observed Accuracy Rates***

The distribution of the observations across key variables is presented in mosaic plots in Figure 1. The top plot shows that the three target articles are approximately equally represented.

As seen in the top left plot, definites tend to be specific, i.e. refer to certain existing entities, and indefinites non-specific (especially mass and plural). The middle left plot shows that definites are less often premodified than indefinites, and mass nouns are not as frequently premodified as count nouns irrespective of definiteness. As seen in the bottom left plot, mass indefinites tend to be abstract, whereas in other categories abstract and concrete nouns are equally distributed. Turning to the top right plot, definites are more often used in subject position than indefinites. The bottom right plot shows that over half of all discourse-pragmatic contexts for the use of “the” are situational. Figure 1 demonstrates that while there are a number of tendencies and potential correlations between semantic/discourse functions and individual articles, there is no categorical or one-to-one mapping, thus confirming the learning challenge.

**Figure 1**

*Distribution of Observations across Key Variables: Target Article, Specificity, Modifier, Abstractness, Syntactic Position, Discourse-Pragmatic Context*

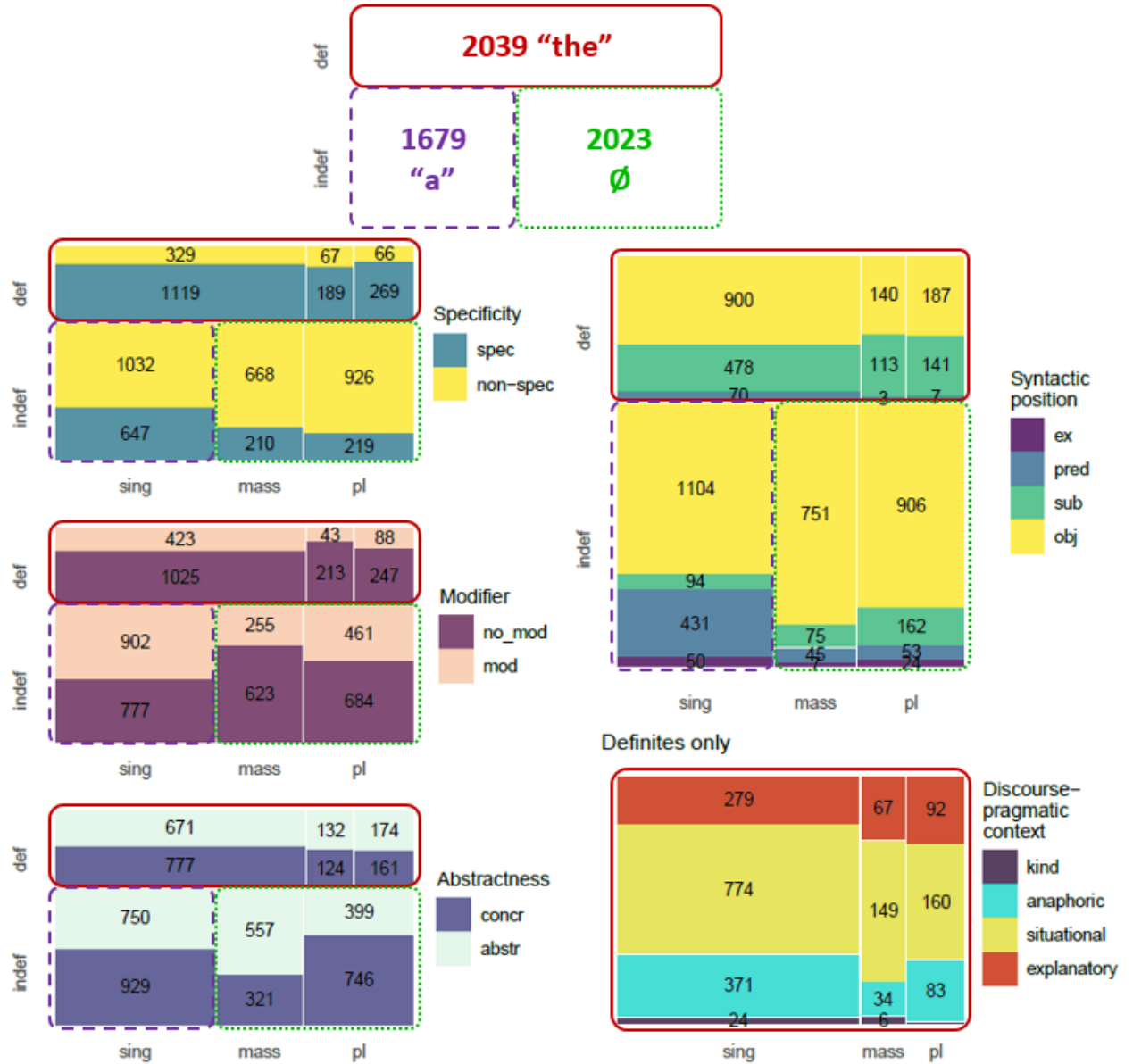


Figure 2 (right) demonstrates article accuracy development across EF levels as measured by the target-language use (TLU) calculated according to the following formula (Pica, 1983), which excludes correct  $\emptyset$  contexts (e.g. "I bought milk") but includes overuse errors (e.g. "I bought a milk")<sup>11</sup>:

<sup>11</sup> In the formula, correct article supplants are appropriate uses of "a/the", obligatory contexts are contexts where "a/the" is required, overuse errors are instances of overusing "a/the" where  $\emptyset$  is required.

$$TLU = \frac{\text{correct article suppliance}}{\text{obligatory contexts} + \text{overuse errors}}$$

**Figure 2**

*Development across EF Levels as Measured by Accuracy Rate (Left) and TLU (Right)*

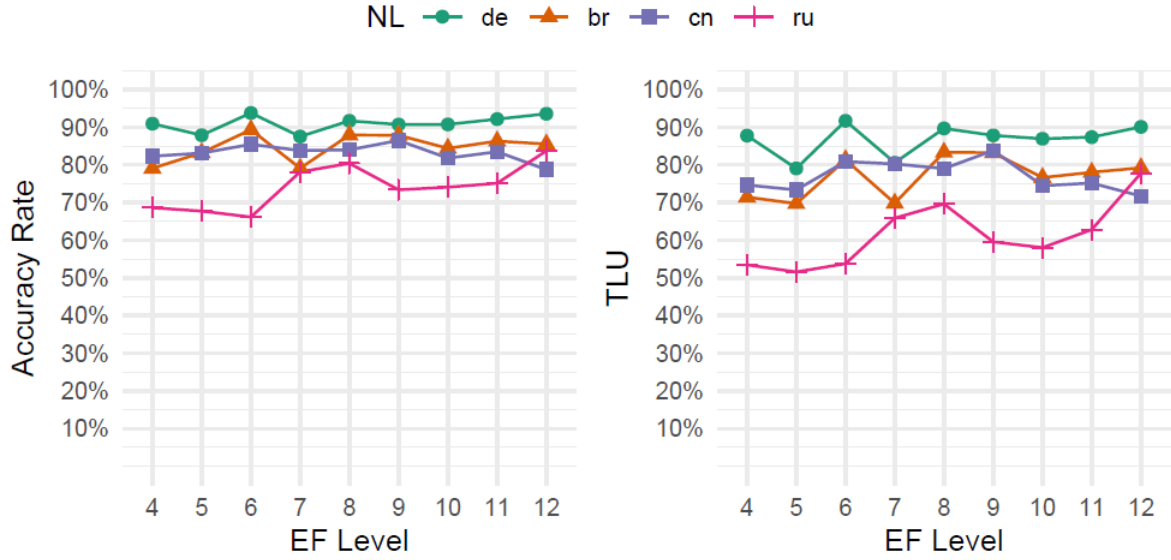


Figure 2 (left) shows accuracy rate measured as the number of correct uses (including correct  $\emptyset$ ) divided by the total number of obligatory contexts (including those where the target is  $\emptyset$ ). Here  $\emptyset$  is just another option, which can be used correctly or incorrectly, whereas TLU only penalises errors in target  $\emptyset$  contexts without rewarding correct  $\emptyset$ , which explains the lower scores in the TLU graph.

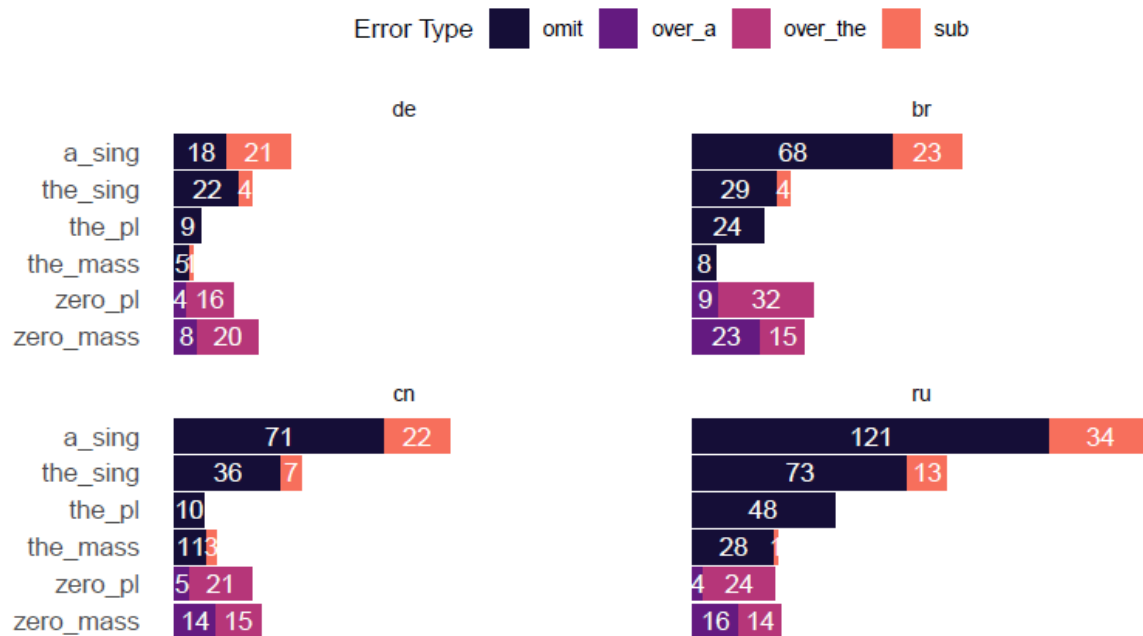
Note that L1-Chinese and L1-Russians, both [-art], do not pattern together (see Murakami & Alexopoulou (2016b) for similar findings). Thus, I further analyse each NL separately without combining them into [+/-art] types.

The distribution of error types is detailed in Table 6 and Figure 3.

**Table 6***Error Type Distribution*

| Error type   | Number of errors | % of total errors |
|--------------|------------------|-------------------|
| Omission     | 581              | 61%               |
| Overuse      | 240              | 25%               |
| Substitution | 133              | 14%               |
| <b>Total</b> | <b>954*</b>      | <b>100%</b>       |

\*16% of all observations

**Figure 3***Error Type Distribution across NL, Target Article, and Noun Type*

*Note.* a\_sing – count singular indefinites (target “a”), the\_sing – count singular definites (target “the”), the\_pl – count plural definites (target “the”), the\_mass – mass definites (target “the”), zero\_pl – plural indefinites (target  $\emptyset$ ), zero\_mass – mass indefinites (target  $\emptyset$ )

In obligatory article contexts, 81% of errors are omissions (see Table 3 above for examples). In target  $\emptyset$  contexts, overuse of “a” and “the” is mostly equal for mass nouns, whereas “the” overuse is more common in count plurals. Note that in general the pattern of errors in each bar is the same across all NLs, the difference is quantitative. The only exception is the use of “a” (a\_sing)

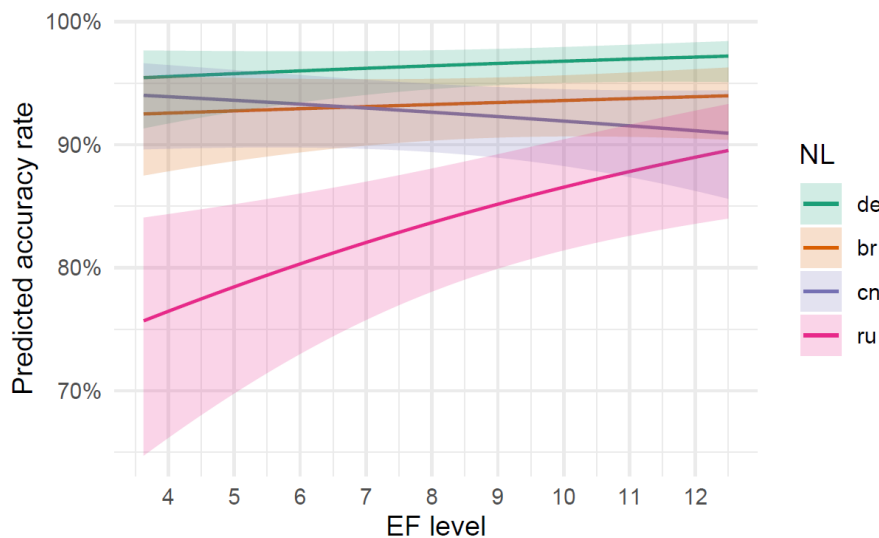
where Germans seem to have a higher proportion of substitution errors than other NL groups, but the overall number of errors for L1-Germans is small.

### ***Predictors of Accuracy in Definite and Indefinite Contexts***

The Full Accuracy Model reveals significant effects of NL, specificity (as reference to a certain existing entity), and modifier presence, which vary by target article and noun type, as well as effects of proficiency level, syntactic position, and abstractness<sup>12</sup>. The full list of estimates produced by the model along with model performance metrics is presented in Table 7.

**Figure 4**

*The Effect of Level across NLs (with 95% Confidence Intervals)*



**Proficiency Level.** Proficiency level has a significant effect only on L1-Russians ( $p = 0.005$ ), whilst other NLs start at ceiling when results are averaged over other variables (Figure 4, note that the scale starts at 65%). Nevertheless, L1-Germans, L1-Brazilians, and L1-Russians all show an upward trend, while the slightly downward trend of L1-Chinese is not statistically significant ( $p = 0.3$ ). Indeed, post-hoc pairwise comparisons do not show any significant differences between the four slopes at  $p < 0.01$ .

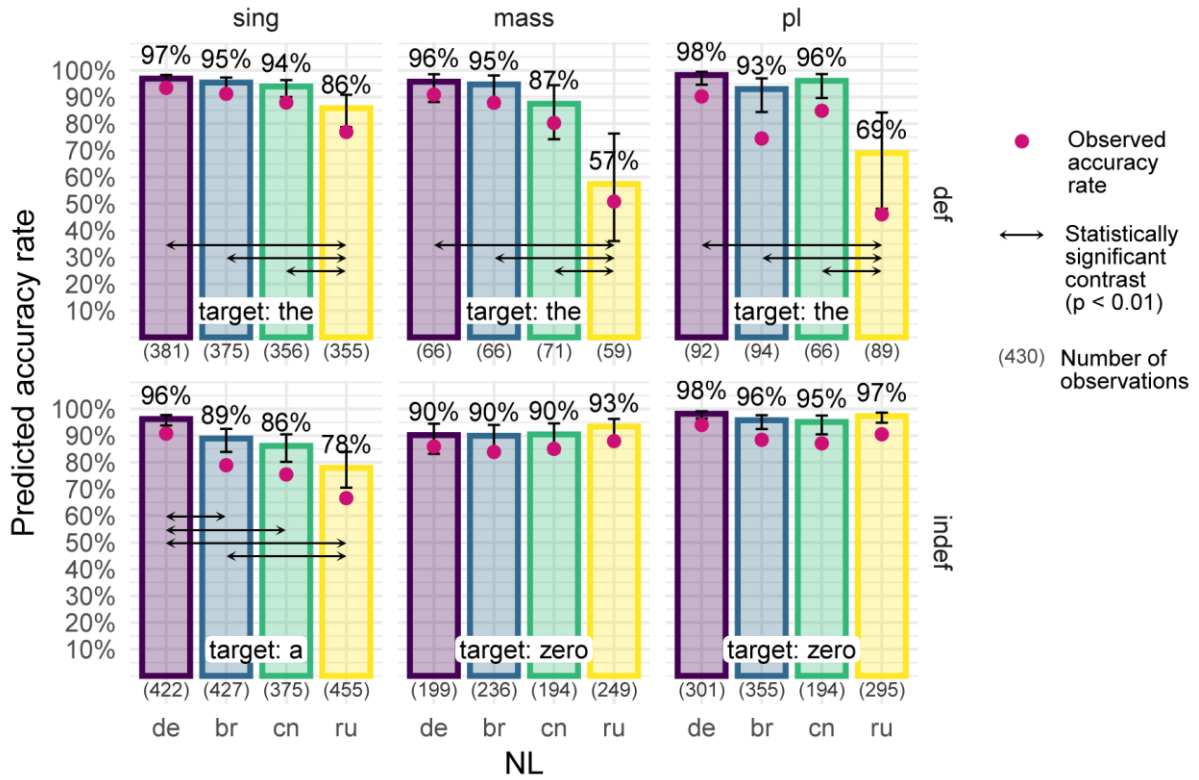
<sup>12</sup> The effect of abstractness, although statistically significant at  $p = 0.02$ , is very small (92% predicted accuracy for concrete nouns, 95% CI [89%, 94%], and 94% predicted accuracy for abstract nouns, 95% CI [91%, 95%]) and, thus, not discussed in more detail here.

**NL.** Figure 5 illustrates the effect of NL on article accuracy across definiteness and noun type.

There is a significant interaction between the three variables, stemming from the fact that the target is different across combinations of variable levels (“the” for all definites, “a” for count singular indefinites, and  $\emptyset$  for mass and plural indefinites). Thus, the top three facets and the bottom left facet of Figure 5 reflect the rate of suppliance of “the” and “a”, while the bottom middle and right facets show accuracy rates in  $\emptyset$  contexts, where any errors would be overuse.

**Figure 5**

*The Effect of NL across Definiteness and Noun Type*<sup>13</sup>



All but L1-German learners are less accurate with “a” than with “the” in singular contexts.

The contrast is statistically significant for L1-Brazilian and L1-Chinese learners ( $p = 0.002$  for both)

but barely so for L1-Russians ( $p = 0.047$ <sup>14</sup>).

<sup>13</sup> In all the figures, error bars represent 95% confidence intervals.

<sup>14</sup> I consider post-hoc pairwise comparisons to be statistically significant at  $p < 0.01$  rather than  $p < 0.05$  to adjust for multiple testing and reduce the risk of Type I error.



The main NL effect concerns the significantly lower accuracy of L1-Russians in obligatory “the” and “a” contexts, which drops further in definite mass (27) and plural (28) contexts, showing a sensitivity to the type of noun not observed in the other NLs. Note that the number of observed mass and plural definites is much lower than that of indefinites and singular definites.

(27) My dinner was horrible! [...] *[The] Red wine* was too sour and *[the] coffee* was bitter.

(L1-Russian, A2, ID 417629)

(28) This property is ideal for investors on a Greek Island. *[The] Apartments* may be fitted with what you would like.

(L1-Russian, B2, ID 114705)

There are some weaker NL effects, namely L1-Brazilian, L1-Chinese and L1-Russian learners are all predicted to be less accurate than L1-Germans in singular indefinites, while L1-Russians are also significantly less accurate than L1-Brazilians (29-31).

(29) It is [...] an example for other women and 'prove' that is possible to get *[a] good position* [...].

(L1-Brazilian, B2, ID 1030969)

(30) Then I will take *[an] air plane* to Madrid.

(L1-Chinese, B1, ID 3441)

(31) I think it helps people to get *[a] well-paid and interesting job* in the future.

(L1-Russian, B1, ID 786665)

In target Ø contexts, one cannot distinguish between genuinely correct article non-use and coincidentally correct article omission, which is likely the case for L1-Russians.

**Specificity.** Specificity, as defined in this study, affects accuracy only in indefinite singulars (Figure 6, left), where accuracy is significantly lower ( $p = 0.001$ ) for non-specific (32) than for specific (33) indefinites. There are no interactions with NL.

(32) I have many dreams [...] I'd make *[a] career* in my business and have a fulfilled and balanced life.

(L1-German, B1, ID 249369)

(33) Alexander had *an accident* last summer when he was arrived cinema.

(L1-Russian, A2, ID 807885)

**Modifier Presence.** A prenominal modifier (Figure 6, right) decreases accuracy in singular definites (34) and mass indefinites (35). Note, however, that in the case of definite singulars a modifier increases article omission but has the opposite effect in mass indefinites, increasing overuse errors.

(34) I first met my friend, Kolya, when I was working in advertising project five years ago. [...]

Kostya and I enjoy working on *[the] advertising project* together.

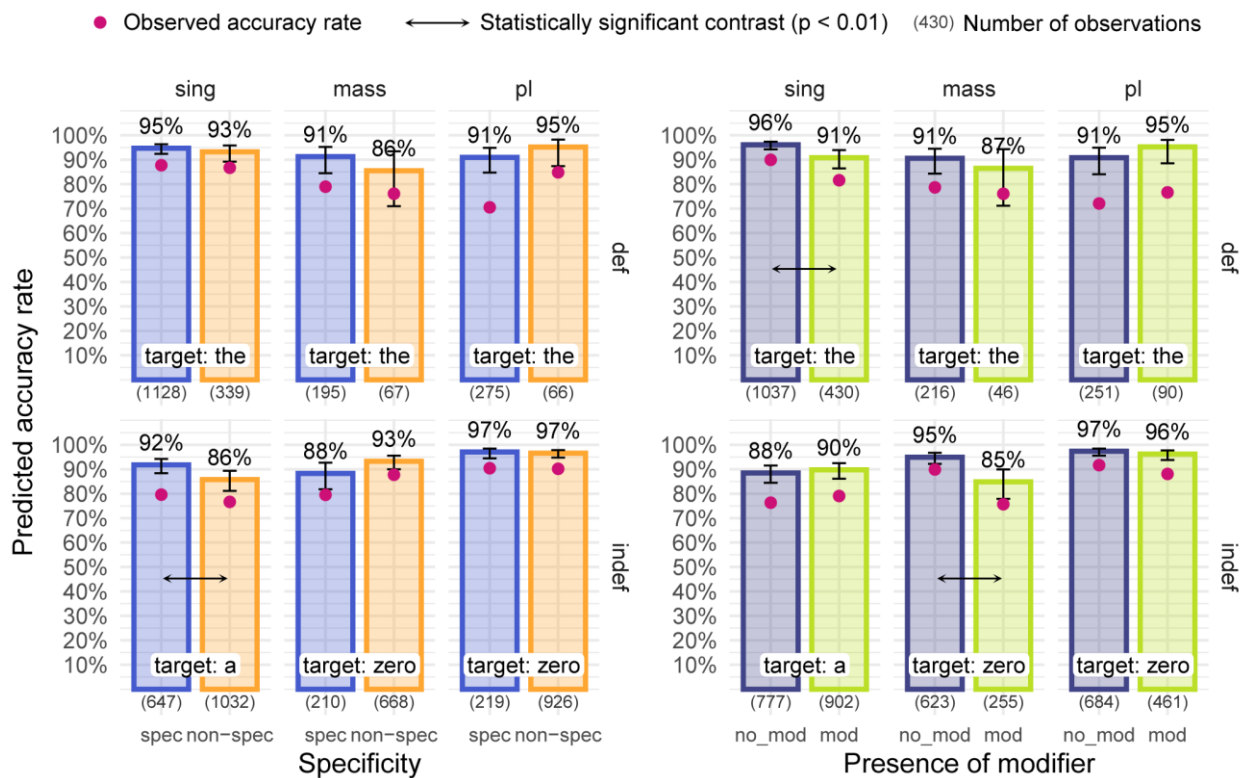
(L1-Russian, A2, ID 887991)

(35) Nowadays there is *a [Ø] great respect* and not be seen anyone smoking indoors.

(L1-Brazilian, B2, ID 664765)

**Figure 6**

*The Effect of Specificity (Left) and Modifier Presence (Right) across Definiteness and Noun Type*



**Syntactic Position.** Errors are significantly more likely in subject and object positions (both at 89% predicted accuracy, 36-37) than in predicate (94% predicted accuracy (38)), or existential (96% predicted accuracy, 39) positions at  $p < 0.01$  (Figure 7, note that the scale starts at 60%). This effect does not interact with any other variable.

(36)[An] *Online study program* give me opportunitaty to learn when I have free time and desire.

(L1-Russian, B1, ID 15851)

(37)I like watching them [reality TV programmes] [...]. I can learn ~~the~~ [Ø] *life experience* from other people.

(L1-Chinese, B1, ID 135026)

(38)[...] we were supposed to have a cosy and comfortable vessel but in fact that was just a *terrible little boat*.

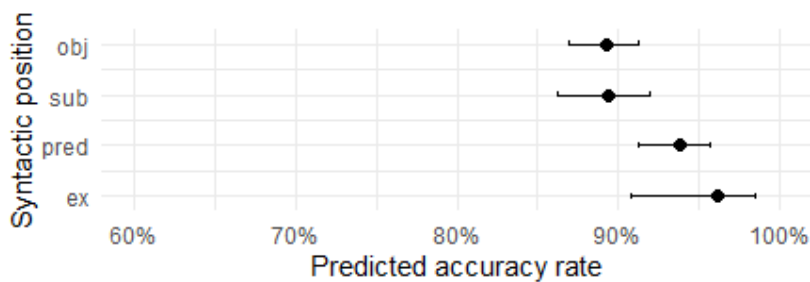
(L1-Chinese, B1, ID 372641)

(39)First, there was *an insect* in my soup!

(L1-Russian, B1, ID 157548)

**Figure 7**

*The Effect of Syntactic Position*



**Table 7**

*Full Accuracy Model Results and Performance Metrics*

| Predictors             | Log-Odds | Accuracy Rate |              |
|------------------------|----------|---------------|--------------|
|                        |          | std. Error    | 95% CI       |
| Intercept (grand mean) | 2.55 *** | 0.16          | 2.23   2.88  |
| def: definite          | -0.06    | 0.09          | -0.23   0.10 |
| Ntype: singular        | -0.12    | 0.10          | -0.33   0.08 |

| <i>Predictors</i>                                  | <i>Log-Odds</i> | <b>Accuracy Rate</b> |               |
|--|-----------------|----------------------|---------------|
|  |                 | <i>std. Error</i>    | <i>95% CI</i> |
| Ntype: mass  | -0.36 **        | 0.13                 | -0.61   -0.10 |
| NL: German   | 0.79 ***        | 0.13                 | 0.53   1.04   |
| NL: Brazilian                                      | 0.10            | 0.11                 | -0.12   0.32  |
| NL: Chinese  | -0.06           | 0.12                 | -0.30   0.17  |
| spec: specific                                     | 0.03            | 0.08                 | -0.13   0.18  |
| mod: no modifier                                   | 0.17 **         | 0.07                 | 0.04   0.30   |
| abstr: concrete                                    | -0.12 *         | 0.05                 | -0.22   -0.02 |
| synt: existential                                  | 0.68            | 0.36                 | -0.02   1.37  |
| synt: predicate                                    | 0.17            | 0.17                 | -0.16   0.51  |
| synt: subject                                      | -0.42 **        | 0.16                 | -0.72   -0.11 |
| Level  | 0.09            | 0.06                 | -0.02   0.20  |
| def: definite * Ntype: singular                    | 0.39 ***        | 0.10                 | 0.20   0.58   |
| def: definite * Ntype: mass                        | -0.07           | 0.12                 | -0.30   0.17  |
| def: definite * NL: German                         | 0.25            | 0.13                 | -0.00   0.50  |
| def: definite * NL: Brazilian                      | 0.25 *          | 0.11                 | 0.03   0.47   |
| def: definite * NL: Chinese                        | 0.20            | 0.12                 | -0.02   0.43  |
| def: definite * spec: specific                     | -0.00           | 0.07                 | -0.15   0.14  |
| def: definite * mod: no modifier                   | -0.07           | 0.07                 | -0.20   0.06  |
| Ntype: singular * NL: German                       | 0.13            | 0.15                 | -0.16   0.41  |
| Ntype: mass * NL: German                           | -0.32           | 0.19                 | -0.69   0.06  |
| Ntype: singular * NL: Brazilian                    | 0.03            | 0.13                 | -0.22   0.29  |
| Ntype: mass * NL: Brazilian                        | 0.25            | 0.17                 | -0.09   0.59  |
| Ntype: singular * NL: Chinese                      | -0.08           | 0.13                 | -0.34   0.18  |
| Ntype: mass * NL: Chinese                          | -0.03           | 0.17                 | -0.36   0.30  |
| Ntype: singular * spec: specific                   | 0.19 *          | 0.09                 | 0.02   0.36   |
| Ntype: mass * spec: specific                       | -0.04           | 0.11                 | -0.24   0.17  |
| Ntype: singular * mod: no modifier                 | 0.02            | 0.08                 | -0.12   0.17  |
| Ntype: mass * mod: no modifier                     | 0.23 *          | 0.10                 | 0.03   0.44   |
| NL: German * Level                                 | 0.05            | 0.11                 | -0.16   0.26  |
| NL: Brazilian * Level                              | -0.03           | 0.09                 | -0.20   0.15  |
| NL: Chinese * Level                                | -0.22 *         | 0.10                 | -0.40   -0.03 |
| def: definite * Ntype: singular * NL: German       | -0.47 **        | 0.15                 | -0.76   -0.18 |
| def: definite * Ntype: mass * NL: German           | 0.33            | 0.18                 | -0.02   0.69  |
| def: definite * Ntype: singular * NL: Brazilian    | -0.10           | 0.13                 | -0.35   0.16  |
| def: definite * Ntype: mass * NL: Brazilian        | 0.23            | 0.16                 | -0.09   0.55  |
| def: definite * Ntype: singular * NL: Chinese      | -0.08           | 0.13                 | -0.34   0.19  |
| def: definite * Ntype: mass * NL: Chinese          | -0.22           | 0.16                 | -0.54   0.09  |
| def: definite * Ntype: singular * spec: specific   | -0.09           | 0.08                 | -0.25   0.07  |
| def: definite * Ntype: mass * spec: specific       | 0.29 **         | 0.10                 | 0.09   0.50   |
| def: definite * Ntype: singular * mod: no modifier | 0.33 ***        | 0.08                 | 0.18   0.48   |
| def: definite * Ntype: mass * mod: no modifier     | -0.13           | 0.10                 | -0.33   0.07  |
| <b>Random Effects by writing ID</b>                |                 |                      |               |
| SD (Intercept)                                     | 0.50            |                      |               |
| SD (def: definite)                                 | 0.53            |                      |               |
| SD (Ntype: singular)                               | 0.51            |                      |               |
| SD (Ntype: mass)                                   | 0.76            |                      |               |
| SD (abstr: concrete)                               | 0.42            |                      |               |
| SD (spec: specific)                                | 0.50            |                      |               |
| SD (def: definite * Ntype: singular)               | 0.60            |                      |               |

| <i>Predictors</i>  | <b>Accuracy Rate</b>                 |                   |               |
|--|--------------------------------------|-------------------|---------------|
|  | <i>Log-Odds</i>                      | <i>std. Error</i> | <i>95% CI</i> |
| SD (def: definite * Ntype: mass)   | 0.00                                 |                   |               |
| N <sub>wr_id</sub>   | 632                                  |                   |               |
| Observations   | 5772                                 |                   |               |
| Marginal R <sup>2</sup> / Conditional R <sup>2</sup>                           | 0.198 / 0.254                        |                   |               |
| Log-likelihood ratio test comparing to the null model<br>(only random effects) | $\chi^2 (43) = 334.76, p < 10^{-15}$ |                   |               |
| C-statistic  | 0.92 (strong predictive power)       |                   |               |
| VIFs   |                                      |                   |               |
| noun type * NL   | 5.6                                  |                   |               |
| definiteness * noun type * NL  | 5.9                                  |                   |               |
| other variables  | < 5                                  |                   |               |
| Overdispersion ratio   | 0.580 ( $\chi^2 = 3315.192, p = 1$ ) |                   |               |

*Note.* \*  $p < 0.05$  \*\*  $p < 0.01$  \*\*\*  $p < 0.001$ . The omitted levels of multi-level variables in the model

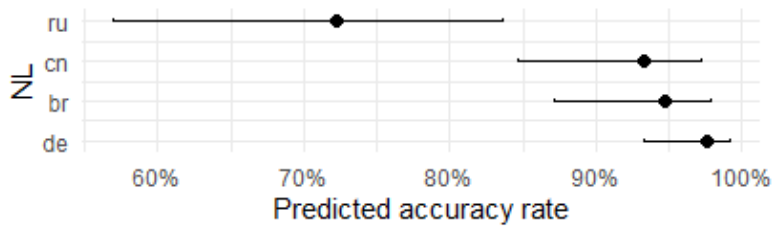
are “Russian” for NL, “plural” for noun type (Ntype), “object” for syntactic position (synt).

### ***The Effect of Discourse-Pragmatic Context on Definite Article Accuracy***

To investigate the effect of discourse-pragmatic contexts on accuracy rate in definite contexts, I fitted a separate generalised linear mixed-effects logistic regression model, including only those contexts where the target article is “the” ( $n = 2039$ ). Additionally, I excluded the “kind” discourse-pragmatic context, as instances of this constitute only 1.5% of the data ( $n = 30$ ). Thus, the following analysis is based on 2009 observations.

As in the Full Accuracy Model, I found a significant effect of NL, and a small effect of modifier presence on count singular nouns. The effect of discourse-pragmatic contexts is barely significant on plural nouns only. The full table of model results and performance metrics is provided in Table 8.

**NL.** As in the Full Accuracy Model, L1-Russians are most likely to make errors in definite contexts (72% predicted accuracy) and are significantly less accurate than any other NL at  $p < 0.001$  (Figure 8, note that the scale starts at 55%). In addition, L1-Chinese learners are also significantly less accurate than L1-Germans, although the difference is not as considerable (93% vs. 98%,  $p < 0.01$ ).

**Figure 8***The Effect of NL in Definite Contexts*

**Discourse-Pragmatic Context in Interaction with Noun Type.** The interaction between noun type and discourse-pragmatic context is limited to a barely significant difference ( $p = 0.012$ ) between anaphoric (84% predicted accuracy) and situational (95% predicted accuracy) contexts for count plural nouns (examples 40 and 41, respectively).

(40) There were a few problems with the house with I signed the contract with the landlord.

[...] Unfortunately none of *[the] problems* has been fixed after I talked to the landlord.

(L1-Chinese, B1, ID 295314)

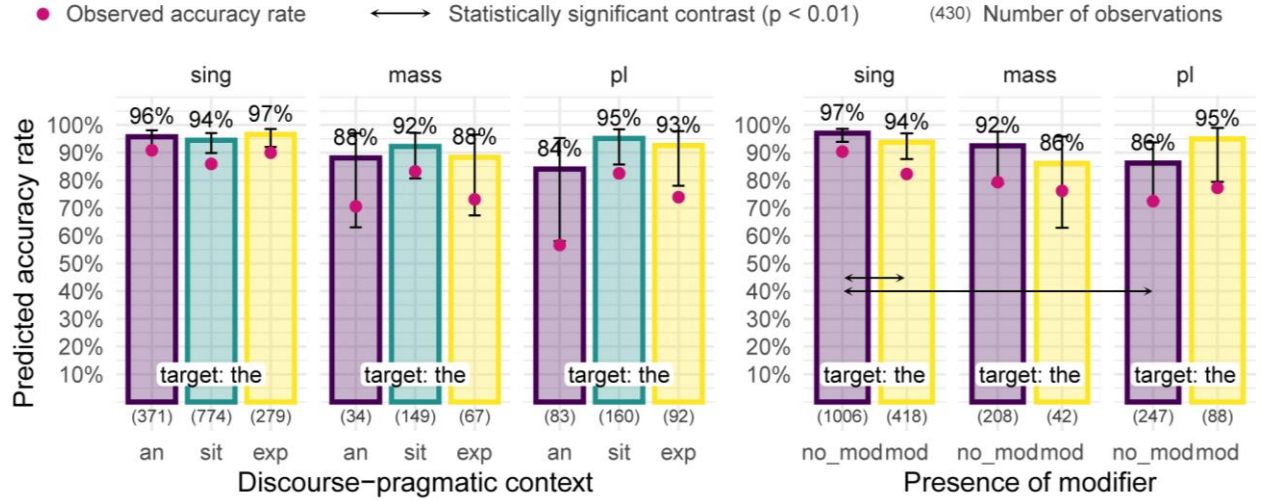
(41) I have always enjoyed having dinner at your restaurant until this time... First, there was an insect in my soup! Second, *[the] grilled lamb chops* with fresh vegetables were too salty and spicy.

(L1-Russian, B1, ID 157548)

Figure 9 (left) shows the predicted accuracy rates and 95% confidence intervals for the other noun types and contexts. Note the relatively large confidence intervals in mass and plural nouns indicating considerable uncertainty about model estimates for these categories.

**Figure 9**

*Interaction between Discourse-Pragmatic Context and Noun Type (Left), and Modifier and Noun Type (Right)*



**Modifier in Interaction with Noun Type.** The interaction between noun type and modifier is driven by two contrasts (Figure 9, right). Firstly, when there is no modifier, learners are significantly more likely to make errors with count plural (40) than with count singular nouns. Secondly, the presence of a pre-nominal modifier has a negative effect on definite article accuracy in count singular context only (42).

(42) Thank you for giving me the opportunity to apply for this job at *[the] sports centre*.

(L1-Brazilian, B2, ID 513902)

**Table 8**

*Definite Accuracy Model Results and Performance Metrics*

| Predictors                         | Log-Odds | Accuracy Rate |              |
|------------------------------------|----------|---------------|--------------|
|                                    |          | std. Error    | 95% CI       |
| Intercept (grand mean)             | 2.55 *** | 0.42          | 1.73   3.36  |
| Ntype: singular                    | 0.54 *   | 0.24          | 0.08   1.01  |
| Ntype: mass                        | -0.38    | 0.28          | -0.93   0.17 |
| mod: no modifier                   | 0.06     | 0.16          | -0.26   0.37 |
| discourse: anaphoric               | -0.29    | 0.19          | -0.66   0.09 |
| discourse: situational             | 0.21     | 0.16          | -0.10   0.52 |
| NL: German                         | 1.17 *** | 0.24          | 0.69   1.64  |
| NL: Brazilian                      | 0.34     | 0.20          | -0.06   0.74 |
| NL: Chinese                        | 0.09     | 0.21          | -0.32   0.50 |
| spec: specific                     | 0.09     | 0.14          | -0.19   0.38 |
| Ntype: singular * mod: no modifier | 0.33 *   | 0.17          | 0.00   0.65  |

| <i>Predictors</i>  | <b>Accuracy Rate</b>                |                   |               |
|--|-------------------------------------|-------------------|---------------|
|  | <i>Log-Odds</i>                     | <i>std. Error</i> | <i>95% CI</i> |
| Ntype: mass * mod: no modifier   | 0.28                                | 0.27              | -0.25   0.82  |
| Ntype: singular * discourse: anaphoric   | 0.31                                | 0.22              | -0.12   0.74  |
| Ntype: mass * discourse: anaphoric   | 0.13                                | 0.31              | -0.49   0.74  |
| Ntype: singular * discourse: situational                                       | -0.47 *                             | 0.18              | -0.82   -0.11 |
| Ntype: mass * discourse: situational   | 0.09                                | 0.25              | -0.40   0.58  |
| <b>Random Effects</b>  |                                     |                   |               |
| SD (Intercept)   | 0.38                                |                   |               |
| SD (Ntype: singular)   | 0.58                                |                   |               |
| SD (Ntype: mass)   | 1.30                                |                   |               |
| SD(mod: no modifier)   | 1.18                                |                   |               |
| SD (spec: specific)  | 1.34                                |                   |               |
| SD (Observations)  | 1.00                                |                   |               |
| N <sub>wr_id</sub>   | 506                                 |                   |               |
| Observations   | 2009                                |                   |               |
| Marginal R <sup>2</sup> / Conditional R <sup>2</sup>                           | 0.290 / 0.319                       |                   |               |
| Log-likelihood ratio test comparing to the null model<br>(only random effects) | $\chi^2 (15) = 35.584, p = 0.002$   |                   |               |
| C-statistic  | 0.97                                |                   |               |
| VIFs   | < 2                                 |                   |               |
| Overdispersion ratio   | 0.385 ( $\chi^2 = 765.462, p = 1$ ) |                   |               |

Note. \*  $p < 0.05$  \*\*  $p < 0.01$  \*\*\*  $p < 0.001$ . The omitted levels of multi-level variables in this model

are “Russian” for NL, “plural” for noun type (Ntype), “explanatory” for discourse-pragmatic context.

### **Predictors of Error Type**

**Error Types in Count Singular Indefinites.** In line with the Full Accuracy Model, the Error Type Model for count singular indefinites shows significant effects of NL, modifier, specificity, syntactic position, abstractness<sup>15</sup>, and proficiency level. Overall, the model predicts 85.5% probability for correct supplings of “a”, 11% omission, 3.5% substitution. The full list of model estimates and performance metrics is given in Table 9.

**NL and Proficiency Level.** In line with the Full Accuracy Model, this model predicts all the NLs to be significantly less accurate than L1-Germans (Figure 10, top; note that rates are given for the average proficiency level, i.e. EF level 8).

<sup>15</sup> The effect of abstractness is marginally significant on predicted accuracy rates (84% for concrete, 87% for abstract nouns) at  $p = 0.06$  and, thus, not discussed in more detail here.



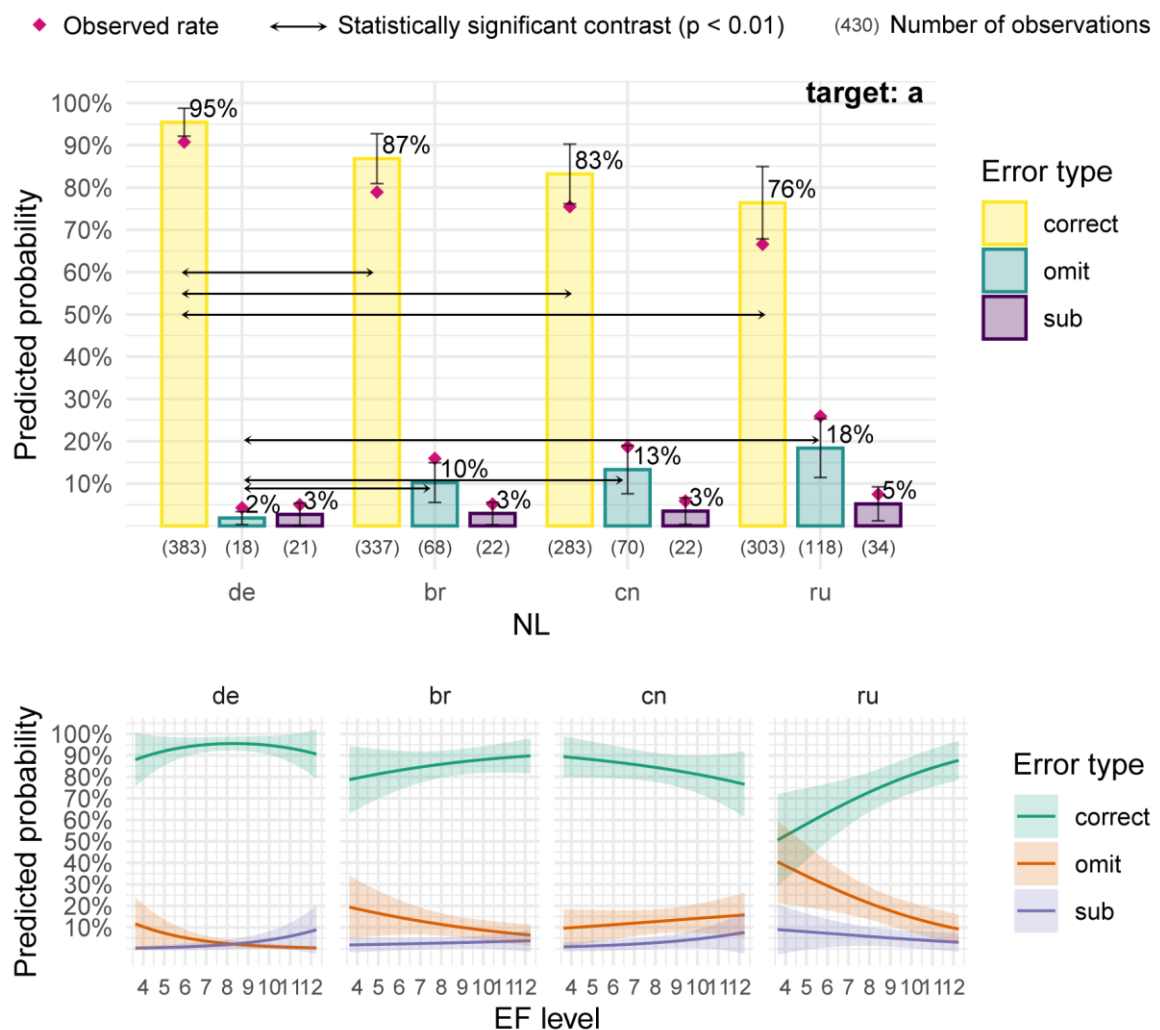
The differences in predicted omission rates between groups mirror the differences in accuracy rates precisely, i.e. all groups are significantly more likely to omit “a” than L1-Germans. Substitution error rates are similar across NLs.

As in the Full Accuracy Model, NL also interacts with level (Figure 10, bottom). The significant accuracy improvement of L1-Russians is achieved through a significant decrease in the omission rate ( $p < 0.01$ ). L1-Germans also show a significant decrease in the omission rate ( $p < 0.01$ ).

When development slopes are compared across NLs, none of the differences reach statistical significance at  $p < 0.01$ .

**Figure 10**

*The Effect of NL Alone (Top) and in Interaction with Level (Bottom) on Predicted Probabilities of Error Types in Count Singular Indefinites*



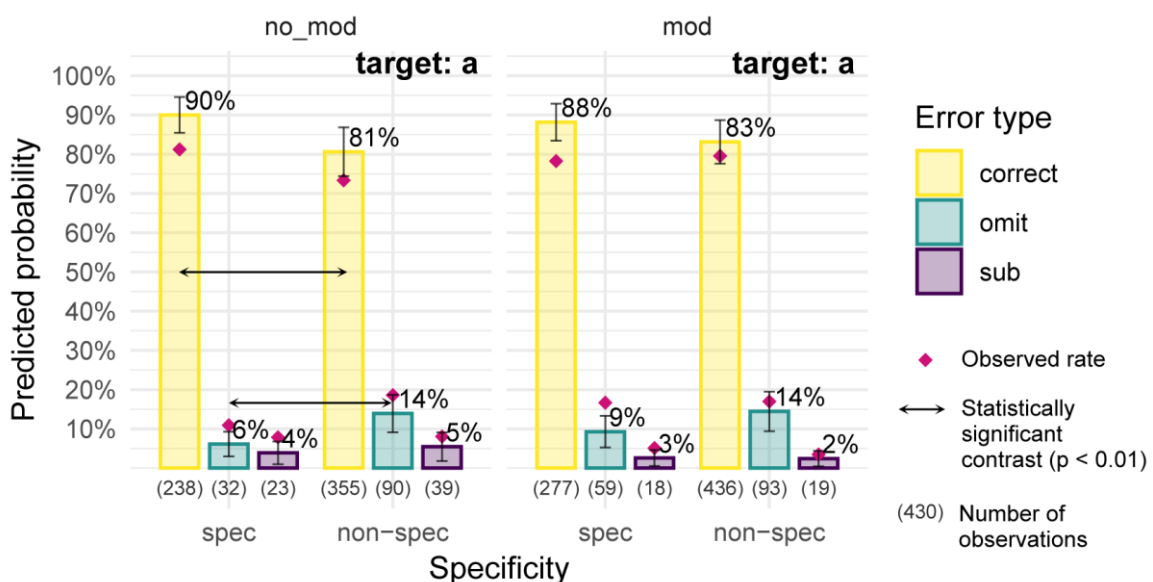
**Interaction between Specificity and Modifier.** There is a significant effect of specificity, as defined in this study, in non-premodified nouns, with the odds of omitting “a” increasing to 14% (43), as opposed to 6% for specific referents (Figure 11, left). The pattern is similar in premodified nouns but does not reach significance (Figure 11, right).

(43) I have many dreams [...] I'd make [a] career in my business and have a fulfilled and balanced live.

(L1-German, B1, ID 249369)

**Figure 11**

*The Effect of Specificity in Interaction with Modifier on Predicted Probabilities of Error Types in Count Singular Indefinites*



**Syntactic Position.** The predicted accuracy rates for subjects and objects are significantly lower than those for existentials and predicates ( $p < 0.01$  for object vs. existential and predicate,  $p < 0.05$  for subject vs. existential and predicate, Figure 12), confirming the Full Accuracy Model.

Omitting “a” is more likely before objects (44) than before existentials ( $p = 0.001$ ) or predicates ( $p = 0.03$ ) and more likely before subjects (45) than before existentials ( $p = 0.03$ ).

(44) Then I will take [an] air plane to Madrid.

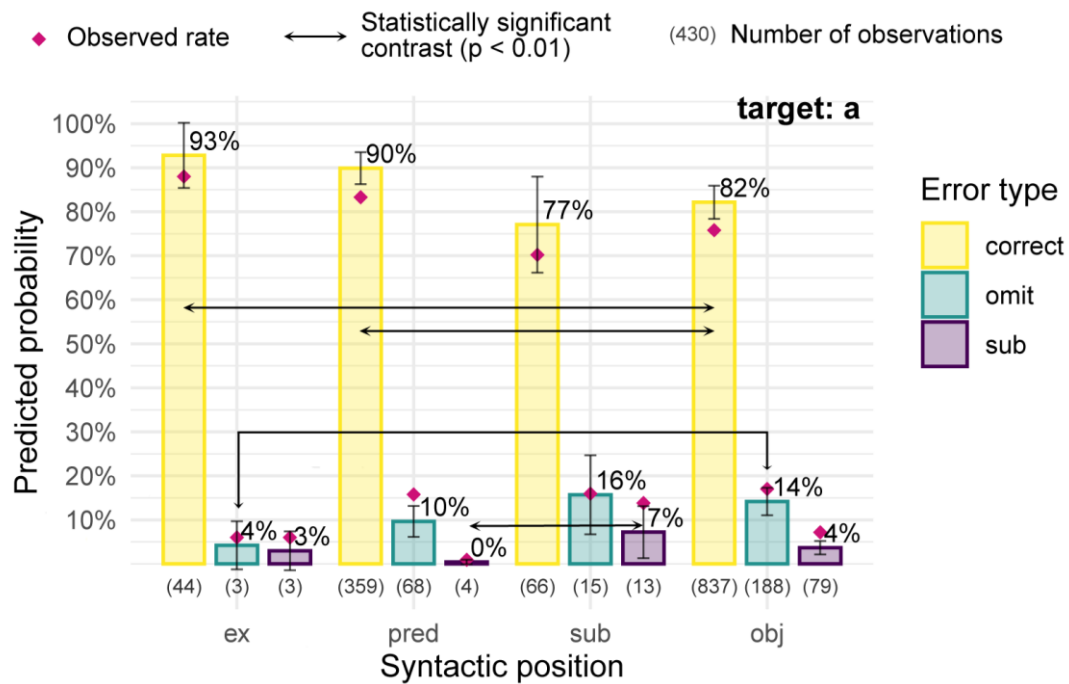
(L1-Chinese, B1, ID 3441)

(45)[An] Online study program give me opportunaty to learn when I have free time and desire.

(L1-Russian, B1, ID 15851)

**Figure 12**

*The Effect of Syntactic Position on Predicted Probabilities of Error Types in Count Singular Indefinites*



Moreover, substitution errors are more probable in subjects (46) and objects (47) than in predicates ( $p < 0.001$  and  $p = 0.025$ , respectively).

(46)~~The~~ [A] large and quickly developed consulting company "ABM Partner" is seekeng for an experienced Financial Analyst.

(L1-Russian, B1, ID 288445)

(47)[...] last friday in my city two thieves stole ~~the~~ [a] motorcycle from ~~the~~ [a] man.

(L1-Brazilian, A2, ID 817284)

**Table 9**

*Results of the Error Type Model for Indefinite Count Singular Nouns*

| Predictors                    | Log-Odds  | Accuracy Rate |               |
|-------------------------------|-----------|---------------|---------------|
|                               |           | std. Error    | 95% CI        |
| omit ~ Intercept (grand mean) | -2.53 *** | 0.24          | -3.00   -2.06 |
| sub ~ Intercept (grand mean)  | -3.68 *** | 0.32          | -4.31   -3.06 |

|  |                                |      |               |
|--|--------------------------------|------|---------------|
| omit ~ NL: German  | -1.66 ***                      | 0.33 | -2.31   -1.01 |
| sub ~ NL: German   | -0.40                          | 0.40 | -1.18   0.39  |
| omit ~ NL: Brazilian   | 0.19                           | 0.23 | -0.25   0.64  |
| sub ~ NL: Brazilian  | -0.20                          | 0.38 | -0.95   0.55  |
| omit ~ NL: Chinese   | 0.51 *                         | 0.23 | 0.07   0.96   |
| sub ~ NL: Chinese  | 0.04                           | 0.39 | -0.72   0.79  |
| omit ~ Level   | -0.44 **                       | 0.14 | -0.70   -0.17 |
| sub ~ Level  | 0.31                           | 0.22 | -0.12   0.75  |
| omit ~ spec: specific  | -0.39 ***                      | 0.11 | -0.60   -0.17 |
| sub ~ spec: specific   | -0.13                          | 0.17 | -0.47   0.20  |
| omit ~ mod: no modifier  | -0.11                          | 0.10 | -0.30   0.07  |
| sub ~ mod: no modifier   | 0.31 *                         | 0.14 | 0.03   0.60   |
| omit ~ abstr: concrete   | 0.12                           | 0.09 | -0.06   0.30  |
| sub ~ abstr: concrete  | 0.29 *                         | 0.15 | 0.00   0.59   |
| omit ~ synt: existential   | -0.97                          | 0.54 | -2.03   0.09  |
| sub ~ synt: existential  | 0.09                           | 0.62 | -1.12   1.30  |
| omit ~ synt: predicate   | -0.08                          | 0.26 | -0.59   0.42  |
| sub ~ synt: predicate  | -1.76 ***                      | 0.49 | -2.72   -0.80 |
| omit ~ synt: subject   | 0.62                           | 0.33 | -0.04   1.27  |
| sub ~ synt: subject  | 1.21 **                        | 0.42 | 0.39   2.03   |
| omit ~ NL: German * Level  | -0.51                          | 0.29 | -1.08   0.07  |
| sub ~ NL: German * Level   | 0.62                           | 0.42 | -0.20   1.44  |
| omit ~ NL: Brazilian * Level   | 0.06                           | 0.21 | -0.35   0.48  |
| sub ~ NL: Brazilian * Level  | -0.15                          | 0.36 | -0.86   0.57  |
| omit ~ NL: Chinese * Level   | 0.64 **                        | 0.21 | 0.22   1.05   |
| sub ~ NL: Chinese * Level  | 0.35                           | 0.40 | -0.43   1.13  |
| omit ~ spec: specific * mod: no modifier                                       | -0.12                          | 0.09 | -0.30   0.07  |
| sub ~ spec: specific * mod: no modifier  | -0.12                          | 0.14 | -0.40   0.16  |
| <b>Random Effects</b>  |                                |      |               |
| SD omit ~ (Intercept)  | 0.53                           |      |               |
| SD sub ~ (Intercept)   | 0.58                           |      |               |
| N <sub>wr_id</sub>   | 541                            |      |               |
| Observations   | 1679                           |      |               |
| Pseudo-R <sup>2</sup> (Nagelkerke, 1991)                                       | 0.165                          |      |               |
| Log-likelihood ratio test comparing to the null model<br>(only random effects) | $D(28) = 259.49, p < 10^{-15}$ |      |               |
| C-statistic  | 0.86                           |      |               |
| VIFs   |                                |      |               |
| NL   | 13.8 <sup>a</sup>              |      |               |
| other variables  | < 3                            |      |               |

Note. \*  $p < 0.05$  \*\*  $p < 0.01$  \*\*\*  $p < 0.001$ ; omit ~ estimates for omission errors vs. correct; sub ~

estimates for substitution errors vs. correct. The omitted levels of multi-level variables in this model

are “Russian” for NL, “object” for syntactic position (synt).

<sup>a</sup> The VIF for NL is very high, which is partly explained by the fact it is involved in an interaction.

However, in a model without interactions, the VIF for NL is also rather high (8.4). Nevertheless,

removing this variable did not significantly alter the estimates or  $p$ -values for any other variables. In

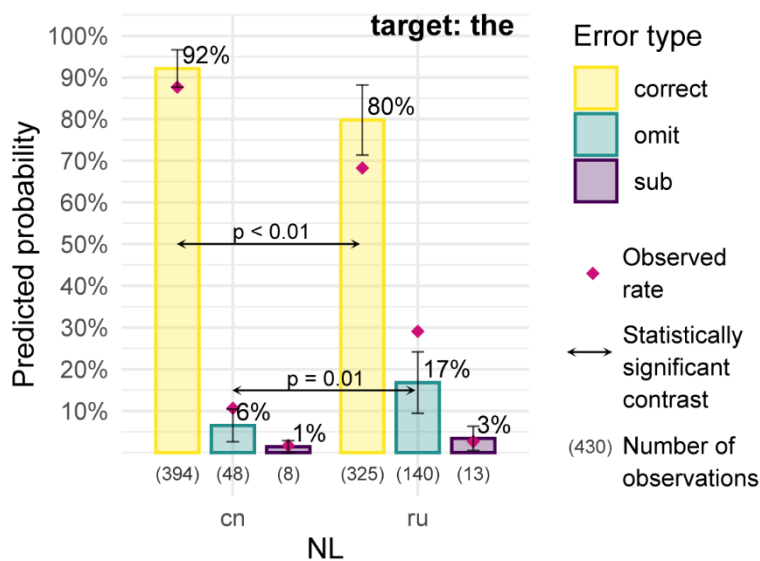
addition, the estimates were stable throughout the model selection process. Since NL is a crucial factor with a significant effect, I decided there were not enough grounds to remove it from the model.

**Error Types in Count Singular Definites in L1-Russian and L1-Chinese learners<sup>16</sup>.** The Error Type Model for count singular definites<sup>17</sup> confirms the results of the Full Accuracy Model and the Definite Accuracy Model, showing significant effects of NL and modifier presence but no significant effect of discourse-pragmatic context on accuracy rate. Overall, the model predicts 86% probability for correct suppliance of “the”, 11.6% omission, 2.4% substitution. However, the level of individual variation (SDs for the random intercepts by writing ID) is considerably higher than what I found in previous models.

The full list of model estimates and performance metrics is provided in Table 10.

**Figure 13**

*The Effect of NL on Predicted Probabilities of Error Types in Singular Definites*



<sup>16</sup> Due to convergence issues most likely resulting from the uneven distribution of errors across noun types and NLs, this model was fitted on L1-Russian and L1-Chinese learners only ( $n = 682$ ).

<sup>17</sup> I excluded the syntactic position variable because its effect in the Full Accuracy Model was driven by the difference between subjects/objects vs. predicates/existentials, and the latter are hardly represented in definite contexts.

**NL.** The effect of NL is driven by omission (not substitution) (Figure 13). L1-Russians are significantly more likely to omit “the” ( $p = 0.01$ ) than L1-Chinese learners and are significantly less accurate ( $p < 0.01$ ), confirming the Full Accuracy Model.

**Modifier.** The effect of modifier presence is also driven by omission (not substitution) (Figure 14, right), with omission significantly increased in premodified nominals (48), which is also reflected in a significantly lower predicted accuracy rate, confirming the Full Accuracy Model (although the effect here is slightly larger).

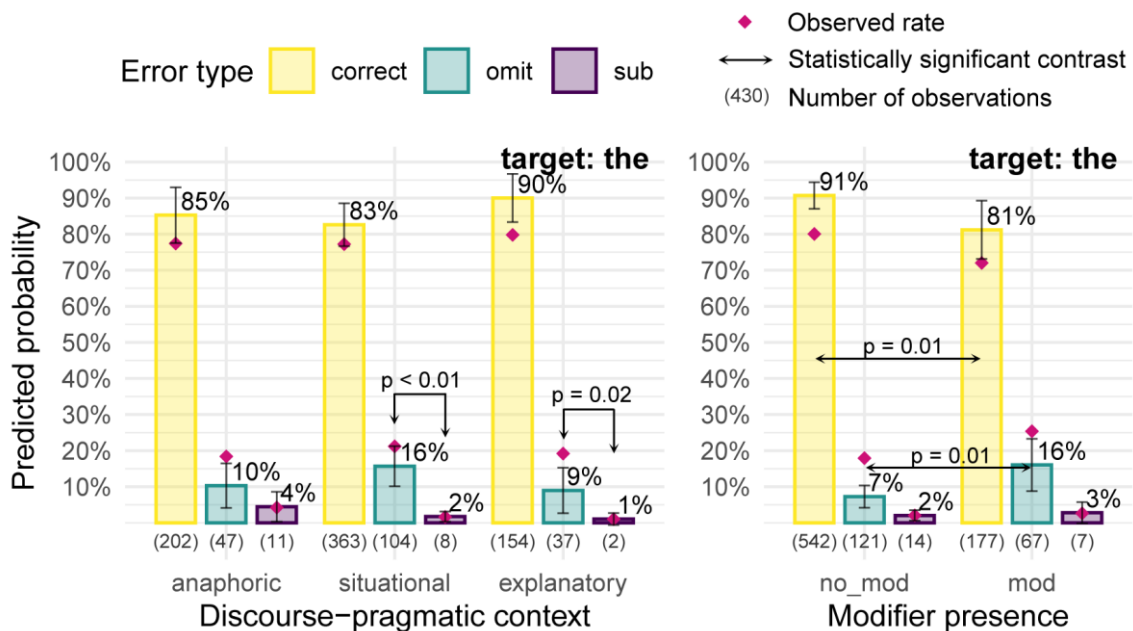
(48) I first met my friend, Kolya, when I was working in advertising project five years ago. [...]

Kostya and I enjoy working on [the] advertising project together.

(L1-Russian, A2, ID 887991)

**Figure 14**

*The Effects of Discourse-Pragmatic Context (Left) and Modifier Presence (Right) on Predicted Probabilities of Error Types in Singular Definites for L1-Russian and L1-Chinese Learners*



**Discourse-Pragmatic Context.** Although discourse-pragmatic context<sup>18</sup> does not significantly affect accuracy rate, confirming the Definite Accuracy Model, omission errors are more likely than

<sup>18</sup> I excluded the “kind” context because there were only 19 instances.

substitution errors in situational ( $p < 0.0001$ ) and explanatory ( $p = 0.016$ ) but not in anaphoric contexts (Figure 14, left).

**Table 10**

*Results of the Error Type Model for Definite Count Singular Nouns (L1-Chinese and L1-Russian only)*

| <i>Predictors</i>   | <i>Log-Odds</i>                     | <b>Accuracy Rate</b> |               |
|---|-------------------------------------|----------------------|---------------|
|   |                                     | <i>std. Error</i>    | <i>95% CI</i> |
| omit ~ Intercept (grand mean)   | -2.81 ***                           | 0.34                 | -3.47   -2.15 |
| sub ~ Intercept (grand mean)  | -4.42 ***                           | 0.57                 | -5.54   -3.29 |
| omit ~ NL: Russian  | 1.13 **                             | 0.42                 | 0.32   1.95   |
| sub ~ NL: Russian   | 1.06                                | 0.66                 | -0.24   2.37  |
| omit ~ discourse: anaphoric   | -0.09                               | 0.24                 | -0.56   0.38  |
| sub ~ discourse: anaphoric  | 0.81 *                              | 0.40                 | 0.03   1.60   |
| omit ~ discourse: situational   | 0.39                                | 0.21                 | -0.01   0.79  |
| sub ~ discourse: situational  | -0.10                               | 0.42                 | -0.92   0.71  |
| omit ~ mod: no modifier   | -0.47 **                            | 0.15                 | -0.78   -0.17 |
| sub ~ mod: no modifier  | -0.23                               | 0.32                 | -0.85   0.39  |
| omit ~ abstr: concrete  | 0.27                                | 0.16                 | -0.04   0.58  |
| sub ~ abstr: concrete   | -0.13                               | 0.28                 | -0.68   0.43  |
| <b>Random Effects</b>   |                                     |                      |               |
| SD omit ~ (Intercept)   | 1.1                                 |                      |               |
| SD sub ~ (Intercept)  | 0.75                                |                      |               |
| N <sub>wr_id</sub>  | 227                                 |                      |               |
| Observations  | 682                                 |                      |               |
| Pseudo-R <sup>2</sup> (Nagelkerke, 1991)                                    | 0.112                               |                      |               |
| Log-likelihood ratio test comparing to the null model (only random effects) | $D(10) = 51.939, p = 1.2 * 10^{-7}$ |                      |               |
| C-statistic <sup>a</sup>  | 0.89                                |                      |               |
| VIFs <sup>b</sup>   | < 4                                 |                      |               |

*Note.* \*  $p < 0.05$  \*\*  $p < 0.01$  \*\*\*  $p < 0.001$ ; omit ~ estimates for omission errors vs. correct; sub ~

estimates for substitution errors vs. correct. The omitted levels of multi-level variables in this model are “Russian” for NL, “explanatory” for discourse-pragmatic context.

**Error Types in Mass Indefinites<sup>19</sup>.** The model predicts 85% probability for correct  $\emptyset$  contexts, 9% overuse of “a”, 6% overuse of “the”. This is slightly lower than the predicted accuracy rate from the Full Accuracy Model, which was above 90%. The full list of model estimates and performance metrics is given in Table 11.

<sup>19</sup> Fitting the model on both mass and plural indefinites produced negative pseudo-R<sup>2</sup> values (Nagelkerke, 1991) pointing to its potential inadequacy. Since there were fewer errors in the plural context, I fitted the model on mass nouns only ( $n = 888$ ).

There are significant main effects of specificity and modifier presence on the type of error (Figure 15), whereas the Full Accuracy Model only found an effect of modifier presence.

First, learners are significantly more likely to overuse “a” with specific (49) referents, i.e. those which refer to a certain existing entity, than with non-specific (50) ones ( $p = 0.015$ ), as seen in Figure 15 (left). However, the difference between predicted accuracy rates is not significant, which is why the effect was not significant for mass nominals in the Full Accuracy Model.

(49)When police got the home they noticed that one servant's face was covered with ~~a~~ [Ø] *red paint*.

(L1-Russian, B2, ID 854608)

(50)In addition, both my brother and I had to read ~~the~~ [Ø] *classical literature* for our school.

(L1-Russian, B2, ID 855790)

Second, learners are significantly more likely to overuse “a” if a noun is premodified (51), which is also predicted to significantly affect their accuracy rate in such instances ( $p < 0.001$  for both comparisons, Figure 15, right).

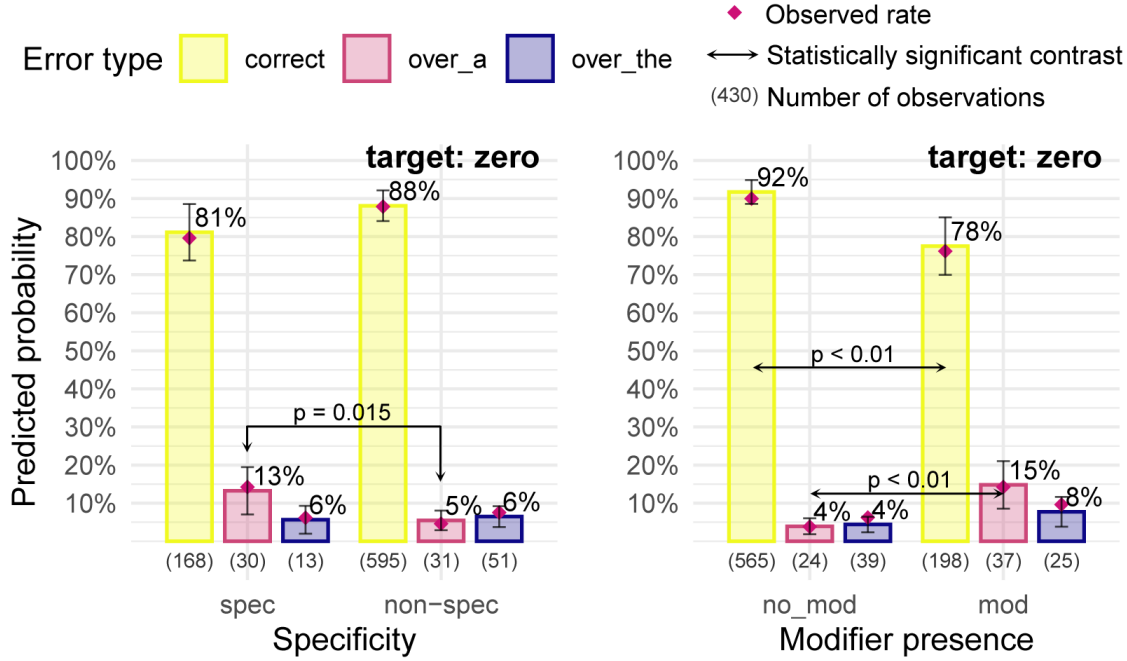
(51)So I become fit, get fresh air and see ~~a~~ [Ø] *beautiful nature*.

(L1-German, B2, ID 1087916)



**Figure 15**

*The Effects of Specificity (Left) and Modifier Presence (Right) on Predicted Probabilities of Error Types in Mass Indefinites*

**Table 11**

*Results of the Error Type Model for Indefinite Mass Nouns*

| Predictors  | Log-Odds                                  | Accuracy Rate |               |
|---|---|---------------|---------------|
|   |   | std. Error    | 95% CI        |
| over_a ~ Intercept (grand mean)   | -2.51 ***                                 | 0.20          | -2.90   -2.11 |
| over_the ~ Intercept (grand mean)   | -2.68 ***                                 | 0.22          | -3.11   -2.24 |
| over_a ~ spec: specific   | 0.50 **                                   | 0.18          | 0.15   0.85   |
| over_the ~ spec: specific   | -0.02                                     | 0.20          | -0.41   0.38  |
| over_a ~ mod: no modifier   | -0.76 ***                                 | 0.16          | -1.08   -0.44 |
| over_the ~ mod: no modifier   | -0.37 *                                   | 0.16          | -0.69   -0.04 |
| <b>Random Effects</b>   |   |               |               |
| SD over_a ~ (Intercept)   | 0.53                                      |               |               |
| SD over_the ~ (Intercept)   | 0.42                                      |               |               |
| N <sub>wr_id</sub>  | 373                                       |               |               |
| Observations  | 888                                       |               |               |
| Pseudo-R <sup>2</sup> (Nagelkerke, 1991)                                    | 0.167                                     |               |               |
| Log-likelihood ratio test comparing to the null model (only random effects) | $D(4) = 57.151, p = 1.15 \times 10^{-11}$ |               |               |
| C-statistic <sup>a</sup>  | 0.86                                      |               |               |
| VIFs <sup>b</sup>   | < 1.5                                     |               |               |

Note. \*  $p < 0.05$  \*\*  $p < 0.01$  \*\*\*  $p < 0.001$ ; over\_a ~ estimates for overuse of “a” errors vs. correct;

over\_the ~ estimates for overuse of “the” errors vs. correct.

## Discussion

### *Summary of Findings*

Using manually coded learner corpus data and statistical modelling, I have broadly confirmed previous findings showing higher article accuracy in [+art] than in [-art] L1 learners (Ionin et al., 2008; Murakami & Alexopoulou, 2016a; Snape, 2008). In this data, L1-Russians improve from about 50% TLU at level A2 to 75% TLU at level B2. By contrast, L1-Germans start almost at ceiling, around 90%, and stay at the same level throughout. In-between, L1-Brazilian and L1-Chinese learners fluctuate between 70% and 80%.

Rather surprisingly, [-art] L1-Chinese learners pattern with [+art] L1-Brazilians and not with [-art] L1-Russians. There is evidence that Chinese is in the process of grammaticalisation of certain morphemes, which function similarly to English articles (Chen, 2004; Huang, 1999).

As expected (Lardi re, 2004), I have found that all learners are overall more accurate in using “the” than “a”, except L1-Germans, who are at ceiling for both. However, it is revealed that a third of all nominals used by learners require  $\emptyset$ . This, taken together with the finding that omission is the most common learner error, leads to 90% accuracy in  $\emptyset$  contexts across all NLs, contrasting e.g. with 60-70% predicted accuracy of L1-Russians for mass and plural definites.

One might argue that the high omission rate in mass and plural definites might be accounted for by the much lower frequency of these contexts in learner production; in Figure 5 mass and plural definites sum up to roughly a third of mass and plural indefinites across NLs (262/878 and 341/845, respectively), suggesting that there may be less opportunity to learn the definite or there might be overgeneralisation from the indefinite. However, there is the opposite pattern in indefinite singulars (Figure 1), which are more often non-specific (1032 non-specific vs. 647 specific), but learners’ accuracy is still significantly lower with non-specific than with specific referents. Thus, frequency does not necessarily lead to higher accuracy. In the case of mass and plural nominals though, the more frequent pattern of mass and plural indefinites is also the pattern with  $\emptyset$  contexts, which might be misanalysed by learners and overextended to definites.

The main question was how the different factors that have previously been shown to affect L2 English article use work together in predicting learner accuracy and errors. I have found that, apart from the well-established large L1 effect, L2 English article use is significantly affected by specificity, prenominal modifier presence, and syntactic position, but not by abstractness or discourse-pragmatic context.

### ***Specificity and Prenominal Modifier***

**Definites.** The results of the corpus study show that definites are not affected by specificity, as defined in this thesis, but there is increased “the” omission with premodified count singulars, i.e. “the” is more likely omitted in “the advertising company” than in “the company”. Trenkic (2007), based on similar findings from L1-Serbians, suggests the article is structurally an adjective for learners, making it optional. So, when a modifier has already sufficiently narrowed down the range of potential referents, an article may be redundant. In the data reported above, in many cases the modifier leaves only one plausible referent option, e.g. “the departure lounge of Oslo airport”, “the following recipe”, “the top score”. One could adopt Trenkic’s cognitive explanation, which suggests increased omission of redundant elements when cognitive resources are limited. However, it still needs to be explained why in this data the redundancy effect is only found in definites, but not in indefinites, which I will address in the following subsection.

Two findings remain unclear. First, definite mass and plural contexts are unaffected by modifier presence, i.e. “the” omission is not increased before premodified mass and plural nominals, e.g. “the red wine”/“the new shoes”. Second, L1-Russians have considerably lower accuracy in mass and plural contexts (predicted 57% and 69%) than in count singulars (86%). Austin et al. (2015) also observe a higher “the” omission rate with plurals than with singulars in 20 intermediate L1-Thai learners completing a prompted story recall task. They attribute this to L1-L2 structural competition, which predicts that cognitively more demanding contexts, such as those requiring the suppliance of multiple functional morphemes (e.g. “the” and plural “-s” in plural definites), impede the suppression of competing L1 forms (i.e. bare plural definites). However, this only explains the higher

omission in definite plurals but not in mass nominals. I cannot fully explain these patterns, which might also be rather uncertain due to lower numbers in these contexts (256 mass, 335 plural) and larger standard errors.

**Count Singular Indefinites.** The first question is how and why count singulars (target “a”) are significantly impacted by specificity. Essentially, “a” is more consistently supplied with specific referents, which refer to certain existing entities, but more often omitted with non-specific ones. I claim that learners may associate “a” with the function of introducing a certain existing referent (i.e. specific, by the definition adopted here) into the discourse. By contrast, in non-specific contexts, where “a” is not introducing an existing referent (as there is none), the semantic contribution of “a” may be unclear to learners.

The second question is why there is no modifier effect in count singular indefinites, i.e. why the modifier does not appear to make “a” redundant in the same fashion as it can make “the” redundant. If one assumes, as suggested above, that learners associate “a” with the function of introducing a specific (existing) referent, one has to admit that a modifier cannot fulfil this function. There is also no evidence that learners use “a” to signal referent identifiability, which is the function of “the”, as there are few substitution errors. Therefore, although a modifier can narrow the range of possible referents, it may still only indicate a type, e.g. “We are seeking an experienced analyst” as opposed to “any analyst”, if it is accepted that learners do not consider “analyst” identifiable in the first place.

This is unlike the findings in Trenkic (2007), whose [-art] L1-Serbian participants tended to omit both “the” and “a” with premodified nominals. The discrepancy is partly explained by the different task types. Trenkic used an oral information gap task (map completion), and a written task asking participants to translate as many stories as they could within the time limit, ensuring less reliance on metalinguistic knowledge. These on-line tasks revealed higher omission rates than the tasks in the EFCAMDAT corpus used here, which were untimed and unsupervised. Nevertheless, in Trenkic’s written task, the modifier effect was overall more pronounced in definite than in indefinite

contexts, which is in the same direction with the pattern reported here of a significant (but smaller) effect in definites and no significant effect in indefinites.

These results also clearly differ from previous studies replicating Ionin et al. (2004), which detect the effect of specificity in learners with [-art] but not with [+art] L1s (García Mayo, 2009; R. Hawkins et al., 2006; Reid et al., 2006), whereas all of the learners in this study demonstrated lower accuracy with non-specific indefinites. Furthermore, in Ionin et al. and their replications, the effect of specificity manifested itself in the form of substitution errors (“the” instead of “a” for indefinites and vice versa for definites), although in Ionin et al. (2008) the effect in L1-Russians was limited to indefinites.

Apart from the apparent differences in defining specificity (see the subsection on “Specificity” in Chapter 2), this discrepancy is probably due to the data used in the different studies, elicitation in Ionin et al. vs. task-based production in EFL context in this study. In fact, in Ionin et al.’s (2004) written production data, the numbers are closer to the findings here (50% overall omission of “a”, 10% substitution of “a” with “the”<sup>20</sup>). Their analysis, however, focused on substitution errors, while I considered omission errors in this study.

**Mass Indefinites.** To understand the effect of specificity on indefinites, one also needs to look at the effect of specificity on indefinite mass nouns, where, unlike with indefinite singulars, specificity, i.e. reference to certain existing entities, leads to overusing “a” and, thus, to lower accuracy (Figure 15). Importantly, overusing “a” is also more likely when a noun is prenominally modified. I argue that this is consistent with the explanation for count singular indefinites above. If learners use “a” to introduce a certain existing referent, they would not use “a” with most mass nouns, which typically denote unbounded or vaguely defined entities. However, when a mass noun is used to refer to something specific, it will often refer to a portion or an instance of the entity, and learners might be using “a” to indicate this (52-53).

---

<sup>20</sup> Ionin et al. observed a higher substitution rate with specific (28%) than with non-specific nominals (4%). I find a similar pattern in this data (11% vs. 5% in L1-Russians). However, I show this contrast is not statistically significant in this much larger sample.

(52) When we seat the server brought us a [Ø] *corn soup* to start.

(L1-Brazilian, B1, ID 1082859)

(53) But one day,i did a [Ø] *great help* to her.

(L1-Chinese, B1, ID 1038505)

A prenominal modifier can additionally specify a subclass or a type of entity, which is arguably more likely to occur when a specific portion or instance is referred to<sup>21</sup>. This might mean that a modifier with a mass noun may be associated with a count interpretation. Consider cases where mass nouns can be interpreted as countable (e.g. “have a beer”, “how many sugars do you take?” etc.). In standard English, this phenomenon is limited, making such examples as “take out two rubbishes” or “get many feedbacks” unacceptable, albeit conceivable. This is not at all obvious to learners (54-55), especially when the equivalent of an English mass noun is count in their L1. This might explain at least some instances of the overuse of “a” with premodified specific referents.

(54)[...] when you do your *works*, you should share your *informations* [...]

(L1-Chinese, B1, ID 17199)

(55)They were very popular in *countrysides*.

(L1-Russian, B2, ID 931737)

### ***Syntactic Position***

The higher article accuracy in existentials and predicates as opposed to subject and object positions is broadly in line with the literature (Hua & Lee, 2005 only for non-specific contexts). One possible explanation is that the discourse and semantic properties, including specificity, of existential and predicate constructions are almost fixed regardless of the noun inserted: existentials state existence and are usually specific, predicates denote properties and are, thus, mostly non-specific (which may account for the virtual non-existence of substitution errors in predicates). They are also explicitly taught early on and may first be learned as formulaic sequences: existentials are often used

---

<sup>21</sup> In fact, over 40% of specific mass indefinites have prenominal modifiers in the data reported here, whereas of non-specific mass indefinites, only ca. 25% are premodified.

for simple descriptions, e.g. “There is a book on the table”, while predicates are used for introductions, e.g. “I am a student”.

Nouns in object positions, based on our data, are almost equally likely to be specific/non-specific and are roughly equally likely to be used with “the”/“a”/∅. Thus, there is no clear pattern learners can infer from the distribution.

As for nominals in subject positions, this data shows these are more likely to be definite and specific, with 70% occurring in target “the” contexts. In English, subjects are also usually topics/themes. I have already shown how a prenominal modifier, which helps to identify the referent, is associated with increased omission of “the”. Similarly, nominals in subject positions may be considered (by learners) sufficiently identified, which would result in article omission (see Trenkic & Pongpaioj (2013), who observed increased article omission before salient nouns in subject position). Indefinite subject nominals, however, are more often non-specific (64% of indefinite singulars in our data), which may explain the high omission rates in these contexts.

### ***Discourse-Pragmatic Context***

The effect of discourse-pragmatic context on the accuracy of “the” is barely significant. Learners are predicted to be more accurate in situational than in anaphoric contexts but, puzzlingly, in plurals only. Previous research has also found situational contexts to be easier than anaphoric ones (Liu & Gleason, 2002), with higher omission rates in the latter potentially due to perceived redundancy of the article (Robertson, 2000).

One reason I may not have found a reliable effect is that the written production data from EFCAMDAT is not as rich in terms of discourse functions as the more controlled data elicited by, for instance, Robertson, whose participants engaged in a picture dictionary task, giving oral instructions to each other to draw specific pictures.

Additionally, considering the large confidence intervals for the predicted accuracy rate in the anaphoric context and the fact that I was unable to include this variable as a random effect, this effect might have emerged in plurals because of considerable individual variation. Thus, I cannot

reach a definitive conclusion about the effect of discourse-pragmatic context on learner article accuracy.

### ***Abstractness***

The effect of abstractness is only marginally significant in count singular indefinites. Previous research suggests learners may be more uncertain about the countability of abstract nouns (Butler, 2002; Hua & Lee, 2005). In the data distribution (Figure 1), indefinite mass nouns were more often abstract than concrete, while for all other noun types (count indefinites and all definites) abstract and concrete nouns are equally represented. I therefore explored the possibility that this contingency between abstract and mass nouns might impact on learner article accuracy. However, I did not find any significant effect of abstractness on article accuracy with indefinite mass or any other type of noun. There are numerous clearly count abstract nouns (e.g. “problem”, “question”) and mass concrete nouns (e.g. “water”, “rice”), which are likely well-known to learners. This may be the reason learners realise early on that abstractness and countability are not linked. As seen above, what seems to matter to learners is whether the referent is specific, or refers to a certain existing entity, regardless of whether it is abstract or concrete.

### ***Limitations***

An important limitation of this study is that the data is drawn from learner writings produced in response to specific task prompts. First, this constrains the range of discourse functions compared to data from conversational exchanges, which may be the reason I was unable to make any conclusions about the role of discourse-pragmatic functions of the definite article. Second, each writing task in EFCAMDAT is completed off-line at the end of a module, which implies preparation, scaffolding, and the possibility to edit the response. As a result, the accuracy rates I observed are probably overestimated and could be considerably lower in spontaneous production.

Another important limitation is that, because L1-Chinese learners turned out to behave similarly to other [+art] L1 groups, it is impossible to be sure that the patterns observed here are not specific to L1-Russians rather than [-art] L1 learners more generally.



Finally, the results of the statistical analysis of error types should be treated with caution, since data had to be split into subsamples for such analysis. While the size of the total sample was justified with a power analysis, meaning that one can be reasonably certain of its statistical power, the analysis of data subsets by error type were likely underpowered. Nevertheless, the fact that the findings within those data subsets were in line with the overall patterns revealed by the Full Accuracy Model suggest that the analysis was probably not misguided.

## Chapter 4: Teaching and Learning Articles in L2 English

### Teaching Articles to English Learners: Previous Research

Articles in L2 English appear to be a persistent challenge even at higher proficiency levels. Therefore, the search for an optimal teaching method for this feature is well justified. At the same time, articles are not usually given as much time and attention in English teaching curricula as other difficult aspects of English, such as tense and aspect. It is possible that articles are viewed as less important for communicative purposes (Master, 2002). In addition, article practice is not easily incorporated into natural task-based or communicative activities, unlike tense and aspect, which allow for a variety of meaning-focused engaging practice tasks (e.g. storytelling, describing future plans etc.). The only function of articles that lends itself to similar kinds of tasks is probably the anaphoric use of the definite article, which could be practised in story-telling activities with first- and second-mention referents. Other aspects of article use tend to be practised in a more grammar-focused style in mainstream English teaching materials.

A series of empirical studies aimed at testing the effects of various types of corrective feedback on SLA chose English articles as their target linguistic feature. Overall, they found a positive effect of explicit corrective feedback on article accuracy.

Muranoi (2000) found a positive effect of Interaction Enhancement (IE), a method based on interaction enhanced by implicit negative feedback, on article accuracy in L1-Japanese English learners, particularly when formal debriefing was also provided. Implicit negative feedback points out learner errors without explicitly stating what the error is. For instance, implicit negative feedback can be given in a form of a clarification request: if a learner makes an error, the teacher asks them to repeat or clarify what they said hoping that this would draw the learner's attention to their error and perhaps lead them to correct it. Another form of implicit negative feedback as a recast, i.e. repeating the learner's words back to them without their error, so that the learner can notice the difference. In pairs, 91 participants rehearsed role-plays, which were designed to elicit the indefinite article. Then

the teacher appointed one of the students and acted out the role-play with them providing implicit negative feedback on the indefinite article while the other participants observed.

Interaction Enhancement was implemented in two experimental groups, which varied according to the type of debriefing they received after each of the three training sessions, i.e. formal (IEF) vs. meaning-focused (IEM) debriefing. In the former, participants received explicit information about their errors and an explanation of the functions of articles with a focus on the communicative value of the indefinite article “a”. Specifically, the teachers explained that the two main functions of “a” were to introduce new referents into the discourse, as well as to classify nouns (i.e. non-referential use of “a”). Importantly, the IEF debriefing did not emphasise the role of countability and number in deciding whether “a” or  $\emptyset$  would be more appropriate. The teachers in this group also told the participants the possible reasons for their difficulties with articles, including the lack of articles in their L1 Japanese and the low salience of articles. By contrast, the IEM debriefing centred around participants’ communicative performance, e.g. how well they were able to express their opinion or disagree with the interlocutor. There was no mention of articles or any other morphosyntactic feedback in this group.

The results of the treatment were measured by a series of tasks, including oral story description, oral picture description, written picture description and grammaticality judgment tasks. The IEF group improved most on both the indefinite and the definite article use and retained the results by the delayed post-test. The IEM group was significantly better than the control group, but significantly worse than the IEF group at the immediate post-test. However, at the delayed post-test the difference between the IEM and the control group became insignificant as the latter also improved. Nevertheless, no practice effect was found, as the improvements of the control group were not significant. Interestingly, no significant differences were found between those students who participated in role-plays with the teacher and those who observed. Muranoi concluded that the IE is generally beneficial and that the IEF is more beneficial than the IEM. Several points should be considered regarding these results. Firstly, the initial proficiency level of the participants was only

measured by their accuracy on the English articles, which appeared to be very low compared to previous research findings (25-42% accuracy). Secondly, the amount of negative feedback actually given was very small. Each experimental group received 9 requests for repetition from the teacher and 5-6 recasts across all three training sessions. Thirdly, it is unclear why training the use of the indefinite article led to improvements in the use of the definite article, considering that they are not mutually exclusive. Finally, setting up IE training is rather complicated from a practical point of view, requires specially designed role-plays and teachers' meticulous attention to article errors.

A few studies compared simpler and more traditional types of corrective feedback and their effect on learner article accuracy. Sheen (2007) found that learners benefited from error correction, especially when metalinguistic feedback was provided alongside. In her experiment, two experimental groups of intermediate-level ESL learners (mostly Korean, Hispanic, and Polish) from a language school in the USA were asked to rewrite narratives and received either error corrections only ( $n = 31$ ) or error corrections together with metalinguistic feedback ( $n = 32$ ). A control group ( $n = 28$ ) followed regular classes with no instruction on articles. They were tested with a speed dictation, narrative writing, and an error correction task. Both experimental groups improved, especially the one receiving metalinguistic feedback, and Sheen also found that higher linguistic aptitude, which she measured as part of the pre-test, further enhanced the effect.

Ellis et al. (2008) showed a positive effect of focused (corrections of article errors only) and unfocused (correction of other types of errors in addition to article errors) corrective feedback on article accuracy of 49 L1-Japanese EFL learners, when compared to a control group who received no feedback. They used narrative tasks, similar to the ones in Sheen (2007) during the treatment stage and narrative writing and error correction as testing tasks. Using narrative picture descriptions, Bitchener (2008) found a significant improvement in all three experimental groups of ESL learners in New Zealand (mostly with East Asian L1s), who received different types of corrective feedback (error corrections with written and oral metalinguistic feedback, error corrections with only written

metalinguistic feedback, error corrections with no other feedback), when compared to the control group receiving no feedback.

The main focus of these studies, however, was on the effects of corrective feedback, and little attention was given to the different factors which determine article use in learner English. In fact, they targeted a very limited set of article functions (only indefinite “a” introducing new information in Muranoi (2000); only first-mention “a” and anaphoric use of “the” in Sheen (2007), Ellis et al. (2008), and Bitchener (2008)).

Another experiment by Kao (2020) focusing on a variety of article uses demonstrated a positive longer-term effect of digital game-based learning in intermediate EFL learners in Taiwan. The first group ( $n = 15$ ) played a digital game where they filled in the gaps with articles while interacting with different characters in a business-related computer game (e.g. they were talking to potential job applicants). The rate of their article accuracy determined whether the game characters agreed to do business with them. Players received immediate corrective feedback and metalinguistic explanations. The second and third group (15 participants in each) completed similar gap-fill activities on paper with the ability to consult metalinguistic explanations on the use of articles, while the second group also received corrective feedback for their pre-test writing (a job application letter). Kao found that the second group initially improved their article accuracy more than the digital game group at the immediate post-test (another job application letter). However, the digital game group improved further by the delayed post-test a month later, while the second group’s performance slightly declined. The third group made no gains throughout. Although this study included a variety of article uses, it still did not include any factors impacting learner article use in their tests, as conclusions were made based on overall TLU scores lumping the different uses of the definite and the indefinite article together. Moreover, the metalinguistic explanations provided in the experiment were too simplistic and potentially misleading. For example, they explicitly equated specificity and definiteness by stating that “in some textbooks specific nouns are called definite nouns” (Kao, 2020, p. 6), making no distinction between specific and non-specific indefinites.

Several other studies managed to incorporate the different linguistic factors influencing article use, such as definiteness, specificity, number and countability, in their training materials and tests. However, they showed very little to no gains in article accuracy.

For instance, Master (1994) tested the effect of instruction based on six major distinctions in the article system in English, such as count/mass, singular/plural, indefinite/definite, specific/generic, common/proper, idiomatic/non-idiomatic. The treatment lasted for 6 hours over 8 weeks and was administered to 14 high-intermediate/low-advanced students with a variety of L1s, while 33 other subjects served as a control group. Master found a moderate effect size of the treatment (0.664), however participants actually gained only 2.29 points, or 6%, on average improving from 26.79 to 29.08 points out of 36, while the control group gained 0.63 points, or 2% (from 26.61 to 27.24), which was insignificant. Master replicated this study (reported in the same paper) increasing the number of hours of treatment but received very similar results. Overall, it appears that such treatment would take a very long time to yield tangible increases in learner accuracy.

Snape and Yusa (2013) focused their treatment on different ways of expressing genericity (non-referentiality) with articles. They distinguished between NP-level and sentence-level generics based on Krifka et al.'s framework (1995). According to this framework, NP-level generics refer to a "well-established kind" or a "natural kind". They are used with kind predicates and can be expressed by definite singulars (e.g. "the pelican is a protected species") or bare plurals (e.g. "pelicans are a protected species") (mass nouns were not included). Sentence-level generics refer to general properties of nouns in subject positions and can be expressed by indefinite singulars (e.g. "a coat is necessary in winter") and bare plurals (e.g. "coats are necessary in winter"). L1-Japanese high-intermediate learners participated in three 70-minute training sessions, in which they made dialogues choosing appropriate articles, practiced interpreting articles in sentences, and constructed sentences with generics. Both control and experimental groups mostly accepted indefinite singulars and bare plurals as generics but did not accept definite singulars as generic (NP-level) on the

acceptability judgment task. However, the treatment did not improve learners' accuracy on the forced-choice elicitation task.

Umeda et al. (2017) replicated Snape and Yusa's (2013) study with some modifications. They hypothesised that the lack of improvement may have been due to insufficient time for treatment, so they increased the number of instruction sessions to nine. They also contested that instruction in L2 may have been too complicated for high-intermediate learners to grasp because of the complexity of the topic, which is why they included instruction in L1. As a result of the treatment, L1-Japanese learners improved on their acceptance of definite singulars as generics on the acceptability judgment task, which was low in the pre-test but significantly improved in the immediate post-test and the delayed post-test (12 weeks after the intervention). The improvement on indefinite singulars (also low at pre-test) was only observed in the immediate post-test. However, none of the improvements were retained by the second delayed post-test one year after the intervention.

A very different approach was adopted by Zhao and MacWhinney (2018), who used instruction to promote cue focusing within the Competition Model developed by MacWhinney (1997). The basic idea of the Model is that different forms constitute cues for functions that they can map on. Cues may have different strength, which explains why certain forms may be chosen over others and why they may be acquired earlier. Zhao and MacWhinney first established cue availability (the number of cues per the number of NPs) and cue reliability (whether a cue always predicts the use of a form) for the English articles using a mini-corpus of different written genres (38 texts, ca. 24000 words, ca. 3700 NPs). Some of the most available cues included plural and mass NPs taking  $\emptyset$  (e.g. "books", "water"). The average reliability of the ten most available cues was rather high (93%), however some cues, particularly plurals and singulars with post-modification (e.g. "the letters I received today", "the man she is dating"), had very low reliability (39% and 43.5% respectively). Sixty-four L1-Chinese intermediate-advanced learners of English participated in two 45-minute training sessions in which they were shown sentences with contrasting cues (23 cues were chosen for the intervention) and required to choose appropriate articles. In case of error, learners were

given metalinguistic feedback and had an opportunity to change their answer. The experimental group participants' accuracy and response time on gap-fill activities improved significantly in the immediate post-test. There were no improvements in the control group. However, no delayed post-test was administered. A more detailed look at the cues and examples of the contrasted sentences used in the intervention reveals that the representation of article functions was highly restricted. In essence, the intervention materials trained participants to associate certain forms with articles. For instance, participants were instructed to "use *the* when a noncountable noun is post-modified by a relative clause or a prepositional phrase" and to "use  $\emptyset$  with unmodified noncountable nouns" (Zhao & MacWhinney, 2018, p. 118). It is not surprising that the participants improved their accuracy on the post-test, which contained the same types of sentences as the training sessions. However, it is questionable whether such training would help learners to deal with any other uses of the articles, which were not presented during the intervention, such as using "the" with unmodified mass nouns in situational contexts (e.g. "pass me the water").

Summing up, articles have been used as a convenient target feature (due to their non-salience and common lack of attention in teaching) in multiple studies of the effects of various types of corrective feedback. Furthermore, most experiments focused on a limited set of article uses. The few studies that did consider the complex interaction of linguistic factors affecting learner article use showed meagre results. An important point raised by Kao (2020), who attempted to address it in their digital game experiment, is the ecological validity of many of the above studies, as their experimental treatments would be impractical in real English classrooms, which strive to be more meaning- and task-oriented.

### **Articles in L2 English: The Learning Problem**

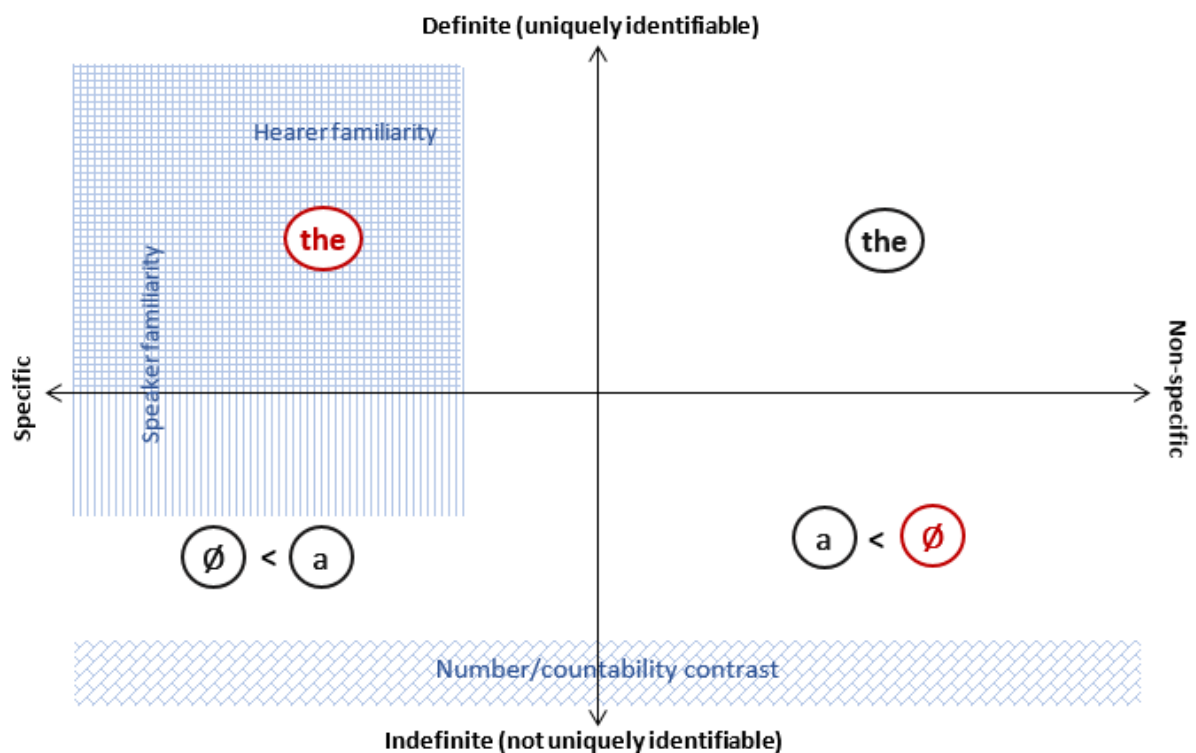
One common shortcoming in many of the studies described in the previous subsection is that they fail to consider the impact of multiple intertwined nominal and discourse-pragmatic features on article use in learner English.



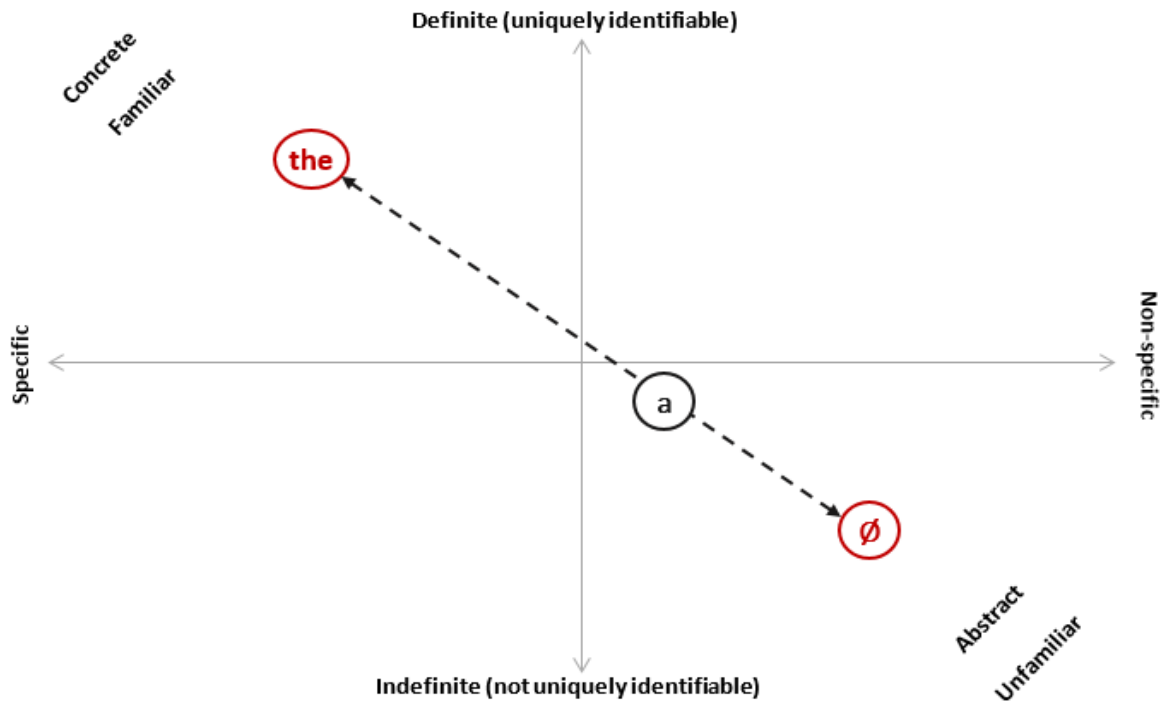
In target English, the distribution of articles on the definiteness-specificity plane can be envisaged as shown in Figure 16. “The” is the only possible option in contexts where specificity (as defined in this study), definiteness and familiarity converge but also in contexts that are definite and non-specific, although the latter are rather infrequent (e.g. “the winner of tomorrow’s race”).

**Figure 16**

*Target English Distribution of Articles on Definiteness-Specificity Plane*



For indefinite contexts, there are two possibilities: “a” and  $\emptyset$ . The distinction between the two in target English is essentially determined by the number/countability contrast, which applies to all indefinite contexts regardless of specificity. However, based on the distribution found in the corpus study reported in Chapter 3 and on the assumption that it is similar to the distribution in native English, I hypothesise that there are certain contingencies in the input. First, non-specific referents, i.e. those that do not refer to certain existing entities, tend to be mass or plural nouns (i.e. target  $\emptyset$ ), while mass indefinites also tend to be abstract. At the same time, specific referents are more often singular nouns and, thus, require “a”. Furthermore, in specific contexts  $\emptyset$  can often be filled by determiners, which indicate specificity, e.g. “some”, “much” etc.

**Figure 17***Hypothesised Distribution of Articles in Learner Interlanguage*

Based on such distribution in the input, learners may incorrectly perceive “the” to be the prototypical article to be used in specific definite contexts, while  $\emptyset$  may be considered the prototypical feature of non-specific indefinite contexts (Figure 17). The article “a” appears to be an intermediate option without a clearly defined place on the definiteness-specificity plane. It may be used in contexts that are either specific or non-specific; the referent in such cases may or may not be familiar to the speaker.

***Interpretability Hypothesis and L2 English Articles***

It is useful to consider the impact of the various features described above in light of the Interpretability Hypothesis (Tsimpli & Dimitrakopoulou, 2007; Tsimpli & Mastropavlou, 2007), which allows to link it to the developmental process of L2 acquisition. The Interpretability Hypothesis makes important claims about the acquirability of interpretable versus uninterpretable features. The former are those features which contribute to the meaning of an utterance and are said to be interpretable at the Logical Form interface level, according to the Minimalist Program (Chomsky, 1995). These include, for example, number or gender on nouns. By contrast, uninterpretable

features are necessary for syntactic derivation but are not interpretable at the Logical Form interface, i.e. they have no semantic contribution. Examples of uninterpretable features include gender on adjectives, or grammatical case. According to the Interpretability Hypothesis, adult L2 learners (beyond critical period, which is assumed under this hypothesis) cannot access uninterpretable features unless they are instantiated in their L1, whereas interpretable features should be available to any L2 learner regardless of their L1 background.

From this perspective, we can state that English articles contain both interpretable and uninterpretable features. On the one hand, both the definite and the indefinite article have the interpretable definiteness feature. On the other hand, the indefinite article “a/an” can also carry uninterpretable features of number and countability. Additionally, the indefinite article may not always have the [-definite] feature in English, as in predicates, for example “She is a doctor”, where it has a purely syntactic role of licensing a noun phrase. Based on this analysis, we could predict that adult learners would have more difficulties with the indefinite article “a” if their L1 is article-less, which is what was found in the corpus study reported in the previous chapter.

However, it is not expected that L2 learners will necessarily display true optionality in their use of items which carry uninterpretable features. Tsimpli & Mastropavlou (2007) suggest that in order to avoid optionality, more advanced L2 learners may associate such items with other interpretable features, especially if the input lends itself to nontarget associations. This is directly relevant to the distribution of English articles discussed above, as there is indeed another interpretable feature of specificity which can be erroneously associated with articles. Such misspecification is probably more likely to affect the indefinite article “a”, which contains uninterpretable features, and the use of which learners might strive to regularise in some way. In other words, learners will want to find a pattern for the distribution of “a” drawing on any available interpretable features, and if the input appears to favour the use of “a” with specific referents and  $\emptyset$  with non-specific referents, this could prompt learners at higher proficiency levels to rely on specificity to constrain their use of “a” vs.  $\emptyset$  in indefinite contexts.

Whilst having targetlike representations and appropriate form-meaning associations generally leads to higher accuracy in both comprehension and production, it is clear that learners are not always accurate even with interpretable features. One reason for this is L1 transfer effects, which can affect interpretable features, such as plural marking on nouns (Murakami, 2013). The effects of L1 transfer on L2 English articles have also been observed in the corpus study reported in the previous chapter. Another reason, which has not been discussed in connection with the corpus study, is cognitive constraints.

### ***Cognitive Constraints and Processing***

Cognitive constraints may affect both the processing of articles in the input and the use of articles in oral or written production. The reason cognitive constraints are particularly relevant here is that using a language other than one's first language is known to be more cognitively demanding. This is because it is believed, based on existing empirical evidence, that both L1 and L2 (and L3 etc. if relevant) are activated in learners' minds at all times (Kroll et al., 2014), although they can be inhibited to varying degrees depending on situation. For example, in an immersion setting the dominant language will be more activated, while any other languages can be more inhibited. Having higher proficiency in a language can also boost its activation in comparison to a language at a lower proficiency level. This means that inhibiting one's first language will be more difficult than one's other languages until a very high level of proficiency is achieved.

On the one hand, such cognitive constraints affect input processing, as learners may not have enough cognitive capacity to attend to all the aspects of the input. Various proposals in the literature point to learners' tendency to rely on lexical or pragmatic information to infer meaning while morphosyntactic cues may be overlooked. This may be due to the way certain features or items are processed in learners' L1, which can be transferred to L2 processing resulting in "attention blocking" (Ellis & Sagarra, 2010). To illustrate, if learners are used to relying on lexical information to infer tense, as in Mandarin Chinese for example, they may transfer this reliance to their L2 processing, which would "block" their attention to morphological tense marking. Another potential cause is a

general difficulty in accessing morphosyntactic information in on-line processing, as suggested under the Shallow Structure Hypothesis (Clahsen & Felser, 2006).

Importantly, these difficulties in processing articles in the input can hinder the development of targetlike L2 representations even if the features are interpretable and available in principle. In a similar vein, Input Processing theory (VanPatten, 1996, 2002) suggests that learners will prioritise more meaningful morphological features in their input processing. The theory further claims that inappropriate processing or lack of processing of morphological features will result in their slower or more problematic acquisition.

On the other hand, cognitive constraints can impact on-line production of articles even if L2 representations are targetlike. This is particularly true for any features that differ between L1 and L2, as constant activation of the L1 in a learner's mind will lead to competition between L1 and L2 forms. In the case of articles, this could be a competition between bare noun forms allowed in L1 and "article + noun" forms required in English, and the former could be more activated at lower English proficiency levels. Crucially, English "article + noun" forms may even be processed appropriately in comprehension but may still be competing with bare nouns in production (Trenkic et al., 2014).

In addition, processing can be significantly affected by task complexity. According to the Limited Attentional Capacity Model (Skehan, 1998), higher complexity of a task forces learners to divert their limited attentional resources towards task content and away from monitoring their linguistic accuracy. However, Robinson's Cognition Hypothesis (2005) suggests a more intricate view distinguishing between various dimensions of task complexity. The hypothesis claims that increase in complexity of certain task aspects, such as reasoning demands, may direct learners' attention to linguistic form, although there are other aspects of complexity (e.g. planning time or task structure) which can have a negative impact on linguistic accuracy.

### ***Processing of L2 English Articles***

According to the Input Processing theory (VanPatten, 1996, 2002), appropriate processing of a feature in the input (which is critical for making correct form-meaning mappings and forming targetlike representations) depends on its meaningfulness and communicative value, which comprises inherent semantic value and redundancy.

Considering the above overview of cognitive constraints on L2 processing and production, one might expect that learners would encounter difficulties in processing English articles, especially if their L1 is article-less. This is because articles are function words and, thus, are not likely to be prioritised if cognitive resources are limited. Nevertheless, articles do have some communicative value and contribute to meaning in certain contexts, and the definiteness feature itself is considered to be interpretable, although the indefinite article in some contexts performs only a structural function and carries some uninterpretable features (number and countability). There is empirical evidence that learners can be sensitive to articles' semantic contribution, even if their L1 is article-less.

In a visual-world eye-tracking study, Trenkic et al. (2014) demonstrate that L1-Mandarin-Chinese learners of English behave similarly to native speakers in utilising articles "a" and "the" to decide what nominals refer to. In this experiment, 48 intermediate-level L1-Mandarin English learners and 56 L1-English participants were presented with audio stimuli of the form "The [agent] will put the [theme] inside the/a [goal]" (Trenkic et al., 2014, p. 7) along with visual displays each showing an agent (e.g. a pirate or a queen) looking at a "theme" object (e.g. a cube or a pencil) surrounded by five other objects – three containers (e.g. a basket or a can) and two non-container distractors. Of the three containers, two were potential "goals" and one was a distractor. In two-compatible trials both potential "goals" were available (e.g. two open cans or two empty baskets), whereas in one-compatible trials one was unavailable (e.g. a closed can or a full basket). This was counterbalanced with the article preceding the "goal" noun in the audio stimulus – definite "the" or indefinite "a", resulting in four different experimental conditions. The researchers used eye-tracking

to record the timing and location of participants' fixations. The task was to click on the object where participants thought the "agent" would put the "theme". As expected, L1-English participants decided sooner which object to click on (as measured by the timing when they started fixating on one of the potential "goals" significantly more than on the other one) in two-compatible "a" trials and in one-compatible "the" trials. In other words, they could choose a container more quickly when they heard "The pirate will put the cube inside a can" and saw two open cans, and when they heard "The pirate will put the cube inside the can" and saw only one open can (with the other one closed). L1-Mandarin participants were slower overall but showed the same pattern as native speakers.

Trenkic et al. conclude that these results support the Competition Model of L2 processing (MacWhinney, 1997) in that their participants were able to process articles in well-formed sentences in a target-like manner due to the lack of a competing morphological element in their L1. The authors go on to explain that appropriate processing in comprehension does not imply accurate production, since during production bare nominal structures from learners' L1 are expected to be activated and competing for selection with "article + nominal" structures.

They also clarify that learners' accuracy in interpreting articles in well-formed sentences does not necessarily entail sensitivity to article omissions, as evidenced by the results of a self-paced reading experiment with L1-Mandarin and L1-Croatian learners of English in Study 1 in O'Reilly (2018). In other words, while learners may interpret "the" and "a" appropriately when they are present in the input, this does not mean they would not treat a bare nominal as a viable alternative and interpret it just as swiftly based on context (as they would in their L1). The fact that English allows bare nominals (mass and plural indefinites), which can account for up to 50% of nominals in the input (Master, 1997), only strengthens this misconception.

Therefore, from the point of view of VanPatten's Input Processing theory, articles can have inherent semantic value, which learners can recognise and interpret. However, there is no theoretical reason to suggest that this translates into accurate production. Moreover, it appears that the claim is justifiable for certain types of contexts which are captured by the visual-world eye-

tracking method, i.e. visible concrete objects in immediate situations. It is not clear whether articles' semantic value would be as transparent to learners in more abstract contexts. Finally, there are common instances in English when articles do not carry much semantic value, such as the use of "a" to license predicates (e.g. "She is a doctor") or idiomatic use (e.g. "in the morning").

As for redundancy, the second criterion for defining communicative value according to VanPatten, article misuse or omission rarely lead to communicative failure (Master, 2002). In fact, articles can be difficult to notice in the input, as they are usually phonologically non-salient and can even be dropped in colloquial speech (Rastall, 1995). In sum, it is reasonable to assume that articles "a" and "the" are of low priority for learners when it comes to processing, with the exception of contexts where their semantic contribution is transparent, such as those created in the visual-world eye-tracking paradigm.

In addition, the use of "a" is complicated by the number/countability feature, which requires the learner to distinguish between count singular, count plural, and mass nouns. Since, according to the input processing principles (VanPatten, 2002), learners tend to rely on meaning in the first instance, one might expect that they would initially prioritise referent resolution as a communicative goal, which may result in a focus on semantic and discourse-related features, such as specificity, familiarity, identifiability, hindering the recognition of structural cues, such as number and countability. This is especially true for learners with [-art] L1s, as they cannot rely on transferring a similar form-meaning mapping from their L1. Thus, [-art] L1 learners are unlikely to map "a" onto uninterpretable structural features. Given learners' natural propensity to attach meaning to forms and the fact that  $\emptyset$  may co-occur with non-specificity and abstractness in the input, as observed in the corpus study, inappropriate processing can be expected. This would also be expected under the Interpretability Hypothesis, which, as noted above, suggests that higher-proficiency learners may resort to compensatory strategies to regularise their production by associating forms with other available interpretable features (in case of the indefinite article, this could be specificity, for



example). Inappropriate form-meaning mapping, in turn, prevents input from becoming intake and, thus, according to VanPatten, leads to acquisition problems.

### ***Sources of Different Types of Errors***

When discussing issues with article use or article accuracy, it is important to consider that these can manifest in several ways in learners' processing and production. As was observed in the corpus study, learner language production can show omission of "the" or "a", substitution of "the" with "a" and vice versa, or overuse of "the" or "a" (i.e. using an article where  $\emptyset$  is expected).

Somewhat similarly, in comprehension learners could "miss" the article or be insensitive to omission errors, i.e. not process it at all, or they could misinterpret "a" as "the" and vice versa or be insensitive to substitution errors.

In production, it is known that omission is by far the most common concern compared to substitution or overuse. Based on what is known from article processing studies, learners are not sensitive to either omission or substitution errors in online processing, although they are more likely to notice omission errors in more metalinguistic tasks, such as grammaticality judgment (O'Reilly, 2018). Moreover, they can appropriately process articles (i.e. they do not necessarily "miss" them) when articles are present in the input and contribute to the meaning (as in the visual-world eye-tracking paradigm in Trenkic et al., 2014).

Considering the various factors which impact on article choice by English learners discussed so far, it is unlikely that all these errors and issues can be attributed to the same cause. On the one hand, omission errors in production and insensitivity to omission in the input is closely related to L1 effects, as this issue is typically more pronounced in learners with article-less L1s, as well as to cognitive constraints, which are often invoked to account for optionality in learner language. Even when the level of proficiency increases, omission might still persist, albeit to a lesser extent, as advanced learners may create inappropriate form-meaning mappings.

There is a slight difference between omitting "a" and omitting "the", as the former may result from ignoring structural cues, or uninterpretable features of number and countability,

whereas the latter may involve failure to mark definiteness (e.g. “time” vs. “the time”), which is an interpretable feature. At the same time, both omission errors could simply be attributed to failure to supply an article with a noun phrase regardless of its definiteness status due to limited cognitive resources, which force the learner to focus on more meaningful elements. One could argue that systematic omission which persists at higher levels and on tasks with lower cognitive demands would point towards an inappropriate form-meaning mapping, whereas more random omission which increases in more complex and online tasks might require a cognitive explanation. However, many learners are probably susceptible to both issues. Thus, one would expect less omission and more systematicity (even if non-targetlike) in less demanding contexts and the opposite for more demanding ones. Whilst omission could have two different explanations, article overuse is more likely due to an inappropriate form-meaning mapping, as producing an article where it is not needed requires additional cognitive resource and is more likely intended rather than accidental.

As for substitution errors, i.e. using “the” instead of “a” and vice versa, this type of error always involves inappropriate marking of (in)definiteness and, thus, has to do with an interpretable feature. An important caveat is that deciding whether a nominal is definite or indefinite is not as straightforward as deciding whether it is singular or plural. For example, in “I didn’t have (the) time to do it”, both versions may be judged as acceptable, as native speakers can accommodate either option by imagining an appropriate scenario. By contrast, “I didn’t have a time” would simply be unacceptable. This is because definiteness is determined either by discourse, which in some cases can be interpreted differently depending on the interlocutor’s perspective, or by prior/context knowledge, which can vary from person to person depending on their background and life experience. Thus, even a slightly unexpected statement, for example “Close a door” rather than “Close the door”, can be accommodated if the listener imagines an appropriate context. This is supported by the findings in O’Reilly (2018), whose L1 English participants were insensitive to substitution errors (but not to omission errors) on a self-paced reading task. Debatable cases aside, substitution errors in contexts that are clearly definite or indefinite (for example, cases of

unambiguous anaphoric reference) appear to point towards inappropriate form-meaning mappings, at least in production where article presence itself demonstrates that the learner had sufficient cognitive resource to produce it. However, it could also be argued that cognitive demands might make it more difficult to follow discourse, especially at lower levels of proficiency. Similarly, insensitivity to substitution errors in the input during online processing can readily be explained by cognitive constraints. Certainly, given the interpretability of definiteness, one would expect substitution errors to reduce significantly with proficiency, even in article-less L1 learners.

In sum, inappropriate form-meaning mappings are a more likely explanation for the omission and overuse of “a”, where uninterpretable features are involved, particularly at higher levels, where learners should have more cognitive resources available during L2 use and probably strive to regularise any optionality. In addition, all kinds of omission, particularly at lower levels, can also be explained by cognitive constraints and learners consequent focus on content and more meaningful elements. With the omission and overuse of “the” and substitution errors, both cognitive and form-mapping issues can be invoked, although the latter are less expected due to the interpretability of the definiteness feature.

### **Processing Instruction Intervention for L2 English Articles**

Assuming that learner problems with English articles may stem from inappropriate form-meaning mappings, which may be caused by a combination of L1 effects, processing constraints and a certain distribution of articles in the input, articles appear to be a good candidate for treatment with Processing Instruction (PI). PI is a type of intervention which focuses on structuring input in such a way that would amend learners’ erroneous processing strategies, i.e. help them make correct links between forms and their meanings (VanPatten, 2015). Once the learner becomes aware of the correct form-meaning mapping and continues processing the feature correctly in further input exposure, this is supposed to trigger successful acquisition.

The centrepiece of a PI intervention is a series of structured input activities, which modify input in such a way that forces correct processing of a linguistic feature leading to the formation of

correct form-meaning mappings. This is achieved by stripping away any context that might help learners to derive the intended meaning without attending to the target feature. For example, to ensure appropriate processing of the past tense “-ed” in L2 English, learners may be presented with a sentence, such as “I called Mary” without any temporal lexical cues (e.g. “yesterday”) and asked whether the sentence refers to the past, the present, or the future. Without access to other sources of meaning, learners must rely on the form (past tense “-ed”) to derive meaning creating appropriate form-meaning mapping.

Structured input activities can be of two types. At first, learners usually complete referential structured input activities, which always have a right or wrong answer, such as the example in the previous paragraph. Learners receive immediate correctness feedback after each response. Then, learners are given affective activities, in which they are asked for an opinion, where correctly processing the target feature is necessary in order to respond. For instance, they might be asked to mark statements such as “I called my parents at the weekend” as true or false. Importantly, they are not required to produce any language throughout the entire intervention, as the activities are all based on comprehension. This is due to the stipulation that PI will trigger intake, which will form correct representation and, thus, lead to accurate production.

In addition to the core structured input activities, PI usually involves giving learners explicit information on the target feature and the common processing issues at the beginning of the intervention.

The effectiveness of PI has been demonstrated for various linguistic features in different L2s: object pronouns in L2 Spanish (VanPatten & Cadierno, 1993; VanPatten & Sanz, 1995), preterit in L2 Spanish (Cadierno, 1995), simple past tense in L2 English (Benati, 2005), passives in L2 English (VanPatten & Uludag, 2011), causative constructions in L2 French (Wong & Ito, 2018). Positive effects on accuracy were reported not only on comprehension tasks but also in production, despite the fact that production is not required during the intervention (VanPatten & Uludag, 2011),

supporting the claim that transforming input into intake is sufficient to initiate a cascade of processes eventually leading to target-like production.

The added benefit of this kind of intervention is that it is not difficult to set up once the materials are ready and it does not require much involvement from the teacher, especially during the structured input phase.

### **Rationale and Research Questions**

The next step in this study is the development and implementation of a PI intervention targeted at articles, which will incorporate most of the linguistic variables investigated in the corpus study. Based on the results of the corpus study, the intervention materials will probably not include discourse-pragmatic context as a factor, as this was not found to have a significant impact on article accuracy, while also being a rather cumbersome multi-level variable. However, abstractness should probably be included despite its small effect because of the potential association between abstractness and non-specificity evident in the distribution of items in the corpus study. It would, thus, be important to counter this perception by crossing abstractness and specificity factors in materials design. The main goal is to explore whether a linguistically motivated PI intervention can significantly improve learner article accuracy.

Using PI as opposed to any other kind of intervention appears appropriate, since it connects input, which it aims to convert to intake, with representation. As far as articles are concerned, there is clear evidence, on the one hand, that target English input may be misleading, and on the other hand, that learner representations reveal certain non-targetlike patterns (i.e. article errors are not random). It is hoped that PI can provide the missing link.

It is important to acknowledge that there are contexts in which articles do not contribute much to meaning. In fact, there is an argument that that articles' primary function is to signal a noun phrase, while their association with definiteness marking may be incidental (Hawkins, 2004; Trenkic, 2009). Therefore, using PI, which forces form-meaning mappings, is in some sense and in some contexts misleading, e.g. "This is a problem" is most likely used to describe a situation rather than

state that there is only one problem, “I want an apple” does not necessarily imply strictly one apple etc. Nevertheless, it appears that PI might exploit learners’ natural tendency to attach meanings to forms and help direct them to more appropriate mappings by artificially highlighting certain aspects of articles, such as the link of the indefinite article to number and countability, and the definiteness feature distinguishing between the definite and the indefinite articles. In other words, PI could prove to be an effective approach from a pedagogical perspective, although it may not reflect the true nature of articles in English.

Thus, the main research questions for the intervention study are the following.

1. Will learner article accuracy benefit from a PI intervention which incorporates the linguistic factors affecting learner article use?
2. Will any of the linguistic factors have differential effects on the outcome of the intervention?

I hypothesise that most of the gains will be observed in reducing “a” omission and overuse, as these are the error types where inappropriate form-meaning mappings are a more likely cause, as explored above.

3. Will the effect of the intervention be mediated by learner variables, such as proficiency level and L1-type?

On the one hand, I would expect less gain in lower-level learners, as instilling appropriate form-meaning mappings might still not help as it does not remove the significant cognitive constraints that such learners are subject to, as discussed above. On the other hand, higher-level learners will have less room for improvement, so their gains might be less considerable by comparison.

I would also hypothesise more improvement for [-art] L1 learners, especially those at higher levels, as they would be most likely to have formed inappropriate form-meaning mappings to regularise their use of the indefinite article, which is dependent on uninterpretable features.

## Chapter 5: Processing Instruction Intervention Pilot Study

In this chapter, I describe the initial design of the intervention materials and report the pilot study trialling those materials. Conducting a pilot was deemed essential as all the materials, including treatment and assessment, were new and created specifically for this study. The results of the pilot revealed several issues and were used to make the necessary improvements before the main data collection, which is reported in Chapter 6.

### Methodology

#### *Intervention Materials*

**Experimental Group Materials.** The PI materials for the experimental group in this study were designed according to the Processing Instruction guidelines (VanPatten, 2015) and included explicit information given at the beginning of some of the activities and the structured input activities themselves<sup>22</sup>.

Explicit information was given in English and consisted of some brief guidelines for choosing the correct articles (in a form of an algorithm, example in Appendix C) drawing learners' attention to some typical processing issues (for example, learners are explicitly told that article choice does not depend on whether they refer to a specific item or an idea in general).

The explicit information was followed by structured input activities, which were divided into two large blocks. In the first block learners practised using the indefinite article, i.e. "a" vs.  $\emptyset$ . In the second block they practised choosing between the definite and the indefinite article, "the" vs. "a"/ $\emptyset$ . The main reason for such order of presentation is that the latter implies the ability to distinguish between "a" and  $\emptyset$  unless the materials are limited to count singular nouns only. It is also justified by the fact that the use of "a" involves number and countability features, which are more straightforward to grasp, whereas to distinguish between the definite and the indefinite article the learner needs to interpret discourse or situational context, which can be more debatable.

---

<sup>22</sup> Only referential structured input activities were included in this experiment (but not affective ones) due to time constraints and practical considerations of conducting an online intervention.

**Indefinite Article Block.** The centrepiece of the first block was a set of tasks involving low-frequency English words that learners would not be expected to know. Figure 18 shows an example.

**Figure 18**

*Example Item from the Structured Input Activity (Indefinite Contexts)*

Sludge was used to fertilise this soil.

*You could use this sentence to talk about... (choose all possible answers)*

|                          |                                   |                                   |                                 |
|--------------------------|-----------------------------------|-----------------------------------|---------------------------------|
| ...one bag of sludge     | ...more than one bag<br>of sludge | ...less than one bag<br>of sludge | ...part of one bag of<br>sweets |
| <input type="checkbox"/> | <input type="checkbox"/>          | <input type="checkbox"/>          | <input type="checkbox"/>        |

In this example, learners are driven to realise that the absence of an article indicates the indefinite quantity of the item and choose three options (“one bag of sludge”, “more than one bag of sludge”, and “less than one bag of sludge”). Low-frequency vocabulary was used to ensure that learners could not rely on word meaning to derive countability information. There is always a distractor option (“part of one bag of sweets”), and the order in which options are presented is randomised.

The target items were varied along the following parameters:

- Noun type: count singular, e.g. “moat”, count plural, e.g. “beacons”, or mass, e.g. “sludge”
- Abstractness of the noun: abstract, e.g. “duress”, or concrete, e.g. “burrow”
- Modification: modified by a preceding adjective, e.g. “professional demeanour”, or not
- Specificity: specific (56) or non-specific (57), i.e. referring or not referring to a certain existing entity, except for nominals denoting a property in predicate position, which were always specific<sup>23</sup> (58).

<sup>23</sup> It is recognised that predicates, as in (58), denote properties, which are not specific per se; however, I considered such examples specific for the purposes of the intervention because they describe a specific referent. Clearly, examples can be constructed where predicates describe non-specific referents, e.g. “anyone can be a craven”, but these were not included in the intervention materials.



(56) There is *a huge boulder* next to the road.

(57) I can't stand *tantrums* in public.

(58) He is *a real craven*.

- Syntactic position: subject (59), object (57), predicate (58), existential (56)

(59) *A wide moat* protects the castle.

- Semantic transparency of the verbs (for indefinite nominals in object position).

Semantically transparent contexts (59) were unambiguous as to whether the referent was specific or non-specific and included verbs such as “buy”, “spend”, “read”, “wear” etc. On the contrary, semantically opaque contexts (57) were ambiguous in this respect and included verbs such as “need”, “want”, “look for”, “will”, as well as verbs in negative, interrogative, and conditional constructions (60).

(60) If you have *a sudden qualm*, call me.

Inanimate nouns were used for all items except for some nominals in predicate positions (58).

At first, learners practised with high-frequency words to familiarise themselves with the type of activity and then moved on to activities with low-frequency words. Low-frequency words were chosen based on Paul Nation’s BNC/COCA headword lists (Nation, 2017) as words that do not appear on the first 4 band lists (i.e. not within 4000 most frequent words). I also avoided using words which were cognates with Romance, Germanic, and Slavic languages, for the same reason that I employed low-frequency words, i.e. to ensure that learners cannot rely on word meaning to derive countability.

At the end of the first block learners were introduced to some examples of ambiguous nouns in English, which could be used in both count and mass interpretations (e.g. “cake”, “glass”, “charity”), and were given similar practice activities with those (Figure 19). One reason for including this type of items was to demonstrate to learners that countability is not necessarily a fixed feature of a noun and that using “a” or Ø in front of certain nouns can make them countable or uncountable,

respectively. Another reason was ecological validity, as it would have been misleading to present only unambiguous nouns while ambiguous cases are so common in English.

For ambiguous items prenominal modification was not a parameter but was controlled, so that modified and non-modified items were equally distributed in each set. In addition, ambiguous nouns in non-specific contexts were not presented in the subject or predicate syntactic positions as it was not possible to construct natural sets of examples for those conditions<sup>24</sup>.

### Figure 19

*Example Item from the Structured Input Activity with an Ambiguous Noun*

**There is a paper on the table.**

You could use this sentence to talk about...

*(Choose all possible answers)*

- ☐ ...one newspaper
- ☐ ...more than one newspaper
- ☐ ...some sheets of blank paper
- ☐ ...one black hole

This resulted in 258 items. The block was split into 10 sets of 12, 18, or 24 items presented in the order shown in Figure 20. As seen in the figure, several sets of specific items precede the introduction of non-specific ones. This is based on the corpus study results, which demonstrated that learners were less prone to omitting “a” with specific count singular indefinites. In addition, items in existential position, which were of the highest accuracy in the corpus study, were introduced first, along with object position, whereas items in subject position, which had the lowest accuracy in the corpus study, were introduced later. Items in predicate positions were also introduced later, despite the generally high article accuracy for these in the corpus study, mainly

---

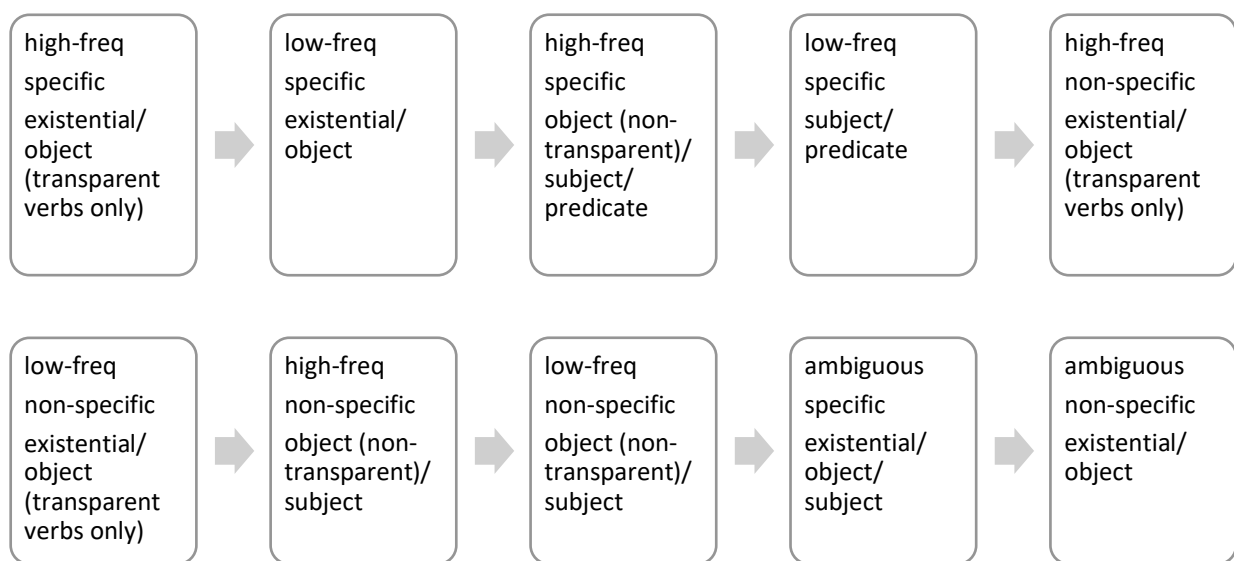
<sup>24</sup> For example, “a watermelon contains a lot of water” does not necessarily refer to one watermelon and, thus, does not contrast with “watermelon contains a lot of water”, as would be required for the activity to work. In object and existential positions, the singular reading is more preferred, e.g. “I need a glass”, so this could contrast with “I need glass”, as expected. This was also not an issue for non-specific unambiguous nouns in subject position, as there was more freedom in word choice, so I could select those words that would be more compatible with a singular reading, e.g. “A laptop is necessary for each participant”.

because the indefinite article in such cases has a purely structural function of signalling a noun phrase. The fact that the intervention is imposing a form-meaning mapping in this context (i.e. “He is a doctor” should be interpreted as a statement about one doctor, not multiple doctors or an unidentified amount) mainly serves a pedagogical purpose, exploiting learners’ propensity to attach meaning to forms, as discussed in the previous chapter.

Each set of items took 10-15 minutes to complete.

**Figure 20**

*Order of Item Presentation for Pilot Experimental Activities in the Indefinite Block (Sets 1-10)*



**Definite vs. Indefinite Article Block.** In the second block learners were given a sentence or a phrase containing the target nominal and asked to match it with the most appropriate context, as in the example in Figure 21.

**Figure 21**

*Example Item from the Structured Input Activity (Definite vs. Indefinite Contexts)*

*In which dialogue does the purple sentence fit best?*

- I'm looking for the new information. - I'm looking for the new information. - I'm looking for the new information.
- Check the papers on my desk. It should be there. - I see. What sort of information are you looking for? - No problem. I can look after your dog while you're gone.



One of the three dialogues was always a nonsense distractor. The other two represented a definite and an indefinite context, so the learner would have to rely on the article to choose the appropriate one. The items represented specific definite contexts only to limit the scope of the intervention and avoid overloading participants, as non-specific definite contexts are relatively rare and would have been more difficult to construct. The indefinite counterparts, however, could be non-specific, especially where non-transparent verbs were used, e.g. the indefinite counterpart of the example in Figure 21, “I’m looking for *new information*. – I see. What sort of information are you looking for?”, is non-specific.

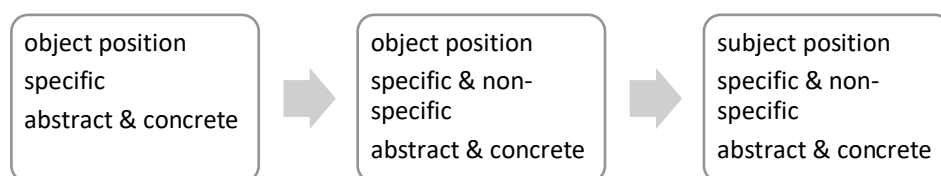
The target items in this block varied along the following parameters:

- Definiteness (definite/indefinite)
- Noun type (count singular/count plural/mass)
- Abstractness of the noun (abstract/concrete)
- Modification (modified by a preceding adjective or not)
- Syntactic position (subject/object)
- Semantic transparency of the verbs (for nominals in object position)

This resulted in 72 items. The block was split into three sets of 24 items (Figure 22), each taking about 10-15 minutes to complete.

**Figure 22**

*Order of Item Presentation for Pilot Experimental Activities in the Definite/Indefinite Block (Sets 11-13)*



The factorial structure of items in the Experimental materials is provided in Appendix D.

**Control Group Materials.** Control group participants were given reading comprehension and vocabulary activities that would roughly take the same amount of time as the Experimental activities. The tasks were based on online lessons from the Breaking News English website

(breakingnewsenglish.com), where each lesson contains a real-world news article adapted to several English learning levels followed by classroom activities. Half of the texts were taken from level 3 and half from level 5 (out of 6 levels) to ensure there would be some easier and some more challenging reading. The texts were used exactly as they appear on the website with explicit permission from the author and creator of the website. I made changes to the reading comprehension and vocabulary tasks drawing on my own teaching experience to fit the timing allocated per activity and the online format of the platform where the experiment was set up. An example of a Control group activity is provided in Appendix E.

### **Assessment**

Participants were asked to take a pre-test before starting the activities, an immediate post-test, and a delayed post-test 3 months after the immediate post-test.

The tests consisted of three tasks: grammaticality judgment task (GJT), elicited imitation (EI), and free written production.

**GJT.** The GJT contained 44 target items and 22 distractors each with a grammatical (correct, C) and an ungrammatical (incorrect, IC) version. Whether the correct or the incorrect version of each item was displayed to a participant was randomly determined by the programme. Target items were varied according to the following factors:

- Definiteness: definite (61) or indefinite (62)

(61) John is in hospital. - I know, I saw *the* (\*a)<sup>25</sup> car crash.

(62) I like this band. I want to buy *an* (\*the) album, but I haven't decided which one yet.

- Specificity: specific (63) or non-specific (64), i.e. referring or not referring to a certain existing entity (this was varied only for indefinite contexts)

(63) (\*A) Cups are in the cupboard.

(64) There are (\*an) interesting events every week.

---

<sup>25</sup> The option with an asterisk in parentheses represents the ungrammatical (IC) version of the item.

- Noun type: count singular (61), count plural (63), or mass (65)

(65) This is (*\*an*) *animal fat*.

- Abstractness of the noun: concrete (63) or abstract (64)

The items were also controlled for:

- Modification: modified by a preceding adjective or noun adjunct (65) or not
- Syntactic position: subject (63) or object (62) for definites and indefinites, as well as predicate (65) or existential (66) for indefinites only

(66) There may be *a* (*\*Ø*) *club* in this area.

- Semantic transparency of the verb: transparent (61) or opaque (66)

Distractors targeted a variety of typical learner errors, such as the use of tenses, prepositions, word order etc.

**Elicited Imitation.** In the EI task participants heard a pre-recorded passage or short dialogue containing 2-3 target nominals (12-13 passages containing 36 total targets per test). In (67) the target nominals are italicised.

(67) There is *a concert* tonight. We can get *an extra ticket*. Do you want to come? – Sorry, there are *important tasks* I need to finish.

After listening to each passage played twice, participants had to type what they had heard as closely to the original as possible. The idea behind this task, which is based on the task used for a similar purpose in VanPatten & Uludag (2011), is that the passages are deliberately made too long to be held in one's short-term memory (19-31 words). Unable to remember the passages word for word, participants would need to reconstruct them, which is hoped to elicit productive use of articles, while partly controlling participants' choice of target items. Participants had a time limit between 70-90 seconds (depending on the length of the passage) to type their responses.

The passages were recorded by native speakers of British English at normal speed without placing any emphasis on the target nominals. The target nominals varied along the same factors as the items in the GJT.

The target nouns for both GJT and EI were all within the 3000 most frequent words in English according to Paul Nation's BNC/COCA headword list (Nation, 2017), and no word was used more than once to avoid any practice effects.

**Free Writing.** In the free written production task participants were asked to write about 100 words in response to a given prompt within a time limit of 10 minutes. The prompts included topics of general interest, such as describing one's hometown or expressing an opinion about modern technology.

The order of tasks and the items within them was randomised for each participant by the programme. I created three equally balanced versions of the test (A, B, C), so no test item was seen more than once by each participant. I used Latin square (split block) design to allocate participants to test versions.

The design structure of the tests is detailed in Appendix F along with example items.

### ***Participants***

Participants for the pilot were recruited through the English First language school in Cambridge. The school welcomes students with a variety of L1 backgrounds to join their immersion language programmes ranging from a couple of weeks to several months in length.

After the experiment was approved by the Research Ethics Committee of the Faculty of Modern and Medieval Languages and Linguistics of the University of Cambridge, the administration of the school spread the information about this experiment among their students and teachers. They also ensured that parental consent was obtained for any potential participants under 18 years of age before they received any registration links.

I visited the school twice to give a short presentation about the study to interest potential participants. Those who chose to participate were offered a certificate of participation upon completion of the immediate post-test.

As a result, 50 students registered their interest, of whom 28 went on to do the pre-test, all the activities and the immediate post-test. It is the data from these 28 students that was included in

the analysis below. Six of them also completed a delayed post-test about three months after the immediate post-test. Note that only 3 participants in the Control group came from a [-art] L1 background, while the proportion of such learners was slightly higher in the Experimental group (5 out of 13). Further details about the participants are given in Table 12.

**Table 12**

*Participants of the Pilot Experiment*

|   | <b>Experimental group</b>  | <b>Control group</b>  |
|---|--|---|
| Number of participants                              | 13   | 15  |
| Age   | Mean = 24.8, range 17-37   | Mean = 23.6, range 17-32  |
| Gender distribution                                 | 9 female, 4 male   | 13 female, 2 male   |
| English level (according to the school's placement) | B1 – 1<br>B2 – 4<br>C1 – 8   | B1 – 2<br>B2 – 8<br>C1 – 5  |
| L1  | [+art]<br>Spanish – 3<br>German, French – 2 each<br>Dutch – 1<br><br>[-art]<br>Chinese – 4<br>Japanese – 1 | [+art]<br>Spanish – 5<br>Italian – 3<br>German – 2<br>French, Dutch – 1 each<br>[-art]<br>Russian, Chinese, Thai – 1 each |

**Procedure**

The entire experiment, from filling out consent forms to delayed post-test, happened online using Qualtrics software (*Qualtrics*, 2020). All the treatment activities were accessible from any mobile device, although the tests could only be done on a laptop or desktop computer due to the technical limitations of the software.

Upon registering their interest and signing the consent form, participants filled out the personal and language background questionnaire, which asked them about their age, gender, level of education, L1, information about their L2/L3 etc. together with the age of onset of each language and self-assessed proficiency level, information about extended residence in English-speaking countries (over 3 months) and about English learning contexts (classroom, tutoring, self-study etc.). The consent form and full background questionnaire is provided in Appendix G.



Then the programme randomly allocated participants to either the Experimental or the Control group and immediately sent out the link to the pre-test (randomly selected from the three versions, A/B/C). After a participant completed the pre-test, which took around 45-60 minutes, they would start receiving links to the relevant activities, no more than one per day. The participants were encouraged with e-mail reminders to finish the whole experiment within 3 weeks (apart from the delayed post-test), but about half of them took longer than that ( $M = 28$  days, range: 15-57 days). There were 13 activities for both Experimental and Control groups each taking about 10-15 minutes. Thus, the total time spent on Experimental or Control treatment was roughly between 2.5-3.5 hours.

On the following day after finishing the last activity, each participant received the link to the immediate post-test. The delayed post-test could be accessed three months after completing the immediate post-test.

In addition, participants received an anonymous feedback survey link after completing the immediate post-test, which was used to improve the materials for further data collection.

### ***Coding***

On the GJT participants received 1 point for correctly identifying whether the item is grammatically correct or contains an error. They received 0 points if they made a mistake or chose "I'm not sure".

The elicited imitation task was coded manually by the researcher. First, I identified whether the three target items were reproduced by each participant. If an item was reproduced but in a different configuration of factors, for example with a modifier added or in a different syntactic position, this was marked accordingly. In case a target item was replaced by a synonym or a different word which still made sense, this was counted as a reproduced target. Any other produced noun phrases, apart from targets, were not considered for analysis. Where it appeared that a participant did not understand the prompt, this was not included in the analysis even if they produced some words resembling the intended targets. Once the reproduced targets were identified, they were

coded for correctness and error types (omission of “a”/“the”, overuse of “a”/“the”, substitution of “a” with “the” or vice versa).

In the free production task, all the nominals were identified manually and then coded for the type of target (“a”/“the”/∅), correctness and error types. I excluded from the analysis proper nouns and nouns that were preceded with the following determiners: “some”, “any”, “no”, “much”, “how much”, “how many”, “most”, “more”, “(a) few”, “(a) little”, “another”, “each”, “every”, “such”, “this”, “that”, “these”, “those”, “what”, “which”, possessive pronouns (“his”, “their”) or nouns (“teacher’s”, “Tom’s”).

### **Analysis and Results**

The purpose of the pilot was to see whether the treatment and assessment materials work and to identify any potential issues with data collection and coding/scoring. The pilot data was also used to derive approximate estimates of the effect of the intervention, which were fed into the power analysis conducted to determine the optimal sample size for the Main Intervention Study reported in Chapter 6.

Keeping these goals in mind and considering that the dataset collected during the pilot was relatively small, I conducted the analysis based on results aggregated by participant using linear regression modelling rather than considering each test item by each participant as a separate data point and using mixed-effects regression modelling. The latter approach would be preferable for a more detailed and robust analysis and is used in the Main Study (Chapter 6).

### ***Comparing Groups at Pre-Test***

First, I checked whether the different versions of the test (A/B/C) were not significantly different from each other in terms of difficulty. A linear regression model with participant pre-test scores as the outcome variable and test version as the predictor variable estimated that test version did not have a significant impact on participant scores. Using ANOVA to compare the models with and without the test version predictor, I found no statistically significant effect on the GJT scores

( $F(2,25) = 0.02, p = 0.98$ ), the EI scores ( $F(2,25) = 1.33, p = 0.28$ ), or the free writing scores ( $F(2,25) = 1.19, p = 0.32$ ).

Second, I modelled participants' scores at pre-test as a function of group (Control or Experimental) and found no statistically significant effect for the GJT ( $t(26) = 0.54, p = 0.59$ ), EI ( $t(26) = -0.51, p = 0.62$ ), or free writing ( $t(26) = 1.3, p = 0.20$ ). This means that the groups did not perform significantly differently from each other at the pre-test.

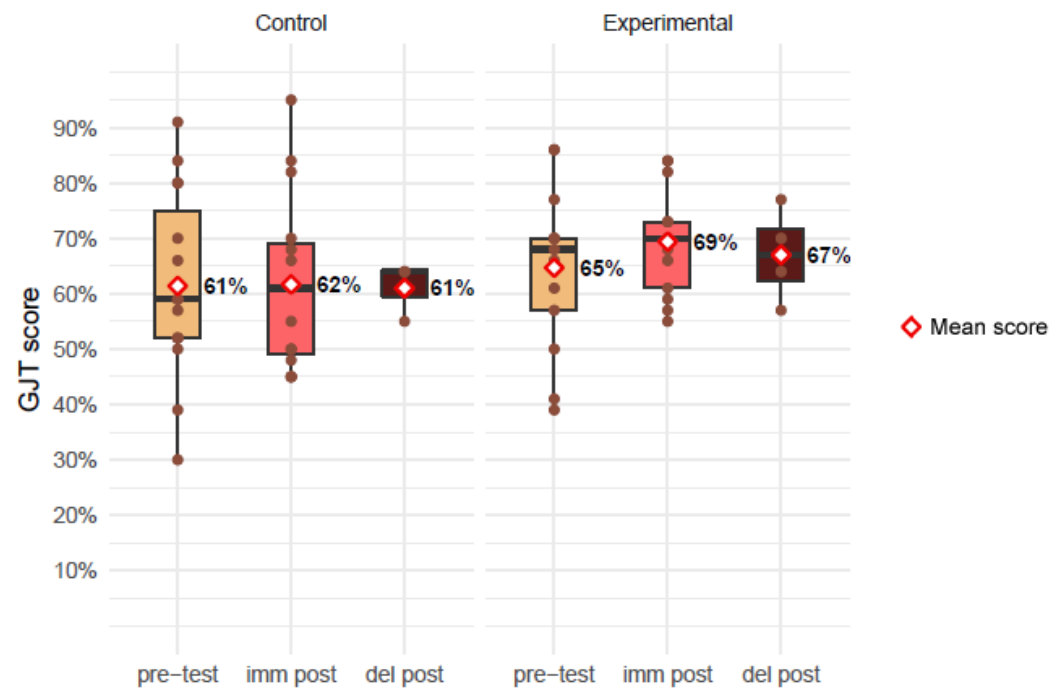
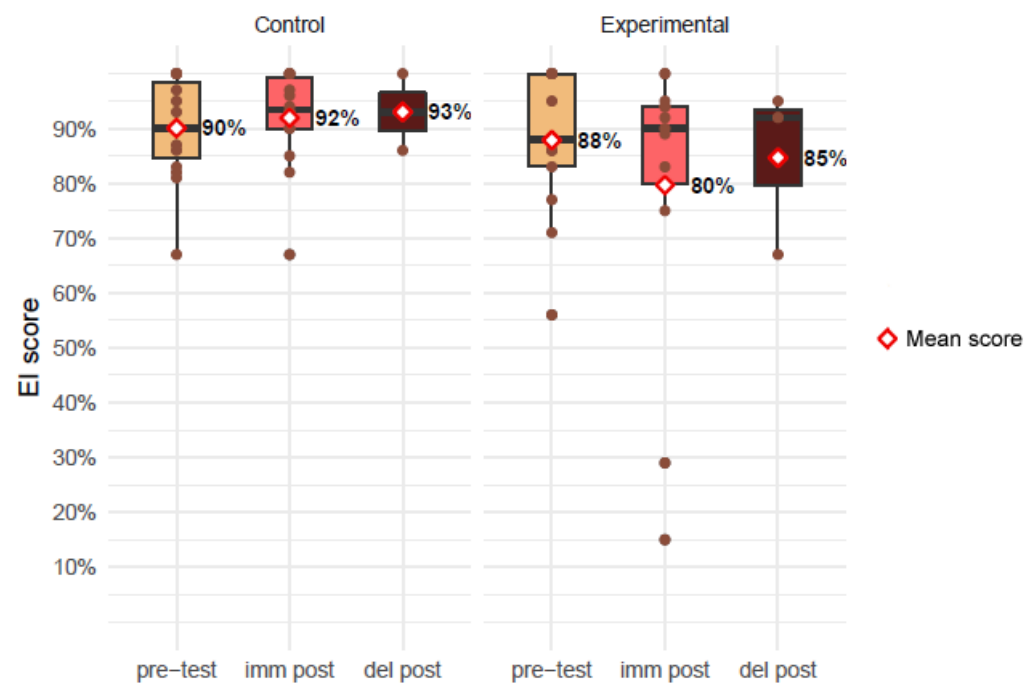
### ***Descriptive Statistics***

The overall scores for the pre-tests, immediate post-tests and delayed post-tests by group are provided in Figure 23 (GJT), Figure 24 (EI), and Figure 25 (free writing). While Figure 23 suggests some improvement for the Experimental group on the GJT, no such tendency can be observed for EI or the free writing task. Moreover, the scores for the latter two tasks are relatively high for both groups leaving little room for improvement.

### ***The Effect of the Intervention on the GJT***

To check whether the intervention had a significant positive effect on the Experimental group's performance, I ran a multiple linear regression model, where the outcome variable was the change in participants' accuracy scores from pre-test to immediate post-test (e.g. if a participant scored 50% at pre-test and 55% at immediate post-test, their score change was 5%).

First, I fitted a model with the following predictors: group (dummy coded as Control – 0, Experimental – 1), L1-type (dummy coded as [+art] – 0, [-art] – 1), proficiency level (coded as a three-level variable: B1 – 0, B2 – 1, C1 – 2), and pre-test GJT score (continuous variable). None of the predictors were significant apart from the pre-test GJT score, which is not surprising, as lower pre-test scores allow more room for improvement than the higher ones.

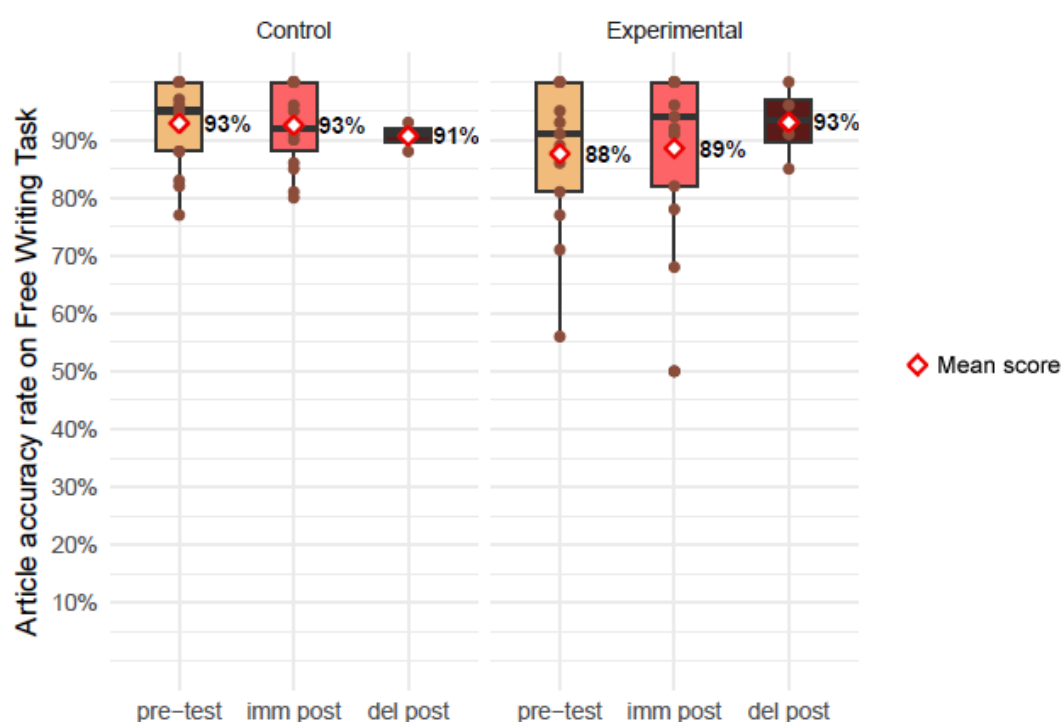
**Figure 23***Pilot GJT Results***Figure 24***Pilot EI Results*

I then attempted two-way interactions between group and each of the other variables to see whether the intervention had a differential effect on participants with different L1s, proficiency levels, and pre-test scores (i.e. starting levels).

Using ANOVA to compare several possible multiple regression models (those which had a statistically significant  $F$ -value), I identified the best fitting model as the one in which group and pre-test score interact to predict the change in the GJT scores from pre-test to immediate post-test and in which the effect of the pre-test score is curvilinear (parabolic). This model explains 54% of the variance in the data (adjusted  $R^2 = 0.54$ ,  $F(4,23) = 8.99$ ,  $p < 0.001$ ) and is a better fit than models including L1-type and/or proficiency level, as well as a model without the interaction or without the parabolic effect. The results of the model are presented in Table 13.

**Figure 25**

*Pilot Free Writing Results*



**Table 13**

*Pilot GJT Model Results*

| Predictors  | GJT score change |            |             |       |
|-------------|------------------|------------|-------------|-------|
|             | Estimates        | std. Error | 95% CI      | p     |
| (Intercept) | 0.73 **          | 0.20       | 0.32 – 1.13 | 0.001 |

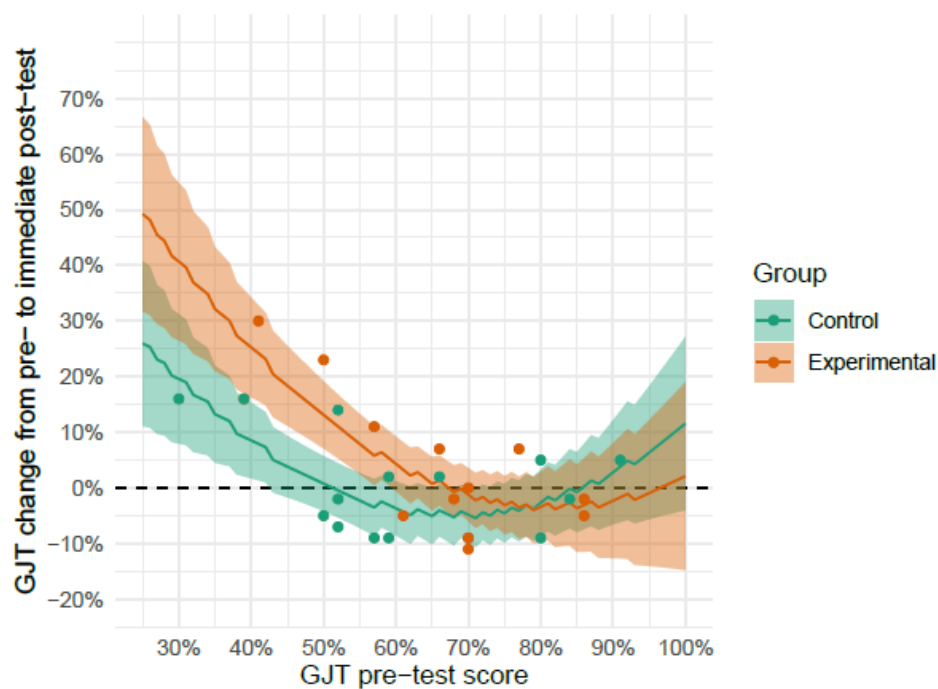
|  |   |      |               |       |
|--|---|------|---------------|-------|
| group: Experimental                      | 0.34 **   | 0.12 | 0.10 – 0.59   | 0.009 |
| GJT pre-test score                       | -2.26 **  | 0.64 | -3.59 – -0.94 | 0.002 |
| (GJT pre-test score)^2                   | 1.65 **   | 0.51 | 0.61 – 2.70   | 0.003 |
| group: Experimental * GJT pre-test score | -0.44 *   | 0.18 | -0.81 – -0.06 | 0.025 |
| Observations                             | 28  |      |               |       |
| R <sup>2</sup> / Adjusted R <sup>2</sup> | 0.61 / 0.54   |      |               |       |
| ANOVA comparing to the null model        | $F(4, 23) = 8.99, p = 0.00016$  |      |               |       |
| Variance inflation factors (VIFs)        | Strong in a model with correlations but low in a no-correlation model (< 1.5 for all variables) |      |               |       |

Note. \*  $p < 0.05$  \*\*  $p < 0.01$  \*\*\*  $p < 0.001$

All the regression coefficients in this final model are statistically significant, including both the linear and the quadratic (parabolic) terms (“(GJT pre-test score)^2” in Table 13). This means that the regression line is curved and declining as seen in Figure 26.

**Figure 26**

*Pilot GJT Score Change as a Function of Pre-Test Score and Group*



Note. Points on the graph represent observed scores, while lines represent scores predicted by the model, with shaded areas showing 95% confidence intervals.

The first important observation about this model is that there is not much change happening for participants with pre-test scores around 65% and higher. This is also the point where the two groups become almost indistinguishable<sup>26</sup>.

Crucially, the difference between the groups is most obvious for learners who score below 65% on the pre-test, which means that those learners are the most likely to benefit from the intervention. This is in line with the 60% threshold for accepting participants into the intervention used by VanPatten and his colleagues in early Processing Instruction experiments (Cadierno, 1995; Cheng, 1995; VanPatten & Sanz, 1995).

As noted above, I found no significant effect of L1-type or proficiency level (as determined by EF school's placement). This is perhaps, not surprising given that [-art] L1 learners were underrepresented in the pilot dataset (8 out of 28), particularly in the Control group (only 3 out of 15). As for the proficiency level, we were only able to collect information on broad proficiency level bands, and most of the participants were within B2 or C1 levels, with only 3 in B1. The pre-test score predictor captured proficiency differences more precisely (and on a continuous scale), which rendered the broad proficiency band relatively less predictive.

### ***The Effect of the Intervention on EI Results***

Using ANOVA to compare linear regression models, I identified that the best predictor of the change in scores on EI is only the pre-test score ( $R^2 = 0.29$ ,  $F(1, 26) = 12.1$ ,  $p = 0.002$ ). Adding the group variable to the equation did not improve the model, which simply reflects the fact that lower pre-test scores leave more room for improvement. Considering that pre-test scores were generally quite high (i.e. there were probably ceiling effects), it is not surprising that the treatment did not have an effect.

In addition, this task may have been particularly difficult as it involved listening comprehension, and the materials were recorded by native speakers without much adaptation to

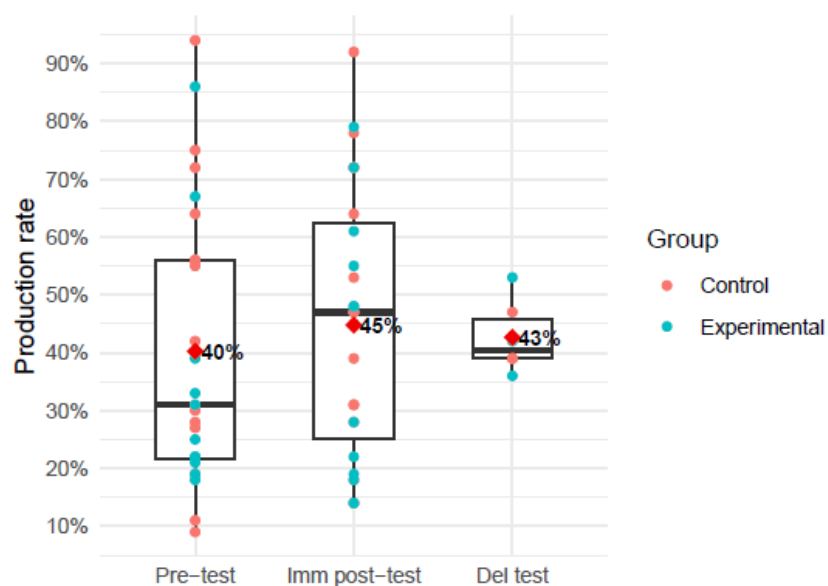
---

<sup>26</sup> The fact that the curves go slightly upwards at the end of the x-axis is meaningless, as clearly learners' scores would not be expected to improve much if their pre-test results were close to 100%.

learners. Quite a few participants reported in the anonymous feedback survey that this task was hard for them because they could not understand or remember what they had heard. As a result, the percentage of intended targets actually produced by learners was generally very low (average 40% at pre-test, as seen in Figure 27). Thus, a participant produced on average only 14 target items out of the intended 36. Within those items that were produced the article accuracy rate was very high, but it is possible that the items they could not reproduce would have been more difficult.

**Figure 27**

*Rate of Intended Target Production on the Pilot EI Task*



### ***The Effect of the Intervention on the Free Writing Task***

Proficiency level, pre-test score and L1-type were all significant predictors of article accuracy on the free writing task in the best fitting model. However, adding group as a predictor did not improve the model, which means that the intervention did not affect article accuracy on this task.

It is important to note that the pre-test score is already very high for most participants (over 80% for 24 out of 28 participants, over 90% for 17 out of 28 participants). However, this overall high accuracy rate masks the fact that over half of all obligatory contexts produced were target  $\emptyset$  contexts (i.e. indefinite mass or plural nominals), in which we cannot distinguish between genuine mastery and coincidental accuracy due to article omission. Considering the low numbers of obligatory contexts in general, especially for those categories where one would expect the most



potential for improvement (i.e. [-art] Experimental group obligatory contexts for “a” and “the” – highlighted cells in Table 14), it is not surprising that no statistically significant effects were found.

**Table 14**

*Numbers of Obligatory Contexts Produced in Pilot Free Writing Task*

| Group                           | Target “the” |           |           | Target “a” |           |           | Target Ø   |            |           |
|---------------------------------|--------------|-----------|-----------|------------|-----------|-----------|------------|------------|-----------|
|                                 | Pre          | Imm       | Del       | Pre        | Imm       | Del       | Pre        | Imm        | Del       |
| <b>Control group total</b>      | <b>81</b>    | <b>90</b> | <b>26</b> | <b>46</b>  | <b>30</b> | <b>11</b> | <b>153</b> | <b>122</b> | <b>25</b> |
| [+art]                          | 69           | 71        | 26        | 41         | 21        | 11        | 120        | 106        | 25        |
| [-art]                          | 12           | 19        | -         | 5          | 9         | -         | 33         | 16         | -         |
| <b>Experimental group total</b> | <b>52</b>    | <b>61</b> | <b>30</b> | <b>34</b>  | <b>42</b> | <b>9</b>  | <b>102</b> | <b>143</b> | <b>35</b> |
| [+art]                          | 34           | 47        | 17        | 22         | 30        | 4         | 71         | 114        | 18        |
| [-art]                          | 18           | 14        | 13        | 12         | 12        | 5         | 31         | 29         | 17        |

*Note.* Pre – pre-test, Imm – immediate post-test, Del – delayed post-test

### Discussion and Issues Identified in the Pilot

The pilot of the intervention demonstrated some potential for Processing Instruction to improve accuracy on the GJT for learners at lower levels (those scoring below 65% on the pre-test). However, it is clear that the ability to make any further conclusions based on this data is limited due to a number of problems. The purpose of the pilot was to reveal such problems, so that they could be addressed for further data collection.

### Sample Size

The most evident issue in the analysis of the pilot data is the small sample size. This is particularly important for exploring the impact of specific variables, which would require splitting data into numerous subcategories, each of which would need to contain a sufficient number of data points (e.g. Table 14). If the aim is to investigate the effect of multiple features of the nominal (e.g. definiteness, number, countability) in addition to several learner characteristics (e.g. L1-type and proficiency level), a considerably larger dataset is needed (as was the case for the corpus study reported in Chapter 3).

### Participants’ Proficiency Level

The results of the GJT show that only those participants scoring below 65% on the pre-test are likely to benefit from the intervention. However, this was the case for only about half of the

participants in the pilot. The accuracy rates on the EI and Free Writing tasks also appeared rather high leaving little room for improvement. Thus, it probably makes more sense to focus on recruiting more lower-level learners (e.g. A2/B1/B2 rather than predominantly B2/C1 levels, as in the pilot).

### ***Assessment***

**GJT.** The participants of the pilot appeared to have the lowest scores on the GJT, and this was the only task where the Experimental group improved after the intervention in comparison to the Control group. However, GJT is arguably metalinguistic, especially when it is untimed.

Another issue identified during test design is the difficulty of creating certain types of ungrammatical items (e.g. making an indefinite interpretation of a context unequivocally unacceptable). Although learners appear to perform as expected on the GJT, it is necessary to test these items with native speakers to ensure reliability.

**EI.** The biggest issue with this task was the fact that learners were unable to reproduce more than half of the intended target nominals, which severely limited the ability to conduct any kind of analysis. The native speaker recordings might have been too fast-paced or unclear, as multiple participants reported in the anonymous feedback forms that they were unable to understand the prompts, although most were placed in B2 or C1 proficiency levels by their language school. Moreover, the density of intended targets (3 per prompt) might have been excessive.

**Free Writing.** In this task the biggest problem was a low number of obligatory contexts, especially of non- $\emptyset$  target contexts.

### ***Experimental Activities***

One important issue with the experimental materials was that participants claimed familiarity with some of the low-frequency vocabulary that was intended as “unknown” for the purposes of the intervention<sup>27</sup>. Perhaps, setting the threshold above 4000 most frequent words was not high enough, especially for higher-level learners.

---

<sup>27</sup> Each activity with low-frequency vocabulary included feedback questions at the end asking participants whether they had known the words used in the activity.

Another potential problem is the length of treatment. It is estimated that participants spent between 2.16-3.25 hours total over a period of a few weeks on the practice activities. Only half of the items on those activities contained low-frequency words, which were supposed to be the most effective in fixing any article processing issues. It is questionable whether one could expect much improvement after such a brief intervention.

## Chapter 6: Processing Instruction Intervention Main Study

### Methodology

#### *Power Analysis*

To determine the optimal sample size for the study, I conducted a simulation-based power analysis. First, I created a large dataset with the same structure as the one I was expecting to collect in the experiment. The dataset was structured as follows:

- 3 experimental groups: Control group and Experimental group as in the pilot, additional Control group (Control1) completing the types of article practice activities typically found in mainstream textbooks for teaching English as a foreign language<sup>28</sup>.
- 500 participants in each group (1500 total)
- 3 tests of 44 target items each completed by each participant (pre-test, immediate post-test, delayed post-test) using Latin squares design, so that items were not repeated across tests (132 items in total per participant)

As the data set was designed for a mixed-effects generalised linear model, each item constituted a separate data point, i.e. a separate row, resulting in 198000 rows of data (1500 participants \* 132 items).

Each item had a score of either 0 or 1 (incorrect or correct response by the participant).

I used the results of the pilot to make approximate estimates of the variable effects, making the adjustment for the expectation to recruit more lower-level participants who might benefit more from the intervention. The mean score for each participant at pre-test was set at 55% (the overall mean score in the pilot was 63%).

The mean scores at immediate post-test were set as follows:

- Control group 0 – 55% (i.e. no change, the mean change in the pilot was +1%)
- Control group 1 – 60% (i.e. +5%)

---

<sup>28</sup> The idea of a second Control group had to be discarded in the end, because recruiting enough participants even just for two groups during the pandemic became a much more challenging endeavour.

- Experimental group – 65% (i.e. +10%, the mean change in the pilot was +4%)

The mean scores at delayed post-test were set as follows:

- Control group 0 – 55% (no change, same as in the pilot)
- Control group 1 – 58% (+3% from the pre-test)
- Experimental group – 62% (+7% from the pre-test, the mean change in the pilot was +2%)

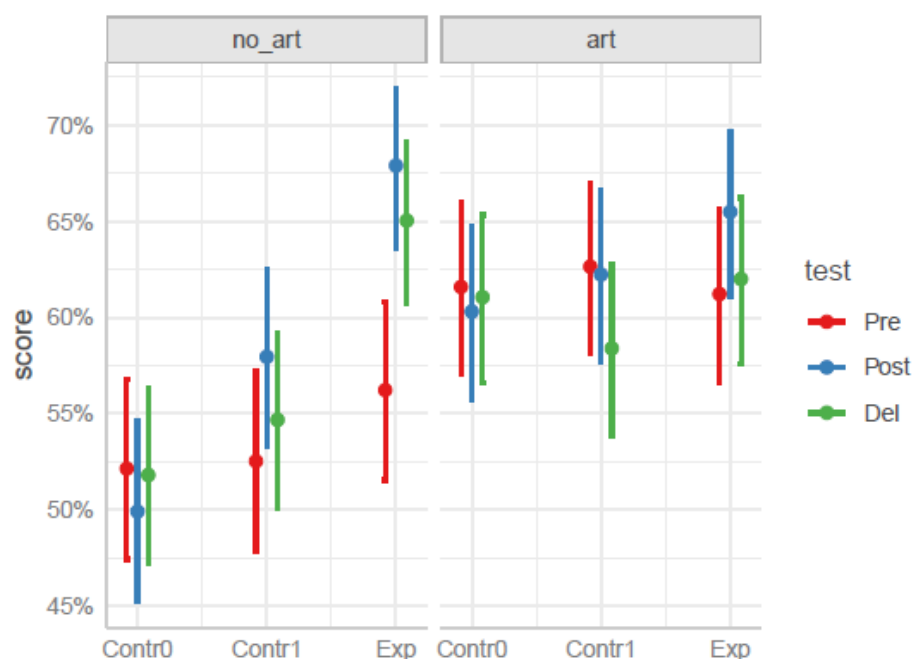
To introduce random variation for the mixed model, three randomly generated numbers were added to each individual data point:

- Participant varying intercept: randomly generated 1500 numbers (one for each participant), normally distributed,  $M = 0$ ,  $SD = 0.15$  log odds
- Item varying intercept: randomly generated 132 numbers (one for each item), normally distributed,  $M = 0$ ,  $SD = 0.05$  log odds

The second step was fitting a rough version of a generalised linear mixed-effects logistic regression model (similar to the one that could be fitted on the data collected in the experiment) on this artificially created dataset. The score on each test item was the outcome variable in the model (binary: 0 for incorrect response or 1 for correct response). The model included three predictors, namely group (Control 0, Control 1, Experimental), L1-type ([+art] or [-art]), and test (pre-test, immediate post-test, delayed post-test), as well as a 3-way interaction between them. The model also included random intercepts by participant and random slopes for the test variable by participant, as well as random intercepts by item. The effect of the intervention predicted by the model is illustrated in Figure 28 (note that the scale starts at 45%). The results of this model are given in Table 15. I took the estimate for the 3-way interaction term “test: immediate post-test \* L1-type: [+art] \* group: Experimental” (i.e. -0.35 log odds) as the basis for power analysis simulations.

**Figure 28**

*The Effect of the Intervention Predicted by the Simulated Dataset Model (with 95% Confidence Intervals)*

**Table 15**

*Simulated Intervention Dataset Model Results*

| <i>Predictors</i>                                       | <b>Accuracy Rate</b> |                   |               |          |
|---|----------------------|-------------------|---------------|----------|
|   | <i>Log-Odds</i>      | <i>std. Error</i> | <i>95% CI</i> | <i>p</i> |
| (Intercept)   | 0.09                 | 0.10              | -0.10   0.27  | 0.374    |
| test: immediate post                                    | -0.09                | 0.09              | -0.27   0.09  | 0.323    |
| test: delayed post                                      | -0.01                | 0.09              | -0.19   0.16  | 0.883    |
| L1: [+art]  | 0.39 ***             | 0.12              | 0.16   0.62   | 0.001    |
| group: Control 1  | 0.02                 | 0.12              | -0.21   0.24  | 0.895    |
| group: Experimental                                     | 0.16                 | 0.12              | -0.06   0.39  | 0.159    |
| test: immediate post * L1: [+art]                       | 0.04                 | 0.08              | -0.12   0.20  | 0.665    |
| test: delayed post * L1: [+art]                         | -0.01                | 0.08              | -0.17   0.15  | 0.909    |
| test: immediate post * group: Control 1                 | 0.31 ***             | 0.08              | 0.15   0.47   | <0.001   |
| test: delayed post * group: Control 1                   | 0.10                 | 0.08              | -0.06   0.26  | 0.218    |
| test: immediate post * group: Experimental              | 0.59 ***             | 0.08              | 0.43   0.75   | <0.001   |
| test: delayed post * group: Experimental                | 0.38 ***             | 0.08              | 0.22   0.54   | <0.001   |
| L1: [+art] * group: Control 1                           | 0.03                 | 0.17              | -0.30   0.35  | 0.859    |
| L1: [+art] * group: Experimental                        | -0.18                | 0.17              | -0.50   0.14  | 0.276    |
| test: immediate post * L1: [+art] * group: Control 1    | -0.27 *              | 0.12              | -0.50   -0.05 | 0.019    |
| test: delayed post * L1: [+art] * group: Control 1      | -0.26 *              | 0.12              | -0.48   -0.03 | 0.027    |
| test: immediate post * L1: [+art] * group: Experimental | -0.35 **             | 0.12              | -0.58   -0.12 | 0.003    |
| test: delayed post * L1: [+art] * group: Experimental   | -0.33 **             | 0.12              | -0.55   -0.10 | 0.004    |
| <b>Random Effects</b>                                   |                      |                   |               |          |
| <i>By participant</i>                                   |                      |                   |               |          |

| <i>Predictors</i>         | <b>Accuracy Rate</b> |                   |               |          |
|---------------------------|----------------------|-------------------|---------------|----------|
|                           | <i>Log-Odds</i>      | <i>std. Error</i> | <i>95% CI</i> | <i>p</i> |
| SD (Intercept)            | 0.72                 |                   |               |          |
| SD (test: immediate post) | 0.08                 |                   |               |          |
| SD (test: delayed post)   | 0.07                 |                   |               |          |
| <i>By item</i>            |                      |                   |               |          |
| SD (Intercept)            | 0.26                 |                   |               |          |
| N <sub>ppt</sub>          | 600                  |                   |               |          |
| N <sub>item</sub>         | 84                   |                   |               |          |
| Observations              | 50400                |                   |               |          |

*Note.* \*  $p < 0.05$  \*\*  $p < 0.01$  \*\*\*  $p < 0.001$ . The omitted levels of multi-level variables in this model are “pre-test” for test, “Control 0” for group.

The final step was estimating statistical power for different sample sizes based on this model and dataset, for which I used the `simr` package<sup>29</sup> in R (Green & Macleod, 2016). The results of the simulations revealed that the effect of the intervention similar to the one illustrated in Figure 28 could be identified in 81% of simulations, 95% CI [72, 88], with 42 test items and 120 participants per group (60 [+art] and 60 [-art] in each group, 45360 total observations). This means that a total of 360 participants would need to be recruited for a statistical power of 81%.

However, if one is only interested in estimating the effect of the intervention on the Experimental group vs. Control 0 on the immediate post-test, roughly the same number of participants per group would be required. Rerunning the simulations on a dataset with only two groups (Control/Experimental) and two test conditions (pre-test/immediate post-test) resulted in 83% power, 95% CI [77, 87], with 42 items and 120 participants per group (60 [+art] and 60 [-art] in each, 20160 total observations). In other words, for an experiment with only two groups, 240 participants would be needed.

Considering that recruiting 240 participants would be challenging, I decided to increase the number of test items where possible, which is detailed in the next subsection. In addition, wherever an increase was made, it was towards the number of items in indefinite contexts, so as to improve the chances of identifying any effects in at least a subset of data. Focusing on indefinite contexts is

<sup>29</sup> I described how the package works in Chapter 3 when reporting the power analysis for the corpus study.

more desirable for this purpose, since this was identified as more challenging in the corpus study, particularly for count singular indefinites, which is also the context more prone to inappropriate form-meaning mappings due to the uninterpretable features involved.

### ***Changes to the Experimental and Testing Materials Based on Pilot Results***

**Experimental Group Activities.** The changes were made in the following aspects of experimental activities.

**Explicit Information.** For this round of data collection, I faced the challenge of recruiting participants for a relatively high-commitment experiment entirely online, as local language schools could no longer support me due to the impact of the pandemic. Therefore, I attempted to boost participant engagement by creating short video lessons (most under 1 minute, with a few up to 2-2.5 minutes long) instead of textual explicit information used in the pilot. The video lessons can be accessed via links provided in Appendix H.

**Number of Items.** I increased the number of activities from 13 to 15, which would increase the total time spent on the intervention to 2.5-3.75 hours. In addition, I removed all the items containing high-frequency vocabulary, which were used in the pilot as a way to help learners familiarise themselves with the type of task, and replaced them with items with low-frequency words, which are supposed to be more effective. I still used examples with high-frequency vocabulary in the short video lessons provided before most activities.

**Low-Frequency Items.** The threshold for low-frequency items was increased from being beyond 4000 most common words to beyond 8000 most common words (Nation, 2017). I replaced any items falling under this threshold while also avoiding cognates (as in the pilot).

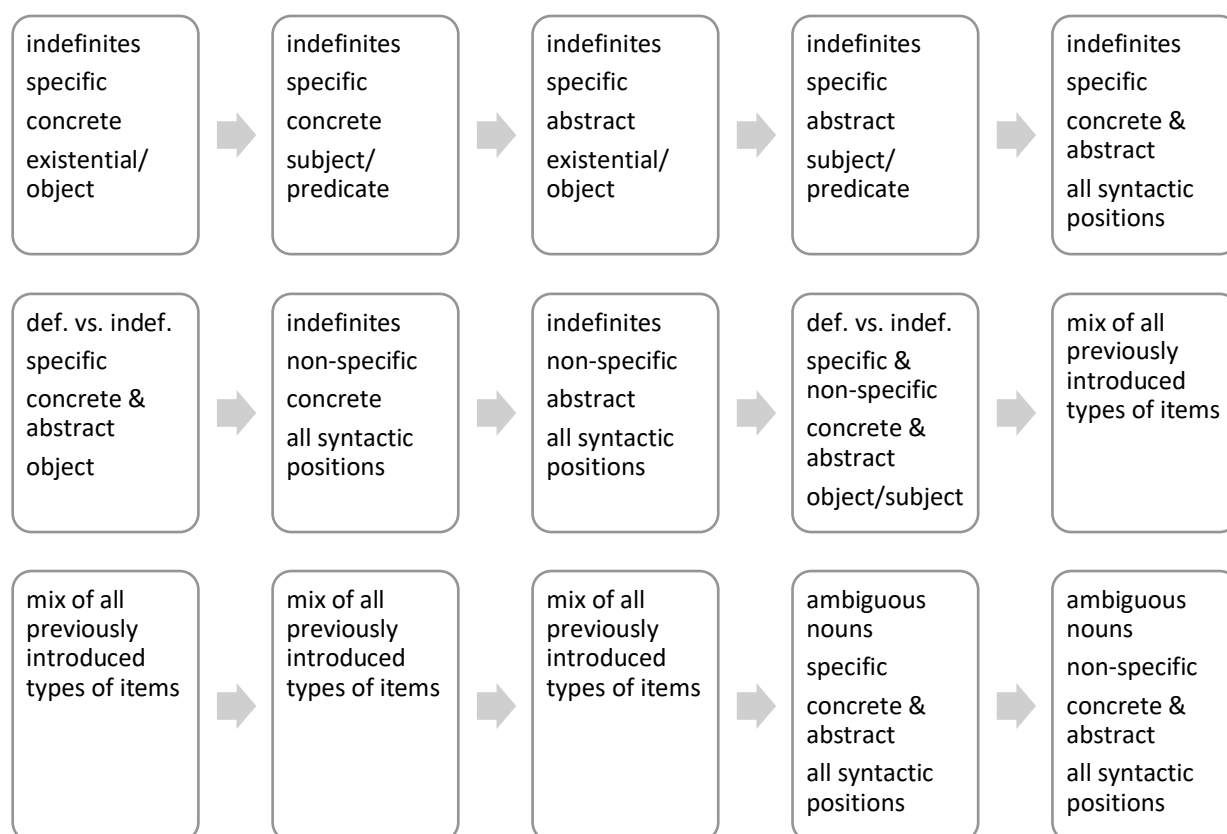
**Order of Activities.** The activities in the pilot introduced different types of items in solid blocks, e.g. one activity had all the specific referents in existential constructions followed by a block with all specific referents in object positions etc. In order to avoid the potential repetitiveness this may create (which could impact on learners' engagement and attention), I mixed the order of items for this experiment, while still gradually introducing new types of items. Figure 29 illustrates the



order for the activities<sup>30</sup>. All the item types were introduced in activities 1-9, while activities 10-13 provided further practice with a mix of all item types. Activities 14-15 introduce ambiguous nouns (e.g. “watermelon”, “cake”), as in the pilot.

**Figure 29**

*Order of Item Presentation for Experimental Activities*



**Singular/Plural Ambiguity in Non-Specific Items.** For this round of data collection, I included items in subject position for non-specific ambiguous nouns. This context was problematic during pilot materials design, because count singular items could not be properly contrasted with their mass counterparts (see footnote 24). To deal with this issue I made both singular and plural options acceptable answers for items such as “a watermelon contains a lot of water”. So, participants could respond that this sentence refers to one watermelon and (optionally) more than one watermelon but not to part of one watermelon. In fact, I extended this rule to all applicable non-specific contexts

<sup>30</sup> Note that semantic transparency of verbs in object position was no longer a factor for indefinite contexts in the new materials. Further explanation of this change is provided below under the “Semantic Transparency Factor” subheading in this section.

(e.g. “I often hear *a chug* late at night”, “She works for *a pittance*”). This change also at least partly addresses the issue of misrepresenting the indefinite article “a” as an equivalent to the numeral “one”.

**Semantic Transparency Factor.** In the new materials, semantic transparency of verbs in object position was no longer a factor for indefinite contexts, although I still ensured that items were varied in terms of using both transparent and non-transparent verbs. This decision was made to reduce the number of factors to be analysed, as analysing each additional factor would reduce the statistical power. The reasons for removing semantic transparency rather than any other factor were that (i) the effect of transparency was not clear in the pilot, (ii) the contrast only applied to nominals in object positions, (iii) constructing pairs of items contrasting by semantic transparency proved to be very difficult. To illustrate point (iii), it was recognised that achieving a specific reading, i.e. an interpretation that the reference was fixed to some existing entity, for items with non-transparent verbs was near impossible. Even attempts such as “I’m looking for *a big box*. It was here 1 hour ago” did not appear convincing upon careful consideration.

**Other Item Changes due to Specificity Readings.** Having discovered the issue with specificity readings and non-transparent verbs described in the previous paragraph, I revisited all the training items scrutinising them for any potential ambiguities in their specificity reading and changed or improved any items that appeared problematic.

**Factorial Structure.** The factorial structure of the new experimental items was similar to the structure in the pilot (Appendix D) with three exceptions detailed above, i.e. (i) removing semantic transparency of the verb as a factor, (ii) adding subject position for non-specific ambiguous nouns, and (iii) replacing all the items which included high-frequency nouns in the pilot with new items containing low-frequency nouns.

Example items demonstrating a mix of different categories from activity 11 and activity 14 are provided in Appendix I.

**Control Group Activities.** The Control group activities were not changed from the pilot, but two additional activities of the same kind were added to match the increased number of activities for the Experimental group.

**Assessment.** The following changes were made to the different test tasks.

**GJT.** For this experiment, the GJT was timed giving participants 10-30 seconds to respond to each item (depending on item length). This was done to mitigate the metalinguistic nature of the task and tap into intuitive judgments (Ellis, 2005).

Considering the results of the power analysis, I increased the number of target items for the GJT from 44 to 58 by adding more of the indefinite test items (i.e. those targeting the contrast between “a” and  $\emptyset$ ). I also increased the number of distractors from 22 to 42 (for a total of 100 items per test).

Moreover, I tested all the items with native speakers: approximately 30 volunteers (PhD candidates in a variety of subjects, mostly non-linguists) responded to randomly selected subsets of items. As a result, I obtained at least 3-4 responses to each item and modified any items that were not consistently rated as intended, which constituted around 35% of the original items. I then iteratively tested new items until a satisfactory response rate was achieved (at least 3/3 or 3/4).

**EI.** Instead of including 3 target items per prompt, as in the pilot, I modified EI prompts to contain only one target item each, making sure that it appeared in the first half of the prompt but not as the first word. I consequently increased the number of prompts to a total of 32 per test (the pilot contained 36 targets across 12 prompts per test).

I transformed all the prompts into sentences (some items in the pilot were short dialogues), which were syntactically complex (i.e. included subordinate clauses) to minimise the possibility of word-for-word recall (68). Prompts contained 20 words on average (range: 15-25).

(68) I can probably get *an extra ticket* for tomorrow's show if you would like to take your new friend with us.

Considering the feedback received from the pilot participants, I explicitly asked the native speakers who recorded the prompts to speak slightly more slowly and clearly as if addressing an intermediate language learner. I tested a subset of the new EI prompts with a small group of learners ( $n = 12$ ) of mostly B1/B2 proficiency levels and obtained an average 77% target production rate (a considerable improvement from the 40-45% on the pilot). Half of these learners heard the prompt played once and the other half heard it twice, but this did not appear to significantly affect their ability to reproduce intended targets. Thus, for the experiment each prompt was played only once. Participants were given one minute to type their response after hearing each prompt.

As in the pilot, the target nouns for both GJT and EI were all within the 3000 most frequent words in English according to Paul Nation's BNC/COCA headword list (Nation, 2017), and no word was used more than once to avoid any practice effects.

**Free Writing.** Participants were asked to produce two pieces of writing for this experiment (as opposed to one in the pilot). The first task involved narrating a short silent (word-less) cartoon<sup>31</sup> to someone who has not seen it. In this task, learners were required to produce five chosen key words (names of objects present in the cartoon, some countable and some uncountable) to ensure comparability across participants for at least a subset of items. This narrative task was intended to elicit the use of concrete nouns in both indefinite (first mention) and definite (anaphoric) contexts. Participants were given 10 minutes to complete this task.

The second task was a short essay prompt on a topic of personal interest, which was intended to elicit more abstract vocabulary (e.g. expressing an opinion on the use of social media or on Covid-19 restrictions in one's country). Participants were given 15 minutes to complete this task.

**Timing.** As the number of items on all the tasks was increased to improve the statistical power of the experiment, it became evident that asking participants to complete the whole test in one sitting would be unreasonable and potentially discouraging. Thus, each test version was split

---

<sup>31</sup> Cartoons from the Simon's Cat series were used for this purpose.

into two parts, each containing half of the GJT and EI items and one of the two free writing tasks. GJT and EI items were divided in a principled manner, so that each part of the test included all types of items. For one part of the test, a participant would spend 12-15 minutes on the timed GJT, 20-25 minutes on the EI tasks, and 10-15 minutes on the free writing, amounting to a total of approximately an hour.

The structure of the test tasks and the full list of test items for each test version is provided in Appendix J.

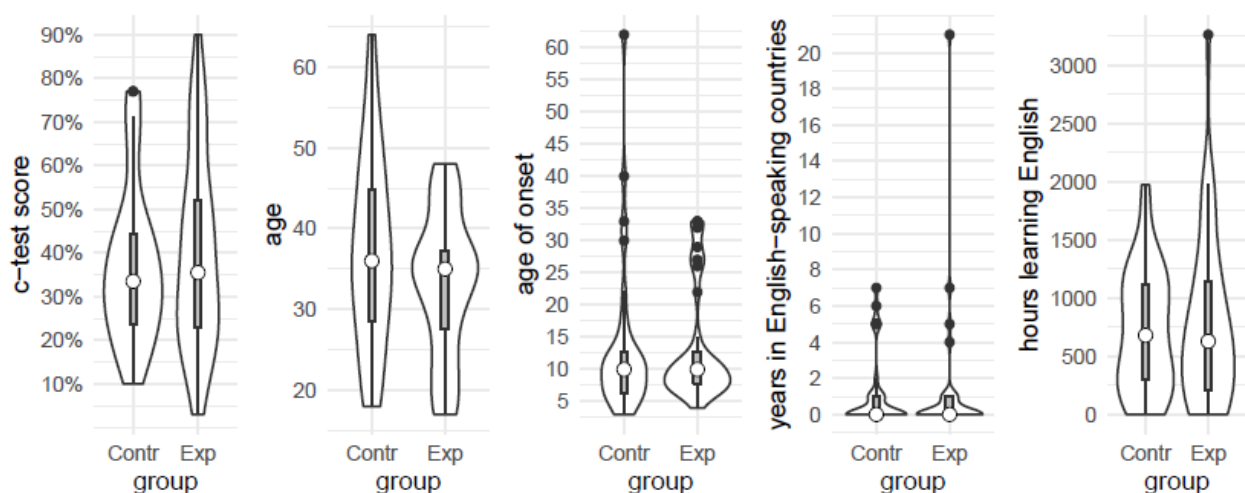
### ***Participants***

After receiving approval from the Research Ethics Committee of the Faculty of Modern and Medieval Languages and Linguistics of the University of Cambridge, I recruited participants online through social media and a network of personal contacts with language teachers and learners. Those who chose to participate were offered a certificate of participation and personal feedback on one of the short essays they would need to write as part of their immediate post-test. This feedback was provided by me to each participant sometime between the immediate and the delayed post-tests and contained comments on the content of the essays as well as grammatical and vocabulary corrections. If any article errors were made, these were not distinguished from any other types of errors in any way and no reference to the experimental materials was made in the comments.

Of the 317 learners who registered for the experiment, 71 proceeded to complete all the activities and the immediate post-test. Of these, 22 participants also completed the delayed post-test three months later. One participant's data had to be discarded due to a technical error (she could not complete all the required experimental group activities). Thus, in this study I analyse data from 70 participants. The information about the participants collected through the personal and language background questionnaire is presented in Table 16 and Figure 30. The questionnaire was similar to the one used in the pilot, only slightly reduced, and is provided in Appendix K in full.

**Table 16***Participant Distribution by L1, Proficiency Level, Education Level, Gender*

| Variable   | Categories                             | Number of Participants |         |       |
|--|--|------------------------|---------|-------|
|  |  | Experimental           | Control | Total |
| L1   | [+art]: Spanish                        | 10                     | 10      | 20    |
|  | Brazilian Portuguese                   | 4                      | 5       | 9     |
|  | [-art]: Russian/Ukrainian <sup>a</sup> | 18                     | 14      | 32    |
|  | Japanese                               | 2                      | 2       | 4     |
|  | Polish                                 | 1                      | 1       | 2     |
|  | Serbian                                | 0                      | 1       | 1     |
|  | Kazakh                                 | 0                      | 1       | 1     |
|  | Kabardian                              | 1                      | 0       | 1     |
| Self-reported proficiency level<br>(according to CEFR) | A1                                     | 1                      | 1       | 2     |
|  | A2                                     | 6                      | 4       | 10    |
|  | B1                                     | 12                     | 12      | 24    |
|  | B2                                     | 14                     | 15      | 29    |
|  | C1                                     | 2                      | 2       | 4     |
|  | Unknown                                | 1                      | 0       | 1     |
| Highest level of education<br>completed                | High school                            | 10                     | 5       | 15    |
|  | Undergraduate                          | 10                     | 6       | 16    |
|  | Postgraduate                           | 16                     | 23      | 39    |
| Gender   | Female                                 | 28                     | 27      | 55    |
|  | Male                                   | 8                      | 7       | 15    |

<sup>a</sup> Russian/Ukrainian are reported together because many of the participants were bilingual.**Figure 30***Participant Distribution by C-test<sup>32</sup> Scores, Age, Age of Onset of English, Years in English-Speaking Countries, Hours Learning English*<sup>32</sup> Further details on the c-test are given in the “Procedure” subsection below.

### ***Procedure***

Participants over 18 years old received a registration link where they could read the Study Information Sheet detailing what they could expect if they decided to participate. They were then asked to complete an online consent form and a questionnaire, which included some general questions (age, gender, education level) and some language background questions (level of English, native language etc.). Finally, they were asked to do a short c-test (10 minutes), which was used to determine their general English level relative to other participants. For this purpose, I used two of the three c-test tasks developed and validated by Park & Choi (2018).

Participants aged 16-17 were recruited through English Teachers in schools as well as private English Tutors, who were informed about the study and asked to contact students' parents in the first instance. Interested parents received a registration link for parents, where they could read the Study Information Sheet and complete an online form acknowledging their awareness of the experiment procedure and providing contact details of their children. The students registered in this way then received the link to the consent form and the questionnaire described above.

The Study Information Sheet, parental acknowledgment form, participant consent form and questionnaire, and c-test are given in Appendix K<sup>33</sup>.

Each registered participant was randomly assigned to either the Control or the Experimental group. The programme also randomly assigned which test variant each participant was going to do at pre-test, immediate post-test, and delayed post-test (e.g. A-B-C, B-C-A etc.). Participants received the link to Part 1 of their pre-test on the following day after registering and the link to Part 2 on the following day after completing Part 1. One day after finishing Part 2, participants received the link to the first Control or Experimental activity. The link to each subsequent activity was distributed one day after the preceding one was completed. Participants received reminders on the third day if they did not finish the most recent activity sent to them. The link to Part 1 of the immediate post-test was

---

<sup>33</sup> In order to capture more lower-level participants (an issue identified in the pilot), I also provided translations of the information sheets and consent forms into Russian and Ukrainian, the languages in which I expected to recruit the most [-art] L1 learners.

send on the following day after completing the final (fifteenth) activity and the link to Part 2 on the following day after completing Part 1. Three months after finishing Part 2 of the immediate post-test, participants received links to their delayed post-tests (similarly, first Part 1 and Part 2 on the following day).

The entire experiment, from registration to delayed post-test, was set up online using Qualtrics software (Qualtrics, 2020). All the forms, activities, and tests were accessible from any mobile device or computer. Participants were encouraged to complete all the fifteen activities within three weeks (21 days), however 40 out of 70 participants took longer than that ( $M = 25$ ,  $SD = 17.7$ , range: 14-103).

### ***Coding and Scoring***

On the GJT, participants received 1 point for correctly identifying whether the item is correct or contains an error and 0 points otherwise.

The elicited imitation task was coded manually by the researcher. First, I identified whether the target items were produced by each participant. In case a prompt was paraphrased without using the target item but clearly showing that it was understood, this was marked as avoidance. Such items were not included in the analysis in the pilot, where I made no distinction between incomprehensible items and avoidance of the intended targets. Where it appeared that a participant did not understand the prompt, this was not included in the analysis even if they produced some words resembling the intended target.

Items were excluded from the analysis if they were changed in ways that would affect article use, such as:

- Making a singular noun plural or vice versa
- Using a mass noun instead of a count noun (69) or vice versa

(69) Prompt: Generally, *an essay* should be marked by two examiners [...]

Response: Actually *the work* should be checked by two examiners [...]

- Changing a definite context into an indefinite one (70) or vice versa



(70) Prompt: He liked *the advice* that he got [...]

Response: He likes to give *advise* [...]

- Adding determiners which would preclude article use, e.g. demonstratives (“this”, “that”), possessive pronouns or nouns (“my”, “teacher’s”), quantifiers (“some”, “any”) etc.

Other changes did not disqualify items from the analysis but were marked accordingly:

- Changing the syntactic position of the target item
- Adding or dropping a prenominal modifier
- Making a specific referent nonspecific or vice versa (71)

(71) Prompt: I can probably get *an extra ticket* for tomorrow's show [...]

Response: I have *an extra ticket* for tomorrow's show [...]

- Using an abstract noun instead of a concrete one or vice versa (72)

(72) Prompt: You must lower *the concentration* of gas in these tubes [...]

Response: You must lower *the gas* in these tubes [...]

- Using a synonym or a similar type of word that made sense in the context.

Any nominals other than the target items were not considered for analysis. Once the reproduced targets were identified, they were coded for correctness and error types (omission of “a”/“the”, overuse of “a”/“the”, substitution of “a” with “the” or vice versa, avoidance). Correct items scored 1 point, items containing errors or avoidance scored 0.

## **Analysis and Results**

### ***Mixed-Effects Logistic Regression***

To investigate the effect of the intervention on learners' article accuracy, I used generalised linear mixed-effects logistic regression models where the dependent variable is binary (score 0/1), using the `lme4` package in R (Bates, Mächler, et al., 2015). I fitted separate models on the GJT and on the EI data. Within each of the two tasks I fitted one model on the results of the pre-tests and immediate post-tests of all the 70 participants and another model on the results of all the tests of just the 21 participants who completed delayed post-tests.

The need for mixed-effects models is justified by the fact that each participant responded to multiple items on the tests and each test item was responded to by multiple participants. Thus, there is likely some by-participant and by-item variance to be accounted for.

In the process of model selection, I always started with a null model and iteratively added variables which I expected to have an effect comparing models with log-likelihood ratio tests. When adding fixed effects to the model, I also included them as random slopes where possible. Following Bates et al.'s (2015) recommendations, I used Principal Components Analysis (PCA) and log-likelihood ratio tests to attain the most parsimonious random effects structures.

### ***Comparing Groups at Pre-Test***

To check that the Control and the Experimental group were not significantly different from each other, I ran a generalised linear regression model to test whether any of the participant characteristics were significantly associated with being allocated to one of the groups. Thus, the "group" was the outcome variable and the following participant characteristics were the independent variables: gender, age, education level, c-test score, self-reported CEFR level, age of onset, years spend in English-speaking countries, hours of learning English, attendance of English classes at the time of pre-test. Apart from the c-test, the values for the independent variables were derived from the participants' responses to the personal and language background questionnaires. As seen in Table 17, the only significant factor is attendance of English classes. More specifically,

participants in the Experimental group were less likely to be attending English classes at the time of the pre-test, but there was no such imbalance in the Control group.

**Table 17**

*Results of the Model Comparing Control and Experimental Groups at Pre-Test*

| <i>Predictors</i>                   | <b>Probability of Being in the Experimental Group</b> |                   |               |          |
|-------------------------------------|---|-------------------|---------------|----------|
|                                     | <i>Log-Odds</i>                                       | <i>std. Error</i> | <i>95% CI</i> | <i>p</i> |
| (Intercept)                         | -1.81   | 2.04              | -6.10   2.33  | 0.376    |
| Gender: female                      | 0.87  | 0.80              | -0.65   2.54  | 0.278    |
| Education level: Undergraduate      | -0.27   | 0.93              | -2.12   1.57  | 0.768    |
| Education level: Postgraduate       | -1.36   | 0.92              | -3.29   0.37  | 0.138    |
| Age                                 | -0.46   | 0.34              | -1.16   0.18  | 0.170    |
| Age of onset                        | 0.31  | 0.37              | -0.43   1.07  | 0.398    |
| Hours of English                    | 0.05  | 0.34              | -0.61   0.76  | 0.881    |
| Years in English-speaking countries | -0.63   | 0.56              | -1.84   0.41  | 0.256    |
| C-test score                        | 0.32  | 0.36              | -0.38   1.07  | 0.374    |
| Attending English classes: No       | 1.60 *  | 0.67              | 0.36   3.03   | 0.017    |
| Self-reported CEFR level: A2        | 1.74  | 1.90              | -2.22   5.72  | 0.360    |
| Self-reported CEFR level: B1        | 1.12  | 1.82              | -2.73   4.95  | 0.539    |
| Self-reported CEFR level: B2        | 0.63  | 1.81              | -3.22   4.43  | 0.727    |
| Self-reported CEFR level: C1        | 1.99  | 2.31              | -2.61   6.78  | 0.388    |

*Note.* \*  $p < 0.05$  \*\*  $p < 0.01$  \*\*\*  $p < 0.001$ . The omitted levels of multi-level variables in this model

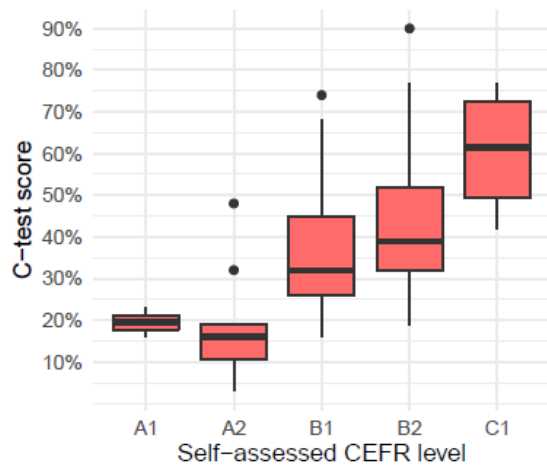
are “High school” for Education level, “A1” for Self-reported CEFR level.

Most importantly, the two groups did not differ in their proficiency levels, either self-reported or measured by the c-test. In fact, self-reported levels appear to correlate well with the results of the c-test, as shown in Figure 31. In further analysis, I relied on the c-test scores as the measure of proficiency level.

Additionally, I checked that participants’ scores were not significantly different on different test versions or parts of the test.

**Figure 31**

*Correlation between Self-Reported Proficiency Levels and C-test Scores*



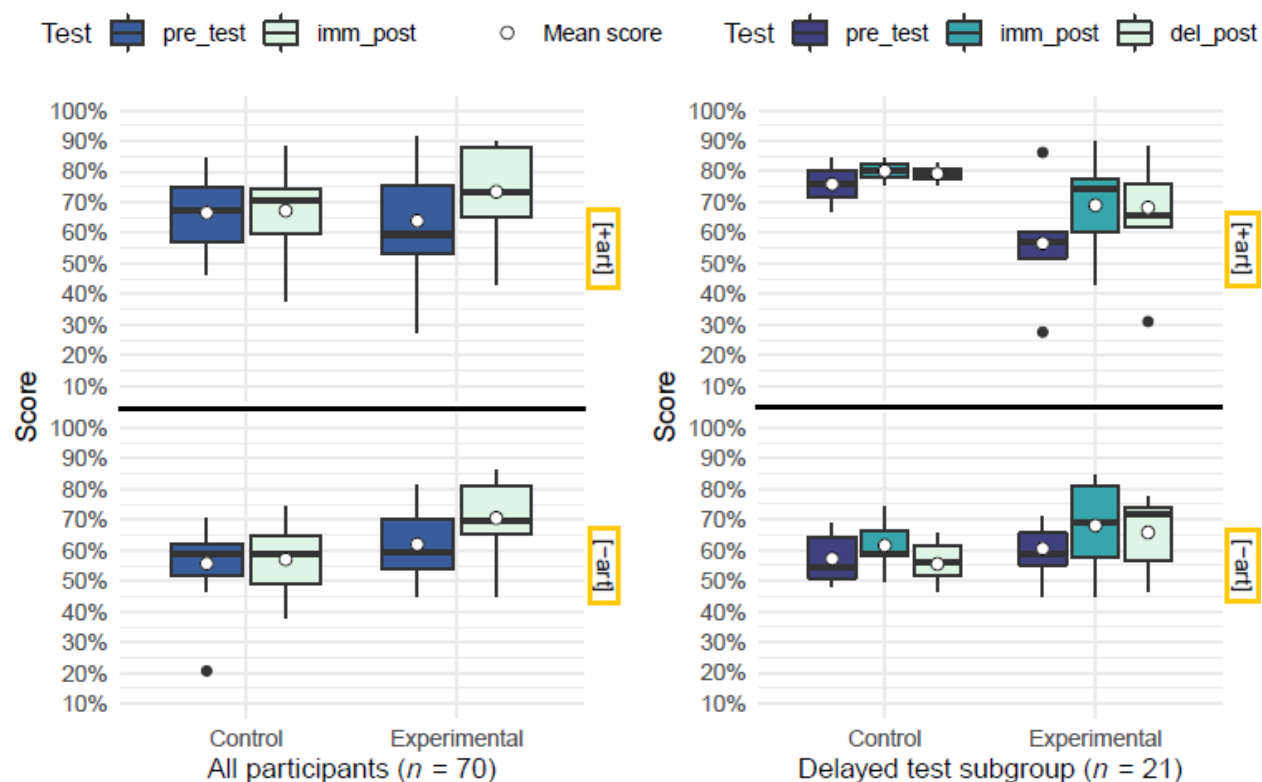
### **GJT Results**

**Descriptive Statistics.** The overall group scores on pre-test, immediate post-test, and delayed post-test split by L1-type are shown in Figure 32 (note that the right-hand side shows data only from the 21 participants who completed the delayed post-test). It appears that for both [+art] and [-art] L1-types the Experimental group improved at the immediate post-test and retained the improvement by the delayed post-test. The Control group, on the other hand, appears to have roughly the same results across all the tests.

**The Effect of the Intervention.** To check whether the intervention had a statistically significant effect and which factors impacted the results, I ran a mixed-effects generalised linear regression model with score as a binary outcome (1 – correct, 0 – incorrect). The results of the model (henceforth, GJT Model) along with model performance metrics are presented in Table 18.

**Figure 32**

*The Overall Effect of the Intervention on GJT Target<sup>34</sup> Items by L1-type*

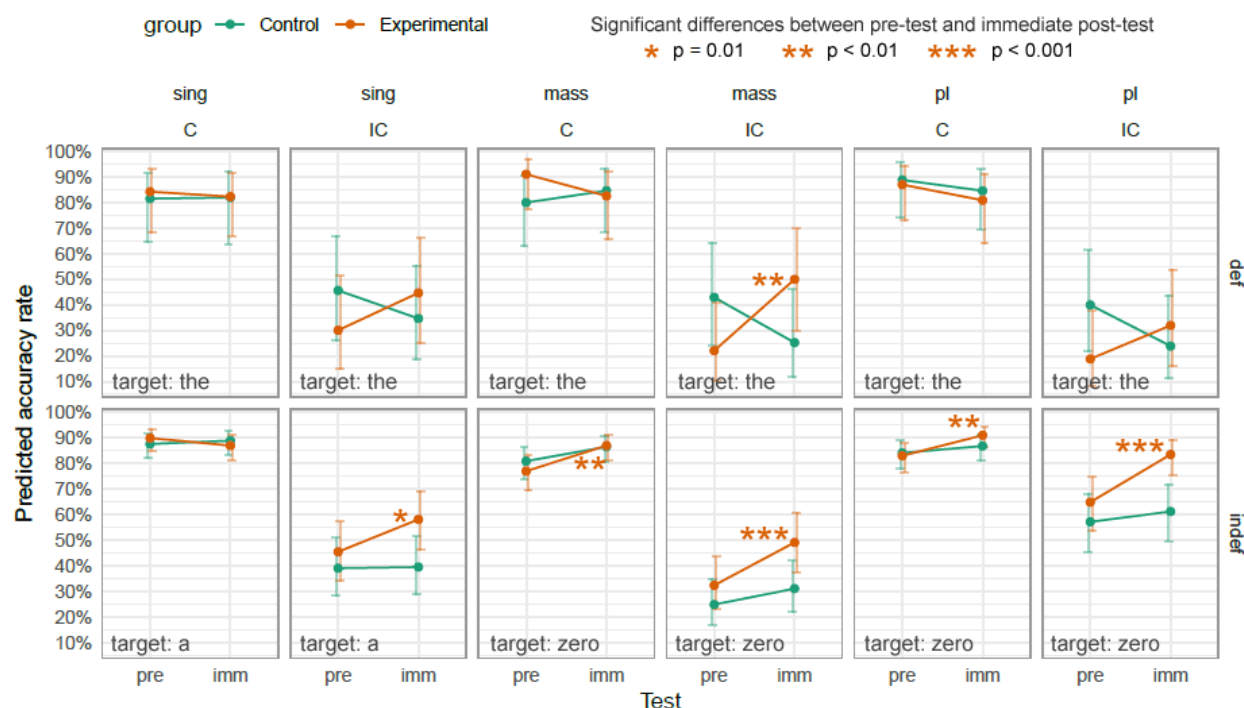


Overall, the model reveals that the Experimental group improved significantly from pre-test to immediate post-test ( $p = 0.0014$ ), but the Control group did not ( $p = 0.58$ ). The Experimental group is predicted to improve by about 8% overall, going from 64%, 95% CI [57%, 70%], at pre-test to 72%, 95% CI [67%, 77%], at immediate post-test, whereas the Control group is predicted to go from 66%, 95% CI [59%, 72%], at pre-test, to 64%, 95% CI [58%, 70%], at immediate post-test. This effect interacts with item variant (grammatical or ungrammatical GJT prompt), definiteness and noun type, as shown in Figure 33.

<sup>34</sup> The results for distractor items are not included in any of the figures or tables.

**Figure 33**

*Predicted Effect of Intervention on GJT Accuracy Rate by Definiteness, Noun Type, and Item Variant*  
(with 95% Confidence Intervals)



Note. C refers to grammatical (correct) and IC to ungrammatical (incorrect) item variants on the GJT.

The first observation about Figure 33 is the clear difference between the grammatical/correct (C) item variants, where learners were required to recognise that there were no errors, and the ungrammatical/incorrect (IC) item variants, where learners had to spot an article error. Learners scored significantly higher on C than on IC variants, which means they were generally more likely to accept an item as correct.

The second observation is that the accuracy of spotting errors in IC variants was generally low (below 50% in most facets of Figure 33). The only exception is indefinite plurals, which is, perhaps, because errors of overusing “a” with plural nouns are very obvious (e.g. “a books”), although this category also includes overusing “the” (e.g. “the books” in an indefinite context).

Figure 33 demonstrates that the Control group is not predicted to improve significantly in either definite or indefinite contexts regardless of noun type or item variant. By contrast, the Experimental group improved significantly in most indefinite contexts. The effect is particularly large

for IC mass (73) and plural items (74), which means learners became better at spotting article overuse.

(73) \*A *snow* covered the playground.

(74) Listening to \*a *different accents* is useful.

To investigate whether there is a significant difference between learners' accuracy in spotting overuse of "a" vs. overuse of "the", I ran a separate model on indefinite mass and plurals only<sup>35</sup>. The results of this model are generally congruent with the results of the GJT Model (Appendix L). In addition, it estimates that learners were generally better at noticing overuse of "a" than overuse of "the", and this difference was particularly drastic for plural nouns (75) (75% predicted accuracy vs. 25%, on average). The effect of error type does not interact with group and test, however, which means that the Experimental group improved their ability to spot both error types.

(75) I had a strange dream. \*The *random people* were asking me questions in different languages. I couldn't understand anything!

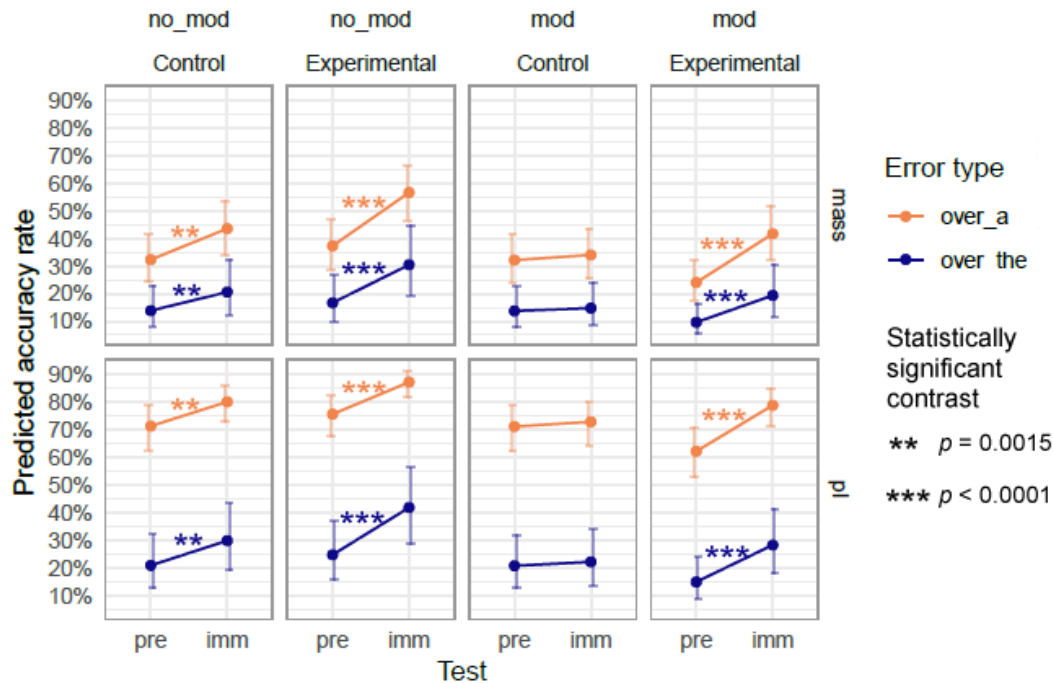
A more interesting result of this model, however, is that it identifies a significant interaction between group, test, and modifier presence. More specifically, it estimates that the Control group, in fact, also improves significantly on spotting article overuse with non-premodified mass and plural indefinites (such as example (73)). The Experimental group, by contrast, is predicted to improve on all types of items, whether premodified or not (Figure 34), and their accuracy gain is consistently larger.

---

<sup>35</sup> Error type models had to be fitted separately on subsets of data because different categories of items have different error type possibilities: omission of "a" or substitution of "a" with "the" in indefinite singulars; omission of "the" or substitution of "the" with "a" in all definites; overuse of "a" or "the" in mass and plural indefinites.

**Figure 34**

*Predicted Accuracy in Identifying Errors on IC Indefinite Mass and Plural GJT Items with or without Prenominal Modifiers*



Returning to the overall effects illustrated in Figure 33, the Experimental group also improved significantly on IC indefinite singular items, i.e. they became better at noticing errors of omission of “a” (76) and errors of substitution of “a” with “the” (77).

(76)There is *\*question* at the end.

(77)Let's sit down. I see *\*the free bench* over there.

A separate model on indefinite singulars only is generally congruent with the results of the GJT Model (Appendix M). There is a significant effect of error type in that learners were better at spotting omission of “a” than substitution of “a” with “the”. However, this effect does not interact with group and test, which means that the Experimental group improved on both types of error. Importantly, this model estimates that the accuracy gain is only statistically significant for [-art] L1 learners ( $p = 0.01$ ), although the tendency is similar for [+art] L1 learners.



In definite contexts, the Experimental group also showed improvement trends on all IC items, but the only statistically significant change was for mass definites, where learners became better at noticing errors of omission of “the” (78).

(78) \**Literature* that you gave me was useful.

**Table 18**

*GJT Model Results*

| <i>Predictors</i>   | <b>Accuracy Rate</b> |                   |               |
|---|----------------------|-------------------|---------------|
|   | <i>Log-Odds</i>      | <i>std. Error</i> | <i>95% CI</i> |
| (Intercept)   | 2.70 ***             | 0.56              | 1.61   3.79   |
| group: Experimental   | -0.17                | 0.60              | -1.34   1.00  |
| test: immediate post  | -0.36                | 0.58              | -1.51   0.78  |
| item variant: IC  | -2.46 ***            | 0.67              | -3.78   -1.14 |
| L1-type: [-art]   | -0.82 **             | 0.30              | -1.40   -0.24 |
| def: indefinite   | -0.86                | 0.59              | -2.01   0.29  |
| Ntype: singular   | -0.78                | 0.72              | -2.20   0.63  |
| Ntype: mass   | -1.02                | 0.71              | -2.41   0.36  |
| c-test  | 0.37 ***             | 0.08              | 0.21   0.52   |
| abstr: abstract   | -0.18 *              | 0.09              | -0.36   -0.00 |
| mod: modified   | -0.29 **             | 0.09              | -0.47   -0.12 |
| group: Experimental * test: immediate post                    | -0.09                | 0.79              | -1.63   1.45  |
| group: Experimental * item variant: IC                        | -0.88                | 0.82              | -2.49   0.72  |
| test: immediate post * item variant: IC                       | -0.39                | 0.77              | -1.90   1.13  |
| group: Experimental * def: indefinite                         | 0.09                 | 0.64              | -1.16   1.34  |
| test: immediate post * def: indefinite                        | 0.58                 | 0.62              | -0.65   1.80  |
| item variant: IC * def: indefinite                            | 1.09                 | 0.70              | -0.28   2.47  |
| L1-type: [-art] * def: indefinite                             | 0.92 **              | 0.32              | 0.29   1.55   |
| group: Experimental * Ntype: singular                         | 0.36                 | 0.80              | -1.22   1.94  |
| group: Experimental * Ntype: mass                             | 1.09                 | 0.84              | -0.56   2.74  |
| test: immediate post * Ntype: singular                        | 0.39                 | 0.81              | -1.19   1.97  |
| test: immediate post * Ntype: mass                            | 0.68                 | 0.79              | -0.86   2.22  |
| item variant: IC * Ntype: singular                            | 0.81                 | 0.89              | -0.94   2.56  |
| item variant: IC * Ntype: mass                                | 0.80                 | 0.88              | -0.93   2.53  |
| L1-type: [-art] * Ntype: singular                             | 0.41                 | 0.39              | -0.36   1.17  |
| L1-type: [-art] * Ntype: mass                                 | 0.69                 | 0.39              | -0.07   1.46  |
| def: indefinite * Ntype: singular                             | 1.38                 | 0.77              | -0.13   2.89  |
| def: indefinite * Ntype: mass                                 | 0.91                 | 0.76              | -0.57   2.40  |
| group: Experimental * test: immediate post * item variant: IC | 1.54                 | 1.09              | -0.59   3.68  |
| group: Experimental * test: immediate post * def: indefinite  | 0.59                 | 0.85              | -1.08   2.26  |
| group: Experimental * item variant: IC * def: indefinite      | 1.29                 | 0.85              | -0.37   2.94  |
| test: immediate post * item variant: IC * def: indefinite     | 0.34                 | 0.83              | -1.28   1.96  |
| group: Experimental * test: immediate post * Ntype: singular  | -0.08                | 1.09              | -2.22   2.07  |
| group: Experimental * test: immediate post * Ntype: mass      | -0.98                | 1.14              | -3.22   1.26  |
| group: Experimental * item variant: IC * Ntype: singular      | 0.02                 | 1.08              | -2.10   2.15  |
| group: Experimental * item variant: IC * Ntype: mass          | -1.01                | 1.11              | -3.19   1.16  |

| <i>Predictors</i>   | <b>Accuracy Rate</b> |                   |               |
|---|----------------------|-------------------|---------------|
|   | <i>Log-Odds</i>      | <i>std. Error</i> | <i>95% CI</i> |
| test: immediate post * item variant: IC * Ntype: singular   | -0.10                | 1.07              | -2.19   1.99  |
| test: immediate post * item variant: IC * Ntype: mass   | -0.73                | 1.07              | -2.82   1.36  |
| group: Experimental * def: indefinite * Ntype: singular   | -0.05                | 0.86              | -1.74   1.64  |
| group: Experimental * def: indefinite * Ntype: mass   | -1.24                | 0.89              | -2.99   0.51  |
| test: immediate post * def: indefinite * Ntype: singular  | -0.48                | 0.87              | -2.19   1.22  |
| test: immediate post * def: indefinite * Ntype: mass  | -0.48                | 0.85              | -2.15   1.18  |
| item variant: IC * def: indefinite * Ntype: singular  | -1.83                | 0.96              | -3.71   0.05  |
| item variant: IC * def: indefinite * Ntype: mass  | -1.97 *              | 0.95              | -3.83   -0.11 |
| L1-type: [-art] * def: indefinite * Ntype: singular   | -1.03 *              | 0.41              | -1.84   -0.22 |
| L1-type: [-art] * def: indefinite * Ntype: mass   | -0.91 *              | 0.42              | -1.73   -0.09 |
| group: Experimental * test: immediate post * item variant: IC * def: indefinite                   | -1.21                | 1.18              | -3.52   1.09  |
| group: Experimental * test: immediate post * item variant: IC * Ntype: singular                   | -0.29                | 1.51              | -3.24   2.66  |
| group: Experimental * test: immediate post * item variant: IC * Ntype: mass                       | 1.58                 | 1.55              | -1.45   4.61  |
| group: Experimental * test: immediate post * def: indefinite * Ntype: singular                    | -0.83                | 1.19              | -3.17   1.50  |
| group: Experimental * test: immediate post * def: indefinite * Ntype: mass                        | 0.75                 | 1.23              | -1.65   3.16  |
| group: Experimental * item variant: IC * def: indefinite * Ntype: singular                        | -0.40                | 1.16              | -2.68   1.88  |
| group: Experimental * item variant: IC * def: indefinite * Ntype: mass                            | 1.21                 | 1.19              | -1.11   3.54  |
| test: immediate post * item variant: IC * def: indefinite * Ntype: singular                       | 0.04                 | 1.15              | -2.21   2.30  |
| test: immediate post * item variant: IC * def: indefinite * Ntype: mass                           | 0.68                 | 1.15              | -1.58   2.93  |
| group: Experimental * test: immediate post * item variant: IC * def: indefinite * Ntype: singular | 0.86                 | 1.63              | -2.33   4.04  |
| group: Experimental * test: immediate post * item variant: IC * def: indefinite * Ntype: mass     | -1.80                | 1.66              | -5.06   1.46  |
| <b>Random Effects</b>   |                      |                   |               |
| <i>By item ID</i>   |                      |                   |               |
| SD (Intercept)  | 0.64                 |                   |               |
| SD (c-test)   | 0.25                 |                   |               |
| <i>By participant ID</i>  |                      |                   |               |
| SD (Intercept)  | 0.31                 |                   |               |
| SD (Ntype: singular)  | 0.30                 |                   |               |
| SD (Ntype: mass)  | 0.18                 |                   |               |
| SD (def: indefinite)  | 0.46                 |                   |               |
| SD (item variant: IC)   | 0.84                 |                   |               |
| N <sub>ppt_id</sub>   | 70                   |                   |               |
| N <sub>item_id</sub>  | 348                  |                   |               |
| Observations  | 8062                 |                   |               |
| Marginal R <sup>2</sup> / Conditional R <sup>2</sup>  | 0.305 / 0.305        |                   |               |

| <i>Predictors</i>   | <b>Accuracy Rate</b>   |                   |               |
|---|--|-------------------|---------------|
|   | <i>Log-Odds</i>  | <i>std. Error</i> | <i>95% CI</i> |
| Log-likelihood ratio test comparing to the null model (only random effects) | $\chi^2 (56) = 336.42, p < 10^{-15}$   |                   |               |
| C-statistic   | 0.86 (strong predictive power)   |                   |               |
| VIFs  | High in a model with correlations but low in a no-correlation model (1.00 for all variables) |                   |               |
| Overdispersion ratio  | 0.837 ( $\chi^2 = 6694.387, p = 1$ )   |                   |               |

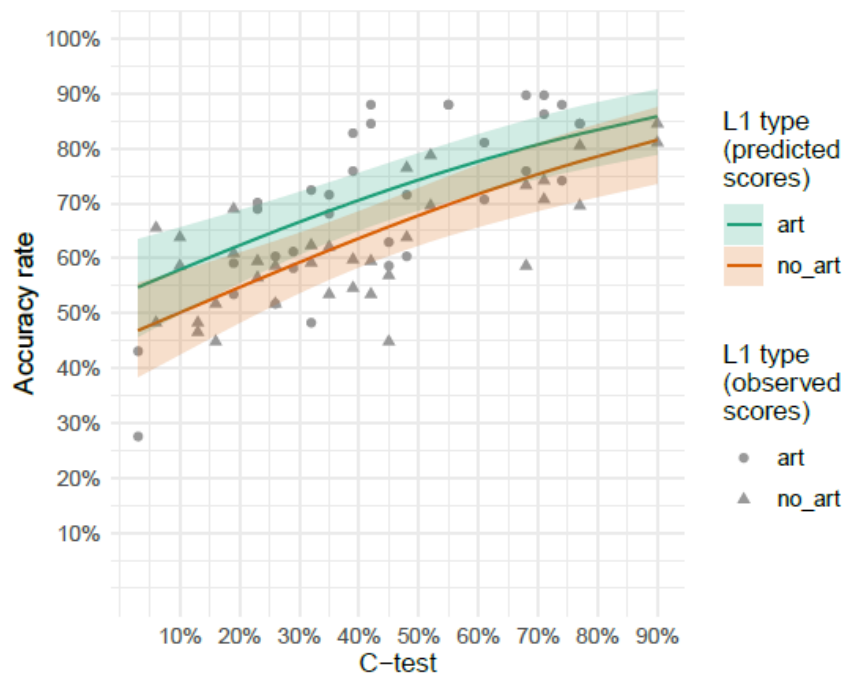
*Note.* \*  $p < 0.05$  \*\*  $p < 0.01$  \*\*\*  $p < 0.001$ . The omitted level of the noun type (Ntype) multi-level variable in this model is “plural”.

**Other Important Variables not Affected by the Intervention.** A few other predictors had a significant effect on learner accuracy, which did not change after the intervention.

**Proficiency Level.** The model reveals a significant effect of proficiency level as measured by the c-test completed by each participant prior to the pre-test. Figure 35 shows that the higher a participant scored on the c-test, the higher their accuracy rate was on the GJT, with [+art] L1 learners generally outperforming [-art] L1 learners. This applies to both pre- and post-test, but the effect of proficiency does not interact with group or test, meaning that regardless of the initial proficiency level, the Experimental group participants improved as a result of the intervention, but the Control group participants did not.

**Figure 35**

*The Effect of Proficiency Level on GJT Accuracy Rate (with 95% Confidence Intervals)*



**L1-Type in Interaction with Definiteness and Noun Type.** As expected, I found an effect of L1-type, which interacts with definiteness and noun type. The crucial contrast driving this interaction is illustrated in Figure 36, which shows predicted accuracy rates for the ungrammatical (IC) GJT items for both pre- and immediate post-test. As above (Figure 33), we see that learners were markedly better at spotting article overuse with indefinite plurals, which was true for both L1-types. The difference is that for [-art] L1 learners, accuracy was significantly lower for both singular and mass indefinites, i.e. they could not reliably recognise either article overuse with mass nouns or article omission/substitution with singular count nouns. As for [+art] L1 learners, they were similarly inaccurate in spotting article overuse with mass indefinites but did not struggle to the same extent with singular count indefinites, where they were significantly better than their [-art] L1 counterparts ( $p = 0.01$ ), although still only at 52% predicted accuracy.

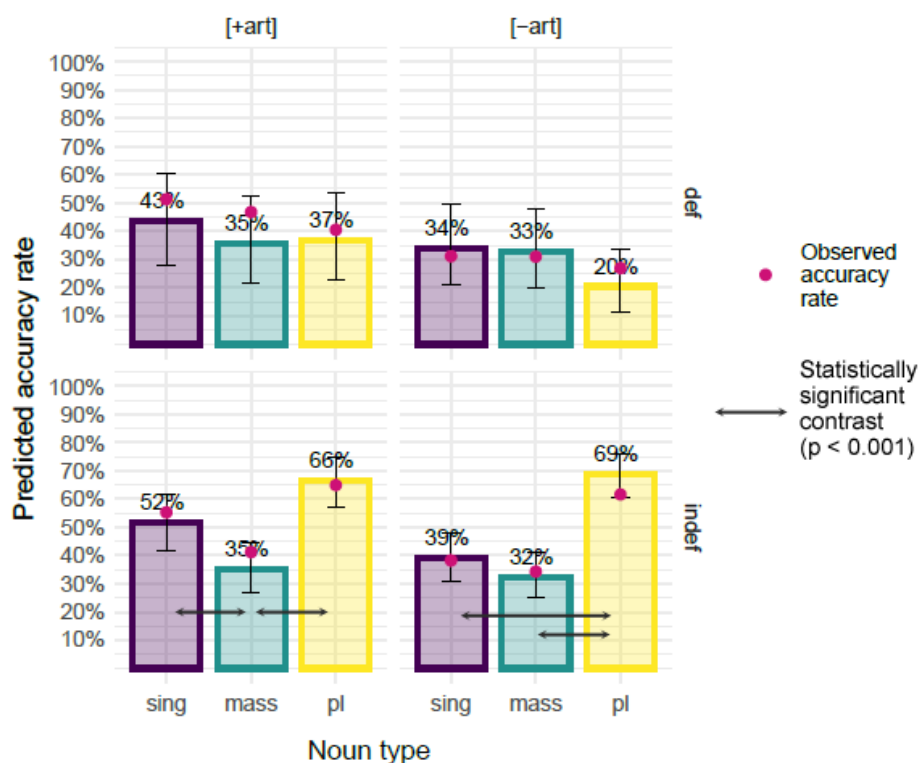
In definite contexts, all learners were well below 50% accuracy, and it appears that [-art] L1 learners performed slightly worse (statistically significant for plurals at  $p = 0.006$ ). However, note the

relatively large confidence intervals, which reflect the fact that the tests contained relatively few definite items (4 C and 4 IC per test).

Importantly, the effect of L1-type does not interact with group or test, meaning that both L1-types were likely to improve their GJT scores after the intervention.

**Figure 36**

*The Effect of Noun Type in Interaction with Definiteness and L1 type on Learner Accuracy on Ungrammatical GJT Items (with 95% Confidence Intervals)*



**Modifier Presence.** The model predicts that all learners are likely to be significantly less accurate ( $p = 0.001$ ) on those GJT items which include prenominal modifiers before target nouns (predicted accuracy 63%, 95% CI [58%, 70%]) vs. non-premodified targets (predicted accuracy 70%, 95% CI [65%, 74%]). This pattern holds for both L1-types and is not affected by the intervention.

**Abstractness.** Learners are predicted to be less accurate on GJT items with abstract target nouns (predicted accuracy 65%, 95% CI [60%, 69%]) vs. concrete target nouns (predicted accuracy 69%, 95% CI [64%, 73%]), although the significance level is marginal at  $p = 0.046$ .

**Specificity.** I explored the effect of specificity, as defined in this study, in the separate models for singular indefinites and for mass and plural indefinites, as I expected a potential interaction between the effect of specificity and the effect of error type. However, the effect of specificity was not statistically significant in either of the two models. This can be seen in the model selection reports in Appendix L and Appendix M.

**Delayed Post-Test Results.** In order to confirm that the Experimental group retained its gain in accuracy on the GJT by the delayed post-test, as appears to be the case from Figure 32, I ran a separate model on the subgroup of the 21 participants who completed the delayed post-test.

As this group is much smaller than the entire sample, it is not surprising that it is not as well-balanced in terms of the split between the Control ( $n = 8$ ) and the Experimental group ( $n = 13$ ). The main differences are that the delayed subgroup is almost exclusively female (only one male participant in the Experimental group) and only contains two [+art] L1 learners in the Control group. The Control group also seems to be older on average ( $M = 40.4$ ,  $SD = 11.7$ ) than the Experimental group ( $M = 30.2$ ,  $SD = 8.7$ ), but the difference is marginally significant,  $t(12) = 2.13$ ,  $p = 0.055$ .

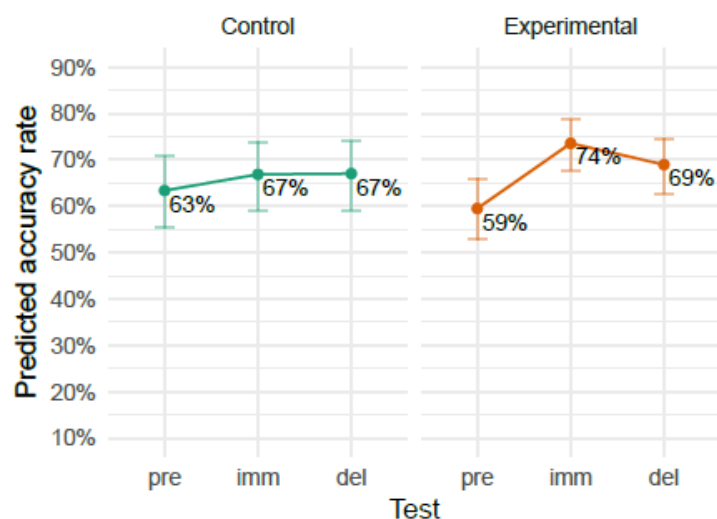
The results of the regression model on the delayed subgroup are presented in Table 19. Despite the imbalance described above, most of the effects found in the GJT Model seem to hold for this subgroup. The only difference is that there is no significant effect of abstractness (which was already only marginally significant for the entire sample).

The main finding is that the Experimental group remained significantly more accurate at the delayed post-test than at the pre-test ( $p = 0.0024$ ), while the Control group's performance did not change (Figure 37). The other effects, i.e. the effect of item variant, definiteness, noun type, L1-type, proficiency, and modifier presence, are similar to those found in the GJT Model.

**Figure 37**

*Predicted Accuracy Rates at Pre-, Immediate Post-, and Delayed Post-Tests for the Delayed Subgroup*

*(n = 21), with 95% Confidence Intervals*

**Table 19**

*Delayed Subgroup GJT Model Results*

| <i>Predictors</i>                                   | <b>Accuracy Rate</b> |                   |               |
|---|----------------------|-------------------|---------------|
|   | <i>Log-Odds</i>      | <i>std. Error</i> | <i>95% CI</i> |
| (Intercept)   | 2.24 ***             | 0.55              | 1.17   3.31   |
| group: Experimental                                 | -0.16                | 0.19              | -0.53   0.21  |
| test: immediate post                                | 0.15                 | 0.17              | -0.18   0.48  |
| test: delayed post                                  | 0.16                 | 0.17              | -0.17   0.49  |
| c-test  | 0.63 ***             | 0.07              | 0.48   0.77   |
| def: indefinite                                     | -0.87                | 0.55              | -1.94   0.20  |
| Ntype: singular                                     | -0.57                | 0.70              | -1.94   0.80  |
| Ntype: mass   | -0.36                | 0.71              | -1.75   1.04  |
| L1-type: [-art]                                     | -0.91 *              | 0.46              | -1.81   -0.01 |
| item variant: IC                                    | -2.67 ***            | 0.51              | -3.68   -1.67 |
| mod: modified                                       | -0.25 *              | 0.10              | -0.45   -0.04 |
| group: Experimental * test: immediate post          | 0.48 *               | 0.22              | 0.05   0.92   |
| group: Experimental * test: delayed post            | 0.25                 | 0.22              | -0.18   0.68  |
| def: indefinite * Ntype: singular                   | 1.46                 | 0.75              | -0.00   2.93  |
| def: indefinite * Ntype: mass                       | 0.53                 | 0.74              | -0.92   1.98  |
| def: indefinite * L1-type: [-art]                   | 1.62 ***             | 0.48              | 0.69   2.56   |
| def: indefinite * item variant: IC                  | 1.75 **              | 0.54              | 0.69   2.82   |
| Ntype: singular * L1-type: [-art]                   | 0.59                 | 0.61              | -0.60   1.77  |
| Ntype: mass * L1-type: [-art]                       | 0.64                 | 0.64              | -0.61   1.89  |
| Ntype: singular * item variant: IC                  | 0.55                 | 0.69              | -0.81   1.91  |
| Ntype: mass * item variant: IC                      | 0.55                 | 0.70              | -0.82   1.92  |
| def: indefinite * Ntype: singular * L1-type: [-art] | -1.83 **             | 0.65              | -3.11   -0.55 |

| <i>Predictors</i>   | <b>Accuracy Rate</b>   |                   |               |
|---|--|-------------------|---------------|
|   | <i>Log-Odds</i>  | <i>std. Error</i> | <i>95% CI</i> |
| def: indefinite * Ntype: mass * L1-type: [-art]                             | -1.40 *  | 0.66              | -2.68   -0.11 |
| def: indefinite * Ntype: singular * item variant: IC                        | -2.16 **   | 0.75              | -3.63   -0.69 |
| def: indefinite * Ntype: mass * item variant: IC                            | -1.71 *  | 0.75              | -3.19   -0.24 |
| <b>Random Effects</b>   |  |                   |               |
| <i>By item ID</i>   |  |                   |               |
| SD (Intercept)  | 0.56   |                   |               |
| <i>By participant ID</i>  |  |                   |               |
| SD (Intercept)  | 0.12   |                   |               |
| SD (item variant: IC)   | 0.46   |                   |               |
| SD (Ntype: singular)  | 0.14   |                   |               |
| SD (Ntype: mass)  | 0.38   |                   |               |
| N <sub>ppt_id</sub>   | 21   |                   |               |
| N <sub>item_id</sub>  | 348  |                   |               |
| Observations  | 3538   |                   |               |
| Marginal R <sup>2</sup> / Conditional R <sup>2</sup>                        | 0.322 / 0.381  |                   |               |
| Log-likelihood ratio test comparing to the null model (only random effects) | $\chi^2 (26) = 240.71, p < 10^{-15}$   |                   |               |
| C-statistic   | 0.85 (strong predictive power)   |                   |               |
| VIFs  | High in a model with correlations but low in a no-correlation model (1.00 for all variables) |                   |               |
| Overdispersion ratio  | 0.910 ( $\chi^2 = 3192.176, p = 1$ )   |                   |               |

Note. \*  $p < 0.05$  \*\*  $p < 0.01$  \*\*\*  $p < 0.001$ . The omitted levels of multi-level variables in this model

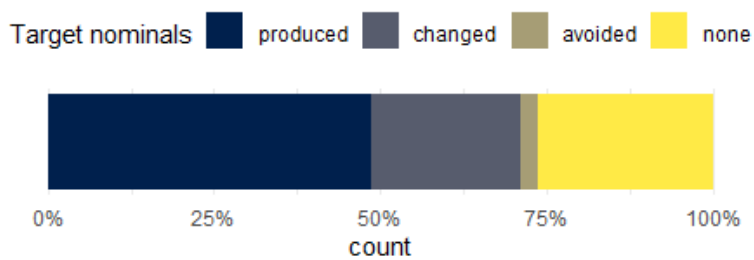
are

“pre-test” for test, “plural” for noun type (Ntype).

## **EI Results**

**Descriptive Statistics.** Before considering the accuracy scores on the EI task, I checked whether the participants were able to produce a sufficient proportion of intended targets. As Figure 38 demonstrates, almost 50% of targets were produced exactly as they were heard in the prompts, and a further 20% were changed in ways that did not affect article use (e.g. put in a different syntactic position or substituted with a synonym), bringing the total proportion of scorable items to about 70%. This is a considerable improvement from the pilot where participants were only able to produce about 40-45% of intended targets on average.



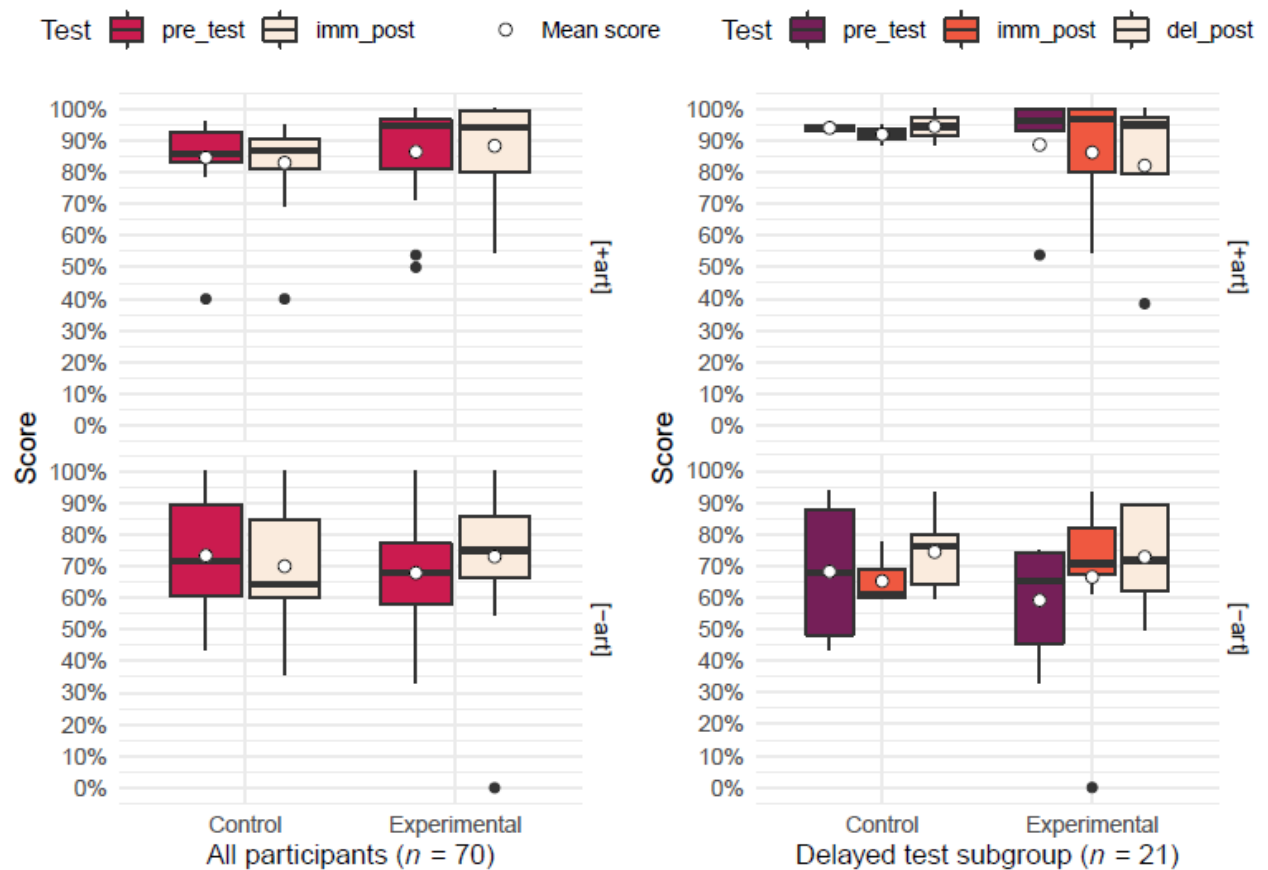
**Figure 38***Production of Target Nominals on the EI Task*

The overall scores for the EI task at pre-test, immediate post-test, and delayed post-test, split by L1-type are shown in Figure 39 (note that the right-hand side shows data only from the 21 participants who completed the delayed post-test). The first observation in Figure 39 is that there is a larger gap in performance between the [-art] and the [+art] groups compared to the GJT results, with [+art] participants' scores approaching ceiling, which means one would not expect much gain from the Experimental [+art] group.

The second observation is that the improvement of the Experimental [-art] group is not as obvious as it appeared on the GJT (cf. Figure 32). In fact, when I started building the regression model for EI and added "group" and "test" variables (with an interaction between the two) to the null model, this did not significantly improve model fit ( $\chi^2(3) = 6.93, p = 0.07$ ), although the  $\beta$ -coefficient for the interaction was significant at  $p = 0.016$ . Thus, I took a different approach to model selection in this case and, instead of the null model, started with a model which included all the variables that, based on the GJT results, I expected to have an effect (without interactions). Indeed, accounting for other variables allowed the model (henceforth, the EI Model) to identify a significant effect of the intervention, which is reported in more detail below. Nevertheless, this implies that the effect is not as robust as the one observed on the GJT.

**Figure 39**

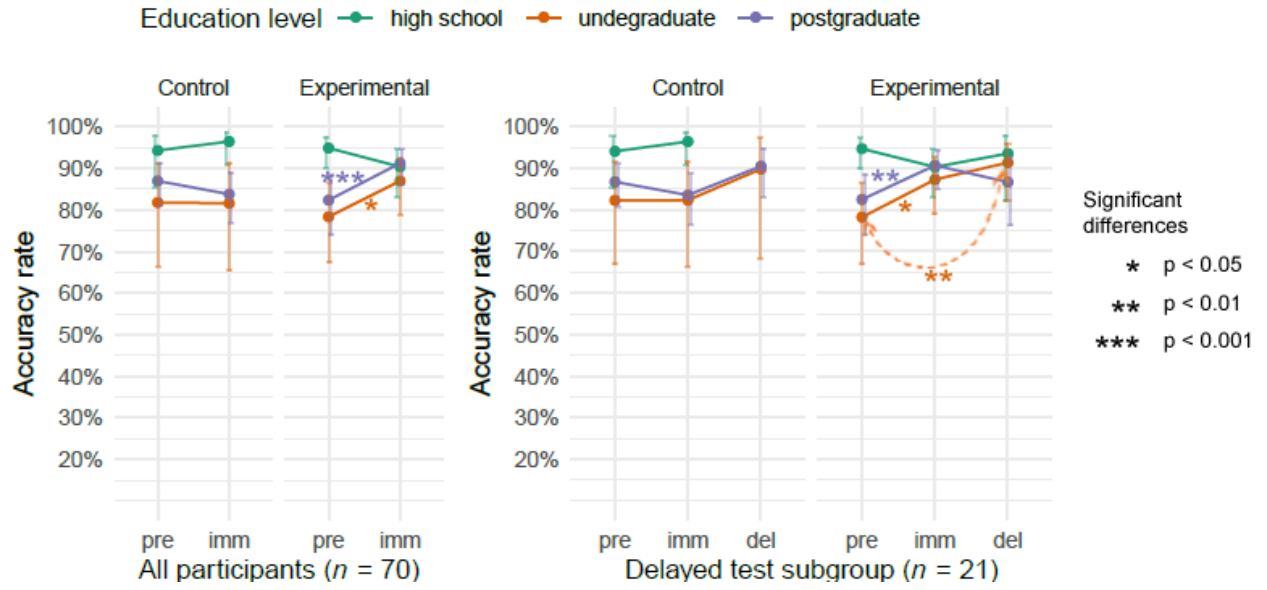
*The Overall Effect of the Intervention on Article Accuracy on the EI Task by L1-type*



**The Effect of the Intervention.** The full results of the EI Model along with model performance metrics are provided in Table 20. The model estimates that the Experimental group improved significantly from pre-test to immediate post-test, however this effect was mediated by the participants' education level, as seen in Figure 40 (note that the right-hand side graph is based on the delayed group subsample containing 21 participants). More specifically, the change was most prominent in those who already had a postgraduate degree (Master's or PhD), reaching statistical significance at  $p = 0.0006$ . The improvement of those with an undergraduate degree was barely significant at  $p = 0.028$ . The rest of the participants, whose highest education level was high school, were already above 90% accurate at pre-test (in both Experimental and Control groups), so, unsurprisingly, their performance did not change significantly at the immediate post-test. The Control group participants' scores did not change significantly for any education level.

**Figure 40**

*Predicted Effect of Intervention on EI Accuracy Rate by Education Level (with 95% Confidence Intervals)*

**Table 20**

*EI Model Results*

| <i>Predictors</i>   | Accuracy Rate   |                   |               |
|---|-----------------|-------------------|---------------|
|   | <i>Log-Odds</i> | <i>std. Error</i> | <i>95% CI</i> |
| (Intercept)   | 4.37 ***        | 0.65              | 3.10   5.65   |
| group: Experimental   | 0.11            | 0.60              | -1.07   1.30  |
| test: immediate post  | 0.50            | 0.57              | -0.63   1.63  |
| education: undergraduate  | -1.31 *         | 0.65              | -2.58   -0.03 |
| education: postgraduate   | -0.91           | 0.54              | -1.97   0.15  |
| c-test  | 0.58 ***        | 0.09              | 0.39   0.77   |
| def: indefinite   | -1.04 *         | 0.42              | -1.87   -0.22 |
| L1-type: [-art]   | -2.28 ***       | 0.39              | -3.05   -1.51 |
| abstr: abstract   | -0.57 *         | 0.22              | -1.00   -0.13 |
| group: Experimental * test: immediate post                            | -1.18           | 0.67              | -2.50   0.14  |
| group: Experimental * education: undergraduate                        | -0.33           | 0.77              | -1.84   1.18  |
| group: Experimental * education: postgraduate                         | -0.47           | 0.66              | -1.77   0.83  |
| test: immediate post * education: undergraduate                       | -0.51           | 0.72              | -1.93   0.91  |
| test: immediate post * education: postgraduate                        | -0.76           | 0.61              | -1.94   0.43  |
| def: indefinite * L1-type: [-art]                                     | 1.53 ***        | 0.42              | 0.70   2.36   |
| group: Experimental * test: immediate post * education: undergraduate | 1.81 *          | 0.85              | 0.14   3.47   |
| group: Experimental * test: immediate post * education: postgraduate  | 2.24 **         | 0.74              | 0.80   3.69   |
| <b>Random Effects</b>   |                 |                   |               |
| <i>By item ID</i>   |                 |                   |               |
| SD (Intercept)  | 0.83            |                   |               |
| SD (L1-type: [-art])  | 0.52            |                   |               |

| <i>Predictors</i>   | <b>Accuracy Rate</b>  |                   |               |
|---|---|-------------------|---------------|
|   | <i>Log-Odds</i>   | <i>std. Error</i> | <i>95% CI</i> |
| <i>By participant ID</i>  |   |                   |               |
| SD (Intercept)  | 0.31  |                   |               |
| SD (def: indefinite)  | 0.47  |                   |               |
| SD (Ntype: singular)  | 0.36  |                   |               |
| SD (Ntype: mass)  | 0.63  |                   |               |
| N <sub>ppt_id</sub>   | 70  |                   |               |
| N <sub>item_id</sub>  | 96  |                   |               |
| Observations  | 2659  |                   |               |
| Marginal R <sup>2</sup> / Conditional R <sup>2</sup>                        | 0.242 / 0.242   |                   |               |
| Log-likelihood ratio test comparing to the null model (only random effects) | $\chi^2 (16) = 101.86, p < 10^{-13}$  |                   |               |
| C-statistic   | 0.86 (strong predictive power)  |                   |               |
| VIFs  | High in a model with correlations but low in a no-correlation model (< 1.5 for all variables) |                   |               |
| Overdispersion ratio  | 0.703 ( $\chi^2 = 1854.275, p = 1$ )  |                   |               |

Note. \*  $p < 0.05$  \*\*  $p < 0.01$  \*\*\*  $p < 0.001$ . The omitted level of the Education Level multi-level variable in this model is “high school”.

The same model specification appeared to be the best fit for the delayed subgroup (Figure 40, right<sup>36</sup>), which implies that the effect of the intervention was overall maintained by the delayed post-test for those participants who took it ( $n = 21$ ), although the results of the postgraduate group dropped slightly making the difference between the pre-test and the delayed post-test insignificant. The undergraduate group, on the other hand, improved further by delayed post-test. The results of the model fitted on the delayed subgroup are presented in Table 21.

**Table 21**

*Delayed Subgroup EI Model Results*

| <i>Predictors</i>        | <b>Accuracy Rate</b> |                   |               |
|--------------------------|----------------------|-------------------|---------------|
|                          | <i>Log-Odds</i>      | <i>std. Error</i> | <i>95% CI</i> |
| (Intercept)              | 4.30 ***             | 0.63              | 3.06   5.54   |
| group: Experimental      | 0.12                 | 0.60              | -1.06   1.29  |
| test: immediate post     | 0.53                 | 0.55              | -0.55   1.61  |
| test: delayed post       | -0.16                | 0.76              | -1.64   1.33  |
| education: undergraduate | -1.24                | 0.64              | -2.50   0.03  |
| education: postgraduate  | -0.89                | 0.53              | -1.94   0.16  |

<sup>36</sup> Note that none of the high school level participants from the Control group took the delayed post-test.

|   |   |      |               |
|---|---|------|---------------|
| c-test  | 0.63 ***  | 0.10 | 0.44   0.82   |
| def: indefinite   | -1.00 *   | 0.39 | -1.77   -0.23 |
| L1-type: [-art]   | -2.28 ***   | 0.37 | -3.01   -1.55 |
| abstr: abstract   | -0.55 *   | 0.22 | -0.98   -0.12 |
| group: Experimental * test: immediate post                                  | -1.18   | 0.65 | -2.45   0.09  |
| group: Experimental * test: delayed post                                    | -0.05   | 0.45 | -0.94   0.84  |
| group: Experimental * education: undergraduate                              | -0.37   | 0.77 | -1.89   1.14  |
| group: Experimental * education: postgraduate                               | -0.44   | 0.66 | -1.74   0.86  |
| test: immediate post * education: undergraduate                             | -0.52   | 0.70 | -1.90   0.85  |
| test: delayed post * education: undergraduate                               | 0.80  | 1.02 | -1.20   2.80  |
| test: immediate post * education: postgraduate                              | -0.78   | 0.58 | -1.93   0.36  |
| test: delayed post * education: postgraduate                                | 0.53  | 0.69 | -0.82   1.88  |
| def: indefinite * L1-type: [-art]   | 1.55 ***  | 0.39 | 0.79   2.31   |
| group: Experimental * test: immediate post * education: undergraduate       | 1.83 *  | 0.83 | 0.21   3.45   |
| group: Experimental * test: delayed post * education: undergraduate         | 0.49  | 0.92 | -1.31   2.29  |
| group: Experimental * test: immediate post * education: postgraduate        | 2.16 **   | 0.71 | 0.77   3.56   |
| <b>Random Effects</b>   |   |      |               |
| <i>By item ID</i>   |   |      |               |
| SD (Intercept)  | 0.84  |      |               |
| SD (L1-type: [-art])  | 0.53  |      |               |
| <i>By participant ID</i>  |   |      |               |
| SD (Intercept)  | 0.42  |      |               |
| SD (def: indefinite)  | 0.40  |      |               |
| SD (Ntype: singular)  | 0.36  |      |               |
| SD (Ntype: mass)  | 0.59  |      |               |
| N <sub>ppt_id</sub>   | 70  |      |               |
| N <sub>item_id</sub>  | 96  |      |               |
| Observations  | 3076  |      |               |
| Marginal R <sup>2</sup> / Conditional R <sup>2</sup>                        | 0.246 / 0.246   |      |               |
| Log-likelihood ratio test comparing to the null model (only random effects) | $\chi^2(21) = 107.96, p < 10^{-12}$   |      |               |
| C-statistic   | 0.86 (strong predictive power)  |      |               |
| VIFs  | High in a model with correlations but low in a no-correlation model (< 1.5 for all variables) |      |               |
| Overdispersion ratio  | 0.714 ( $\chi^2 = 2177.037, p = 1$ )  |      |               |

Note. \*  $p < 0.05$  \*\*  $p < 0.01$  \*\*\*  $p < 0.001$ . The omitted levels of multi-level variables in this model

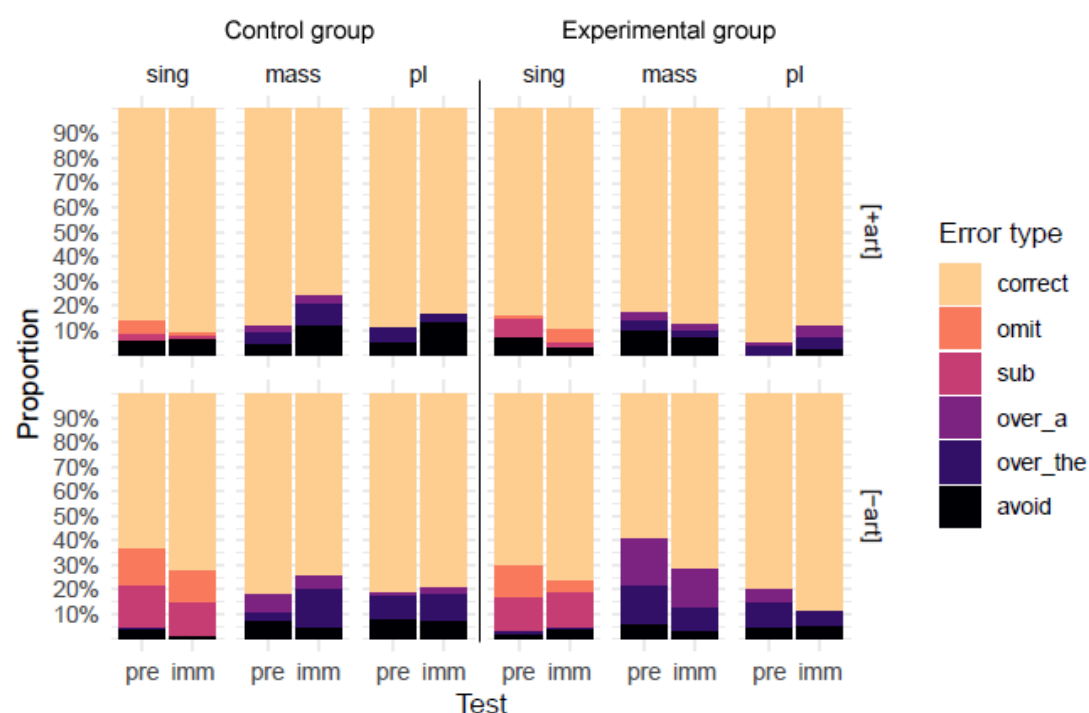
are

“pre-test” for test, “high school” for Education Level.

Visual examination of the data tentatively suggested that there might be differential effect of the intervention on different error types (Figure 41, note that this is for indefinite contexts only<sup>37</sup>). To test whether any of the differences were statistically significant, I used mixed-effects multinomial regression models fitted on singular indefinites and on mass and plural indefinites separately<sup>38</sup>.

**Figure 41**

*Error Types in Indefinite Contexts on the EI Task (Observed Data)*



However, both models were unreliable producing negative  $R^2$  values or impossible random effects estimates. Thus, there is no evidence of any considerable effects of the intervention on particular error types.

**Other Important Variables Not Affected by the Intervention.** A few other predictors had a significant effect on participant accuracy, which did not change after the intervention.

**Proficiency Level.** Figure 42 shows the effect of proficiency level as measured by the c-test. Similarly to the effect on the GJT, higher proficiency predicted higher accuracy on EI, but this effect

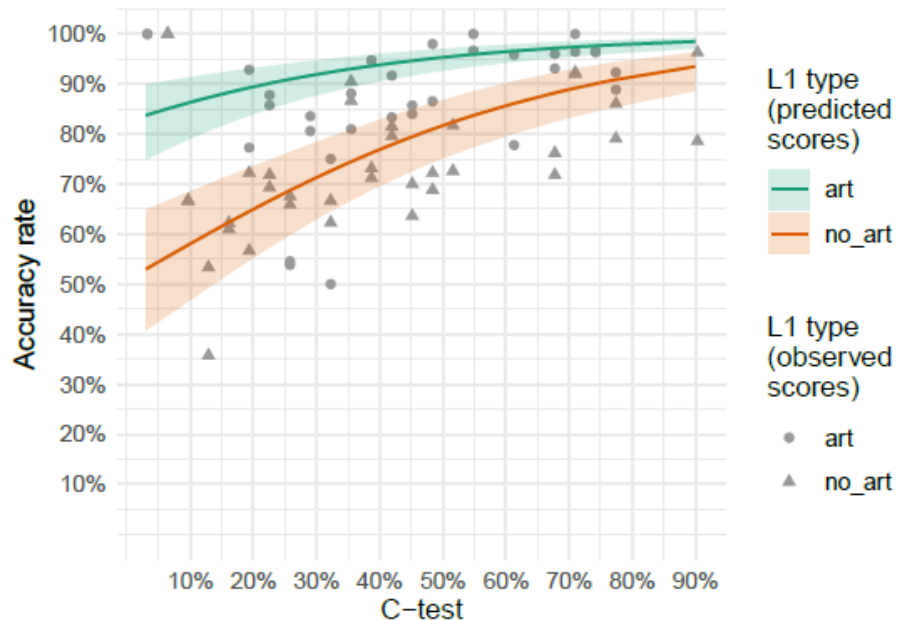
<sup>37</sup> Definite items were too sparse (6 per test) to further subcategorise by error type.

<sup>38</sup> This is because different errors are possible in these contexts. For further explanation of multinomial regression see the Methodology Section in Chapter 3.

did not interact with group or test, meaning that Experimental group participants improved regardless of the initial proficiency level, while the Control group participants did not.

**Figure 42**

*The Effect of Proficiency Level on EI Accuracy Rate (with 95% Confidence Intervals)*



**L1-Type.** The EI task appears to have differentiated between the two L1-types better than the GJT, as a larger gap is seen between [+art] and [-art] L1 learners (cf. Figure 35 vs. Figure 42), which is statistically significant for both definites and indefinites at  $p < 0.001$ .

**Abstractness.** Learners are predicted to be less accurate on EI items with abstract target nouns (predicted accuracy 86%, 95% CI [80%, 90%]) vs. concrete target nouns (predicted accuracy 91%, 95% CI [87%, 94%]) at  $p = 0.011$ . The effect of abstractness is not impacted by the intervention.

**Non-significant variables.** The model estimated that there was no significant effect of specificity, as defined in this study (for indefinite contexts), modifier presence, or noun type.

### **Written Production**

Although the materials were designed to include a written production task and data was collected from all the participants responding to the two writing prompts included in each test, there was not enough time to code and analyse the results of this task within the scope of this dissertation.

## Discussion

### *Key Findings*

The main result of this Processing Instruction Intervention is that it appears to have improved article accuracy for the Experimental group on both GJT and EI, although the effect on the latter task is more limited and less certain. Moreover, the overall effect is to some extent maintained by the delayed post-test three months later (at least for those participants who took the delayed post-test). The Control group participants, on the other hand, did not change significantly in their performance throughout.

Further analysis revealed that the improvement of the Experimental group on the GJT was not uniform but seems to have affected certain contexts more than others. Thus, although the upward trend was seen across both definite and indefinite contexts and all noun types, the gain was statistically significant only for:

- the ability to notice “a” omission, as well as substitution of “a” with “the” in singular indefinites (statistically significant for [-art] L1 learners only),
- the ability to notice “the” omission with mass nouns,
- the ability to notice article overuse (of both “a” and “the”) with mass and plural indefinites.

The first important implication of these findings is that the intervention seems to have positively affected one of the most challenging aspects of article use, i.e. article omission (according to the corpus study in Chapter 3). This is seen for both the omission of “a” and “the” on the GJT.

It is also clear that participants made gains on items involving two different types of contrasts: the number/countability contrast represented by errors of “a” omission and overuse (e.g. “I found [a] bag”, “I like \*a milk”) and the definiteness/indefiniteness contrast represented by substitution errors (e.g. “we are creating \*the [a] new department”), “the” omission and overuse (e.g. “[The] Literature you gave me was useful”, “it took him \*the decades”). The effect appears to be more consistent for indefinite contexts, which may be expected as the use of the indefinite article



involves uninterpretable features, and learners may have formed associations with irrelevant interpretable features (e.g. specificity) to compensate for the lack of regularity in their article use. However, this might also reflect the larger quantity of indefinite items both on the test tasks and in the training materials<sup>39</sup>. The fact that despite this imbalance, there were improvements for mass definites (which were particularly problematic for L1-Russian learners in the corpus study in Chapter 3) is indicative of the potential of Processing Instruction Intervention for targeting all aspects of article use.

Another important result of the intervention is that it was beneficial for both [-art] and [+art] L1 learners, apart from noticing the omission of “a”, which became significantly better only for [-art] L1 participants (although the tendency was the same for [+art] L1 participants). However, this is seen only on the GJT, as [+art] learners were at ceiling on the EI task at the pre-test.

Clearly, the improvements made by the Experimental group on the GJT are sizeable. Nevertheless, it is important to remember that the overall ability to identify ungrammatical items on the GJT is rather low. Even with the gains described above, participants’ accuracy is never predicted to surpass 60%, except for plural indefinites, which is probably due to the prominence of the overuse of “a” (e.g. “\*a dresses”). The first implication of this observation is that learners are generally not good at spotting article errors in the context of a timed GJT. This is, perhaps, evidence that articles are not the primary focus for processing when there is a time constraint, which supports the use of Processing Instruction as an intervention method. The second implication is that learners would likely need to have more practice correctly interpreting input, i.e. making correct form-meaning mappings and, thus, transforming input into intake, before they can achieve higher accuracy levels.

---

<sup>39</sup> This was precisely the reasoning behind increasing the number of indefinite (but not definite) items on tests and in the Experimental group materials: to achieve more robust estimates at least within a subset of data.

### ***The Role of Form-Meaning Mappings and of Cognitive Constraints***

I will first discuss participants' noticing of article omission, where the Experimental group improved significantly for the indefinite article "a" with count singulars and for the definite "the" with mass nouns. For both types of omission, cognitive constraints could be at play, especially for lower levels. However, there was no interaction between proficiency level and group or test time, meaning that regardless of their initial level, all participants in the Experimental group improved at this. In the Experimental group, proficiency level, as measured by the c-test, ranged from below 5% to 90%, with the median just above 35%. Thus, it may be that for most participants, the main reason for their insensitivity to "a" or "the" omission was cognitive constraints. In this case, the intervention has worked for this type of error by directing learners' attention to the article, whereas previously their attention might have been focused on other more meaningful elements. For those at higher levels, abandoning inappropriate form-meaning mappings and establishing new ones is a plausible cause for improvement, although spotting omission still does not exceed 60% in any context, which could mean that a timed GJT places considerable limitations on participants' cognitive resources even at higher levels.

It is noteworthy that for the omission of "a" only [-art] L1 participants made significant gains, whereas for the omission of "the" with mass nouns, this was the case for both [-art] and [+art]. Given that [-art] L1 participants scored lower initially across the board, it is possible that [-art] participants were indeed challenged by the uninterpretable number and countability features on the indefinite article (in addition to the cognitive constraints which applied to everyone), and the form-meaning mapping provided by the intervention enabled them to make quick gains. However, if this was true, I would also expect to see more significant improvement in [-art] than in [+art] participants on spotting "a" overuse with mass and plural indefinites, which was not the case.

Where the results showed no differentiation between [-art] and [+art] L1 participants, the improvements are more likely to stem from the intervention forcing them to allocate more cognitive resources to article processing. This might explain why accuracy increases on error types which

required a definiteness/indefiniteness contrast were not consistent across the board (i.e. improved noticing of “the” omission with mass nouns but not with plurals, no significant changes on spotting the use of “a” instead of “the” with definite count singulars). Deciding whether “the” is appropriate or needed involves making a judgment based on the discourse and context, which is more difficult under the time constraint. By contrast, deciding if “a” is needed in indefinite contexts should become easier for the Experimental group, if the intervention has worked, as they should have learned that they only need to judge the countability and number of the nominal.

In sum, I conclude that the biggest advantage of this PI intervention was that it directed learners’ attention to articles in the input, whereas under normal circumstances and in conditions of limited cognitive resources, such as a timed GJT, they would not have enough capacity to attend to articles. In addition, the intervention may have created new form-meaning mappings helping specifically [-art] L1 learners to spot “a” omission, which is a particular challenge for such learners, as the uninterpretable features of number and countability on articles are not instantiated in their L1s. There is little that can be added to above discussion based on the results of the EI task, as apart from the overall significant improvement of the Experimental group, the statistical analysis was unable to confirm any more detailed patterns.

### ***Testing Practice Effects***

The most robust effect on the GJT is seen in learners’ ability to spot article overuse (of both “a” and “the”) with mass and plural indefinites. However, part of this result is probably explained by the attention raising effect of the test itself, as even the Control group improved significantly on non-premodified items, although the gain was not as large as for the Experimental group. Perhaps, seeing multiple incongruent items, such as “\*a snow”, “\*a dresses”, or “it took him \*the decades”, on the pre-test is sufficient to draw learners’ attention to article overuse on the post-test, at least when an article precedes the noun directly. However, the fact that the Experimental group also improved on premodified items, such as “\*a strong alcohol”, “\*a serious problems”, or “I thought

this seminar would be boring, but we discussed *\*the very interesting concepts actually*", shows an added effect of the intervention.

Within the broad category of article overuse, there are three distinct subcategories, however. First, the overuse of "a" with plural indefinites must be the easiest type of error to notice and is unlikely new knowledge for learners even at lower levels of proficiency (overall, participants scored approximately 70% on the pre-test). The improvement of the Experimental group is probably simply an indication of their increased attention to articles as a result of the intervention (as well as the test training effect, as discussed above).

The next easiest type of error is, perhaps, noticing the overuse of "a" with mass nouns, although this is certainly more challenging, as the distinction between count and mass nouns is much less clear (both semantically and in terms of form) than the distinction between singular and plural. In fact, both groups are predicted to start at approximately 25-30% accuracy for this type of item. Yet, the Experimental group is estimated to improve by ca. 20%, while the Control group only improved by 10% on non-premodified nominals. As the corpus study reported in Chapter 3 has demonstrated, learners are particularly prone to overusing "a" with premodified mass nominals. The fact that the Experimental group became significantly better at recognizing this as an error after the intervention is rather promising.

The third category of article overuse items is the overuse of "the" with mass and plural indefinites. This error type is quite distinct from the first two, as it involves a discourse-semantic definiteness/indefiniteness contrast rather than the more grammatical number/countability contrast. Learners' ability to notice the overuse of "the" is very low to begin with. Nevertheless, the Experimental group is predicted to improve by 10-15% after the intervention. The improvement of the Experimental group here clearly surpasses that of the Control group, as it is both larger and encompasses all types of items, premodified or not. It is still less clear how the Control group was able to become better at noticing overuse of "the" with non-premodified nouns simply from

exposure to such items as “it took him \*the decades” on the pre-test. The testing practice effect is likely still relevant, but this explanation is much less convincing for this error type.

### ***No Effect of Specificity on GJT or EI***

One finding that may seem striking when compared to the results of the corpus study in Chapter 3 is that no effect of specificity, as reference to a certain existing entity, was identified in any of the models for either the GJT or EI.

A key difference between the data used in the corpus study and the data collected in this experiment is that the former was pure production. In other words, the learner writing corpus contained writings where both content and language were generated by learners. Firstly, learners are probably not nearly as engaged with the meaning of GJT items as they are with the meaning of their own writing.

Secondly, it is possible that specificity plays an important role in production because learner writers are committed to building a discourse in their own production. For example, when a learner writes “I completed an English course”, they are referring to something real that happened to them. By contrast, items on a GJT can be perceived out of contexts. When judging the grammaticality or acceptability of “I completed an English course” (intended as specific) or “I would like to complete an English course” (intended as non-specific), specificity may be of little importance to the learner, as the “English course” is not something they experienced either way. The EI task might have a similar effect, although it is technically an elicited production task, since the content of EI items is given and not generated by learners.

Another possible explanation is related to the Interpretability Hypothesis, according to which learners at higher levels of proficiency are likely to try to compensate for any irregularity in their interlanguage grammar by associating forms to other interpretable features that seem relevant and are available. In the case of articles, as I mentioned elsewhere, specificity could be one such interpretable feature that more advanced learners might employ to regularise their use of the indefinite article, which contains uninterpretable features of number and countability deemed

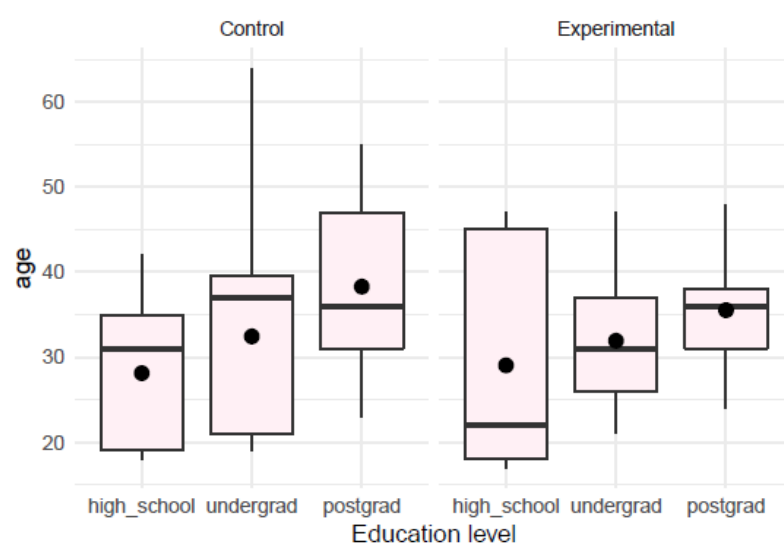
inaccessible to learners from [-art] L1s. Therefore, the effect of specificity, as defined in this study, may be lacking due to the generally lower proficiency level of the participants, who are perhaps not yet resorting to such compensatory strategies. The implications for further research are discussed in the “Limitations” subsection below.

### ***The Mitigating Effect of Education Level on EI***

The effect of the intervention on the EI task was not as robust as that found on the GJT. This is probably because it was only significant for part of the Experimental group, namely those participants who had a degree (especially postgraduate), as they were the ones who improved. Meanwhile, those participants who had only finished high school started at ceiling and, thus, had no room for improvement.

**Figure 43**

*Correlation Between Age and Education Level of Participants*



Considering the nature of the task (reproducing a long complex sentence after hearing it once), it is possible that there was an effect of working memory, since there is a visible correlation between education level and age (Figure 43). Perhaps, younger participants who had not obtained a degree yet, had better working memory and were able to remember the EI prompts word for word (at least better than the older participants), which enabled them to score over 90% on the pre-test

and the immediate post-test. In addition, similar tasks might be more commonly practised in English classes at school.

In this case, however, one would expect that the effect of education level should be replaceable with the effect of age, but this is not the case. Adding an interaction between group, test, and age in addition to the existing interaction between group, test, and education level did not improve the model significantly. Moreover, removing education level entirely did not render the group, test, and age interaction significant either. One explanation for this is probably that a relatively small number of data points does not allow the model to identify an effect of the continuous age variable, whereas lumping data into just three bins makes the effect appear clearer.

### ***Limitations***

One of the biggest limitations of the intervention experiment is that all the conclusions are based mainly on the results of the timed GJT, especially the more nuanced discussion of the effects on different error types and [-art] vs. [+art] L1 participants. The results of the EI task are potentially compatible with those of the GJT and certainly do not seem to contradict them, but the statistical modelling was less informative due to the limitations of the dataset (fewer items, more variability), so there is little one could generalise from the EI task, except that the Experimental group overall significantly improved, and the Control group did not, and that the gains were not as considerable as on the GJT in terms of size.

However, neither GJT nor EI give much indication of whether article accuracy gains would be transferred to free production. Both GJT and EI involve processing given input, and although EI also has an element of production, or rather reconstruction, it does not tap into the same processes as free production. The latter is more cognitively demanding, since the message to be conveyed is not given but needs to be constructed before it is encoded in linguistic forms, and there are more aspects the learner needs to attend to, specifically retrieving and producing the more meaningful linguistic forms leaving even less cognitive resource for articles. Given the results of the GJT and EI, I would expect any gains in writing production to be less sizeable, and I would expect that higher-level

participants, who should generally have more cognitive capacity available, would improve more significantly than lower-level participants. The writing tasks which were included in the tests for this intervention experiment could have supported or debunked these hypotheses; however, the results could not be analysed in time to be included in this dissertation. This would be an important avenue for further research.

With this in mind, it is worth reconsidering the issue of participants' proficiency level. On the one hand, lower-level learners, who constituted the bulk of the participants in this study, clearly benefited from the intervention. As mentioned above, the effect could be largely due to the directing of participants' attention to articles in the input. However, article errors are known to persist into advanced level of proficiency, so it would be useful from an empirical point of view to have a larger group of high- and very high-level learners, as they are more likely to have already formed those inappropriate form-meaning mappings that PI intervention would be so well suited to amend. For this experiment, several practical considerations, such as the switch to online participant recruitment and the requirement for large numbers due to many variables I wanted to be able to analyse statistically, made it difficult to focus on advanced learners only. However, this would be another important question for further research.

Another limitation is the absence of a second Control group completing conventional article practice activities typically found in English textbooks. This would allow more certainty in concluding that whatever effects were observed were likely due to the Processing Instruction intervention rather than due to simply doing any kind of article practice. Although the study would clearly have benefited from an additional control group, it is, perhaps, less critical in light of my main purpose, which was to investigate whether article accuracy could be improved by Processing Instruction intervention structured around relevant linguistic contrasts. I was not primarily concerned with showing that Processing Instruction is the best kind of intervention, as many studies have already demonstrated its benefits for a variety of features (Cadierno, 1995; VanPatten & Cadierno, 1993; VanPatten & Uludag, 2011).



Another serious limitation is that I was not able to recruit the optimal number of participants as identified by the power analysis. Even with the increase in the number of test items, it would probably be necessary to include at least twice as many participants to achieve reasonable statistical power. This means that the experiment is almost certainly underpowered, i.e. there is a risk that some true effects were not identified or that some statistically significant effects identified in this experiment were overestimated.

## Chapter 7: General Discussion

### Summary of the Findings

The empirical findings of this thesis come from two studies. The first study explored the impact of several linguistic factors on article accuracy in learner English based on a large corpus of learner writing. The analysis focused on establishing the relative importance of the different factors and the interactions between them. The results confirmed an already established L1-effect, i.e. lower overall accuracy for [-art] L1 learners, but also revealed a complex pattern of other effects.

The main finding was the differential effect of specificity, as defined in this study, on learner accuracy in definite and indefinite contexts. In indefinite contexts, learners appear to associate “a” with specific indefinites, which leads to “a” omission with non-specific singular indefinites and “a” overuse with specific mass indefinites. Moreover, the overuse of “a” in the latter category is increased when the head noun is premodified, prompting me to suggest that prenominal modification may contribute to a count interpretation of a nominal and possibly makes a referent appear more specific. In definite contexts, however, which are mostly specific, prenominal modifiers have the opposite effect making the omission of “the” more likely. This implies that “the” may be perceived by learners as an optional modifier, which can appear redundant in the presence of other identifying elements, such as adjectives or noun adjuncts. Additionally, “the” omission is more common for definites in subject position, potentially because in this case the referent is likely the topic/theme and is, thus, sufficiently identifiable without an article.

The findings of the corpus study were used to create a teaching intervention aimed at improving learner article accuracy. I hypothesised that the apparent form-meaning mappings characterising learners’ written production summarised above were indicative of input processing biases. Specifically, learners appear to be driven by meaning, as predicted by the lexical preference Input Processing principle (VanPatten, 2002), and to ignore relevant structural cues, such as countability and number. To address those issues, Processing Instruction was chosen as the basis for the development of the intervention materials. In addition to following the principles of Processing

Instruction, these materials were constructed around the linguistic contrasts that were identified as problematic in the corpus study.

The results of the intervention demonstrated that this approach had clear benefits for learners, although these were more obvious on the GJT than on the EI task. The Experimental group improved significantly in their ability to spot the overuse of “a” and “the” in mass and plural indefinites (including premodified ones, which were problematic in the corpus data) and the omission of “the” with mass definites (also a very common error for L1-Russians in the corpus study). Moreover, [-art] L1 learners became significantly better at noticing “a” omission in singular indefinites, which was one of their biggest issues according to the corpus study. Thus, the intervention has shown generally positive results across multiple contexts and error types.

### **Implications for Intervention Studies in SLA**

This study has shown the benefit of combining a detailed linguistic analysis of a feature in learner L2 using corpus-based error analysis with a developmental perspective on SLA, such as input processing, in order to create practical value for learning and teaching.

Most previous intervention studies focused on limited aspects of article use, such as first-mention use of “a” and anaphoric use of “the”, ignoring the effect of important linguistic variables and often lumping together the results across definites and indefinites and different noun types (Bitchener, 2008; Ellis et al., 2008; Muranoi, 2000; Sheen, 2007). Those studies which did incorporate multiple linguistic factors, such as countability, number, specificity, resulted in very little or no improvement (Master, 1994; Snape & Yusa, 2013; Umeda et al., 2017).

It appears from the results of this thesis that careful consideration of the relevant linguistic variables and their incorporation in both the experimental materials and tests can significantly improve the effectiveness of interventions driven by developmental approaches.

### **Acquisition of Articles in L2 English**

The basic hypothesis underlying the development of the intervention study in this thesis was that learner problems with English articles could be traced back to certain processing issues and

inappropriate form-meaning mappings. The use of articles in learners' writing was not random but revealed erroneous mappings between form and meaning, e.g. between "a" and specific indefinite contexts.

Processing Instruction intervention directly addresses such issues by modifying input in ways that steer learners away from incorrect form-meaning mappings and focusing them on relevant structural cues, e.g. showing learners that "a" is required for any indefinite singular nominal, either specific or non-specific. The fact that this type of intervention was effective in improving learner article accuracy across different contexts lends support to the initial hypothesis that English articles are subject to input processing biases.

### **Implications for Learning and Teaching Practice**

The implications of this research for teaching articles in L2 English are two-fold. First, it transpires that article accuracy varies considerably across different contexts. For example, learning to use "the" appropriately with count singular nouns in object position does not necessarily translate to correctly using "the" with mass nouns in subject position. This implies that teaching materials should present article use in a variety of contexts. It should not be expected that appropriate article use will be acquired at the same time across the different contexts, as some are more problematic than others. In line with Processing Instruction guidelines, teachers need to be aware of the common input processing biases that are likely to affect article acquisition.

Second, this study shows that, perhaps, some aspects of English grammar are best practised individually outside of classrooms. Rectifying article processing issues appears to require meticulous practice in interpreting copious amounts of carefully designed input. Doing the activities designed for this intervention in class would be impractical in view of the communicative and task-based focus of most English teaching programmes. The intervention materials used in this experiment do not require any input from the teacher and can be completed in 15-minute increments over an extended period without taking up any teaching time.

A more general implication of this research is that any other aspects of a second language that can be traced back to processing issues and that require Processing Instruction would potentially benefit from a similar approach, where learners could practice with specially designed activities individually, freeing up classroom time for more communicative and content-focused work.

### **Limitations and Further Research**

An important limitation of this thesis is that learners' article processing has not been explored directly. The basic hypothesis about input processing biases that was used in designing the intervention materials comes from indirect evidence, such as the patterns of article use in learner writing. In fact, more recent studies in Processing Instruction have employed direct measures of processing, such as eye-tracking and self-paced reading (Benati, 2020; Chiuchiù & Benati, 2020; Malovrh et al., 2020). Existing studies of article processing have shown that [-art] L1 learners can demonstrate target-like processing of articles in well-formed sentences for purposes of referent resolution when dealing with concrete referents in immediate (visible) situations (Trenkic et al., 2014), although they are not sensitive to article omissions or substitutions in ungrammatical sentences on a self-paced reading task (O'Reilly, 2018). Further research into article acquisition would clearly benefit from extending direct processing measure beyond concrete immediate contexts to confirm the hypotheses based on analysing learner production and help better understand the effects of Processing Instruction.

Another possible direction for future research is to further investigate the impact of discourse-pragmatic factors influencing article use and incorporate it in the intervention. The corpus data employed in this study proved to be relatively poor in terms of the variety of discourse functions. A corpus of interactive language, such as dialogues or picture-dictionary type of activities, may be more appropriate for an in-depth analysis of this aspect of article use. If this aspect was to be included in an intervention experiment similar to the one designed for this thesis, it would probably be best to separate the materials targeting indefinite ("a" vs.  $\emptyset$ ) and definite ("the" vs. "a"/ $\emptyset$ ) article uses into two different interventions or two different stages of an intervention. This

would allow to include more of the relevant items on the pre- and post-tests without making the tests too onerous for participants.

The conclusions from this study would also be better supported if the results of the Experimental group in the intervention study could be compared to the results of another Control group completing traditional article-focused practice activities from EFL textbooks. Including another Control group was unfeasible in this project but could be a worthwhile avenue for future research.

Finally, further development of Processing Instruction intervention materials targeting articles should be extended to include affective structured input activities, which were not included in this study, but which are considered an important component of Processing Instruction (VanPatten, 2015). The fact that this intervention had a positive effect on learner article accuracy even without this component suggests that this could further improve learning outcomes.

## References

- Alexopoulou, T., Geertzen, J., Korhonen, A., & Meurers, D. (2015). Exploring big educational learner corpora for SLA research: Perspectives on relative clauses. *International Journal of Learner Corpus Research*, 1(1), 96–129. <https://doi.org/10.1075/ijlcr.1.1.04ale>
- Alexopoulou, T., Michel, M., Murakami, A., & Meurers, D. (2017). Task Effects on Linguistic Complexity and Accuracy: A Large-Scale Learner Corpus Analysis Employing Natural Language Processing Techniques. *Language Learning*, 67(1), 180–208. <https://doi.org/10.1111/lang.12232>
- Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). *Parsimonious Mixed Models*. 2000.
- Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Benati, A. (2005). The effects of processing instruction, traditional instruction and meaning - Output instruction on the acquisition of the English past simple tense. *Language Teaching Research*, 9(1), 67–93. <https://doi.org/10.1191/1362168805lr154oa>
- Benati, A. (2020). An eye-tracking study on the effects of structured input and traditional instruction on the acquisition of English passive forms. *Instructed Second Language Acquisition*, 4(2), 158–179. <https://doi.org/https://doi.org/10.1558/isla.40642>
- Bickerton, D. (1981). *Roots of language*. Karoma.
- Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing*, 17(2), 102–118. <https://doi.org/10.1016/j.jslw.2007.11.004>
- Bolker, B. (2021). *GLMM FAQ*. <http://bbolker.github.io/mixedmodels-misc/glmmFAQ.html>
- Bolker, B. M. (2014). *Convergence error for development version of lme4*. <http://stackoverflow.com/a/21370041/323785>
- Brown, R. (1973). *A first language: The early stage*. Harvard University Press.
- Butler, Y. G. (2002). Second Language Learners' Theories on the Use of English Articles. *Studies in Second Language Acquisition*, 24(03), 451–480. <https://doi.org/10.1017/S0272263102003042>

- Cadierno, T. (1995). Formal instruction from a processing perspective: An investigation into the Spanish past tense. *Modern Language Journal*, 79, 179–193.
- Chen, P. (2004). Identifiability and definiteness in Chinese. *Linguistics*, 42(6), 1129–1184.
- Cheng, A. (1995). *Grammar instruction and input processing: The acquisition of Spanish ser and estar*. University of Illinois at Urbana-Champaign.
- Chiuchiù, G., & Benati, A. (2020). A self-paced-reading study on the effects of structured input and textual enhancement on the acquisition of the Italian subjunctive of doubt. *Instructed Second Language Acquisition*, 4(2), 235–257. <https://doi.org/https://doi.org/10.1558/isla.40659>
- Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.
- Clahsen, H., & Felser, C. (2006). How native-like is non-native language processing? *Trends in Cognitive Sciences*, 10, 564–570.
- Corder, S. P. (1981). *Error Analysis and Interlanguage*. Oxford University Press.
- Cziko, G. A. (1986). Testing the Language Bioprogram Hypothesis: A review of children's acquisition of articles. *Language*, 62(4), 878–898.
- DeKeyser, R. M. (2005). What Makes Learning Second-Language Grammar Difficult? A Review of Issues. *Language Learning*, 55(S1), 1–25. <https://doi.org/10.1111/j.0023-8333.2005.00294.x>
- Dryer, M. S. (2013). *Definite Articles* (M. S. Dryer & M. Haspelmath (Eds.)). The World Atlas of Language Structures Online; Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/37>
- Elff, M. (2021). *mclogit: Multinomial Logit Models, with or without Random Effects or Overdispersion* (0.8.7.3).
- Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language. *Studies in Second Language Acquisition*, 27(2), 141–172.
- Ellis, N. C., & Sagarra, N. (2010). The bounds of adult language acquisition: Blocking and learned attention. *Studies in Second Language Acquisition*, 32(4), 553–580.
- Ellis, R., Sheen, Y., Murakami, M., & Takashima, H. (2008). The effects of focused and unfocused



- written corrective feedback in an English as a foreign language context. *System*, 36(3), 353–371. <https://doi.org/10.1016/j.system.2008.02.001>
- Fillmore, L. W. (1979). Individual differences in second language acquisition. In C. J. Fillmore, D. Kempler, & W. S.-Y. Wang (Eds.), *Individual differences in language ability and language behavior* (pp. 203–228). Academic Press. <https://doi.org/10.1016/b978-0-12-255950-1.50017-2>
- Frege, G. (1960). On sense and reference. In P. Geach & M. Black (Eds.), & M. Black (Trans.), *Translations from the Philosophical Writings of Gottlob Frege* (pp. 56–78). Blackwell.
- García Mayo, M. del P. (2009). Article choice in L2 English by Spanish speakers: Evidence for full transfer. In M. del P. García Mayo & R. Hawkins (Eds.), *Second Language Acquisition of Articles: Empirical findings and theoretical implications* (pp. 13–35). John Benjamins Publishing Company. <https://doi.org/10.1075/lald.49.05pil>
- Geertzen, J., Alexopoulou, T., & Korhonen, A. (2013). Automatic Linguistic Annotation of Large Scale L2 Databases: The EF-Cambridge Open Language Database (EFCamDat). *Selected Proceedings from 31st Second Language Research Forum*.
- Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.
- Green, P., & Macleod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, 7(4), 493–498. <https://doi.org/10.1111/2041-210X.12504>
- Hawkins, J. A. (1978). *Definiteness and Indefiniteness*. Croom Helm.
- Hawkins, J.A. (2004). *Efficiency and complexity in grammars*. Oxford: Oxford University Press.
- Hawkins, R., Al-Eid, S., Almahboob, I., Athanasopoulos, P., Chaengchenkit, R., Hu, J., Rezai, M., Jaensch, C., Jeon, Y., Jiang, A., Leung, Y. 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. <https://doi.org/10.1075/eurosla.6.04haw>
- Heim, I. (2019). Definiteness and indefiniteness. In P. Portner, K. Heusinger, & C. Maienborn (Eds.),

*Semantics - Noun Phrases and Verb Phrases* (pp. 33–69). De Gruyter Mouton.

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed.). Wiley.

Hua, D., & Lee, T. H. (2005). Chinese ESL learners' understanding of the English count-mass distinction. *Proceedings of the 7th Generative Approaches to Second Language Acquisition Conference*, 138–149.

Huang, S. (1999). The emergence of a grammatical category definite article in spoken Chinese. *Journal of Pragmatics*, 31(1), 77–94.

Huebner, T. (1985). System and variability in interlanguage syntax. *Language Learning*, 35(2), 141–163.

Ionin, T., Ko, H., & Wexler, K. (2004). Article semantics in L2 acquisition: The role of specificity. *Language Acquisition*, 12(1), 3–69.

Ionin, T., Zubizarreta, M. L., & Maldonado, S. B. (2008). Sources of linguistic knowledge in the second language acquisition of English articles. *Lingua*, 118(4), 554–576.  
<https://doi.org/10.1016/j.lingua.2006.11.012>

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning*. Springer.

Kao, C. W. (2020). The effect of a digital game-based learning task on the acquisition of the English Article System. *System*, 95(313), 102373. <https://doi.org/10.1016/j.system.2020.102373>

Krifka, M., Pelletier, F. J., Carlson, G. N., ter Meulen, A., Link, G., & Chierchia, G. (1995). Genericity: An introduction. In G. Carlson & F. Pelletier (Eds.), *The Generic Book* (pp. 1–125). University of Chicago Press.

Kroll, J. F., Bobb, S. C., & Hoshino, N. (2014). Two Languages in Mind: Bilingualism as a Tool to Investigate Language, Cognition, and the Brain. *Current Directions in Psychological Science*, 23(3), 159–163. <http://www.jstor.org/stable/44318765>

Lardiere, D. (2009). Some thoughts on the contrastive analysis of features in second language acquisition. *Second Language Research*, 25(2), 173–227.

<https://doi.org/10.1177/0267658308100283>

Lardi re, D. (2004). Knowledge of definiteness despite variable article omission. In A. Brugos, L.

Micciulla, & C. E. Smith (Eds.), *BUCLD 28 proceedings* (pp. 328–339). Cascadilla Press, MA.

Leech, G. N., Rayson, P., & Wilson, A. (2001). *Word frequencies in written and spoken English: Based on the British National Corpus*. Longman.

Li, C. N., & Thompson, S. A. (1981). *Mandarin Chinese: a functional reference grammar*. University of California Press.

Liu, D., & Gleason, J. L. (2002). Acquisition of the article “the” by nonnative speakers of English: An Analysis of Four Nongeneric Uses. *Studies in Second Language Acquisition*, 24, 1–26.

<https://doi.org/10.1017/S0272263102001018>

L udecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). performance: an R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software*, 6(60), 31–39. <https://doi.org/10.21105/joss.03139>

Lyons, C. (1999). *Definiteness*. Cambridge University Press.

Lyons, C. (2009). Defining definiteness. *Definiteness*, 253–281.

<https://doi.org/10.1017/cbo9780511605789.008>

MacWhinney, B. (1997). Second language acquisition and the Competition Model. In A. de Groot & J. F. Kroll (Eds.), *Tutorials in Bilingualism* (pp. 113–142). Lawrence Erlbaum.

Malovrh, P. A., Lee, J. F., Doherty, S., & Nichols, A. (2020). A self-paced reading study of language processing and retention comparing guided induction and deductive instruction: efficiency and depth of processing in computer-assisted learning. *Instructed Second Language Acquisition*, 4(2), 203–234. <https://doi.org/https://doi.org/10.1558/isla.40636>

Maratsos, M. P. (1976). *The use of definite and indefinite reference in young children: an experimental study of semantic acquisition*. Cambridge University Press.

Master, P. (1987). *A cross-linguistic interlanguage analysis of the acquisition of the English article system*. UCLA.

- Master, P. (1994). The effect of systematic instruction on learning the English article system. *Perspectives on Pedagogical Grammar*, 0(May), 229–252.  
<https://doi.org/10.1017/CBO9781139524605.011>
- Master, P. (1997). The English article system: Acquisition, function, and pedagogy. *System*, 25(2), 215–232. [https://doi.org/10.1016/S0346-251X\(97\)00010-9](https://doi.org/10.1016/S0346-251X(97)00010-9)
- Master, P. (2002). Information structure and English article pedagogy. *System*, 30(3), 331–348.  
[https://doi.org/10.1016/S0346-251X\(02\)00018-0](https://doi.org/10.1016/S0346-251X(02)00018-0)
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*, 22(3), 276–282.
- Michel, M., Murakami, A., Alexopoulou, T., & Meurers, D. (2019). Effects of task type on morphosyntactic complexity across proficiency: evidence from a large learner corpus of A1 to C2 writings. *Instructed Second Language Acquisition*, 3(2), 124–152.
- Murakami, A. (2013). *Individual variation and the role of L1 in the L2 development of English grammatical morphemes: Insights from learner corpora*. University of Cambridge.
- Murakami, A. (2016). Modeling Systematicity and Individuality in Nonlinear Second Language Development: The Case of English Grammatical Morphemes. *Language Learning*, 66(4), 834–871. <https://doi.org/10.1111/lang.12166>
- Murakami, A., & Alexopoulou, T. (2016a). L1 influence on the acquisition order of English grammatical morphemes: a learner corpus study. *Studies in Second Language Acquisition*, 38, 365–401. <https://doi.org/10.1017/S0272263115000352>
- Murakami, A., & Alexopoulou, T. (2016b). Longitudinal L2 development of the English article in individual learners. *Proceedings of the 38th Annual Meeting of the Cognitive Science Society*, 1050–1055.
- Muranoi, H. (2000). Focus on form through interaction enhancement: Integrating formal instruction into a communicative task in EFL classrooms. *Language Learning*, 50(4), 617–673.  
<https://doi.org/10.1111/0023-8333.00142>
- Myles, F. (2012). Complexity, Accuracy and Fluency: The Role Played by Formulaic Sequences in Early

- Interlanguage Development. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: complexity, accuracy and fluency in SLA* (pp. 71–94). John Benjamins Publishing Company.
- Nagelkerke, N. J. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, 78(3), 691–692.
- Nation, I. S. P. (2017). *The BNC/COCA Level 6 word family lists* (Version 1.0.0).
- O'Donnell, M. B., Römer, U., & Ellis, N. C. (2013). The development of formulaic sequences in first and second language writing: Investigating effects of frequency, association, and native norm. *International Journal of Corpus Linguistics*, 18(1), 83–108.  
<https://doi.org/10.1075/ijcl.18.1.07odo>
- O'Reilly, J. (2018). *Processing and production of unique and non-unique-to-L2 syntactic structures: The case of English articles and tense-aspect*. University of York.
- Park, J.-H., & Choi, I. (2018). Development of a C-test for Learners of English as a Foreign Language. *Korean Journal of Applied Linguistics*, 34(2), 57–73.  
<https://doi.org/10.17154/kjal.2018.6.34.2.57>
- Pica, T. (1983). Methods of morpheme quantification: their effect on the interpretation of second language data. *Studies in Second Language Acquisition*, 6(1), 69–78.  
<https://doi.org/10.1017/S0272263100000309>
- Qualtrics (No. 03–2020). (2020). Provo.
- R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.
- Rastall, P. (1995). Definite article or no definite article? *English Today*, 11(2), 37–39.
- Reid, J., Battaglia, P., Schuldt, M., Narita, E., Mochizuki, M., & Snape, N. (2006). *The article choice of learners of English as a second language*. University of Essex, UK.
- Robertson, D. (2000). Variability in the use of the English article system by Chinese learners of English. *Second Language Research*, 16(2), 135–172.

<https://doi.org/10.1191/026765800672262975>

- Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *IRAL*, 43, 1–32.
- Russell, B. (1905). On denoting. *Mind*, 14, 479–493.
- Scott, G. G., Keitel, A., Becirspahic, M., Yao, B., & Sereno, S. C. (2019). The Glasgow Norms: Ratings of 5,500 words on nine scales. *Behavior Research Methods*, 51(3), 1258–1270.  
<https://doi.org/10.3758/s13428-018-1099-3>
- Schatz, I. (2020). Lessons from creating a new corpus from an existing large-scale English learner language database. *International Journal of Learner Corpus Research*, 6(2), 220–236.  
<https://doi.org/https://doi.org/10.1075/ijlcr.20009.sha>
- Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly*, 41(2), 255–283.  
<https://doi.org/10.1002/j.1545-7249.2007.tb00059.x>
- Skehan, P. (1998). *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.
- Snape, N. (2008). Resetting the Nominal Mapping Parameter in L2 English: Definite article use and the count–mass distinction. *Bilingualism: Language and Cognition*, 11(01), 63–79.  
<https://doi.org/10.1017/S1366728907003215>
- Snape, N. (2009). Exploring Mandarin Chinese speakers' L2 article use. In N. Snape, Y. I. Leung, & M. Sharwood Smith (Eds.), *Representational deficits in SLA: studies in honor of Roger Hawkins* (pp. 27–52). John Benjamins Publishing Company. <https://doi.org/10.13140/2.1.2866.5604>
- Snape, N., Leung, Y. I., & Ting, H.-C. (2006). Comparing Chinese, Japanese and Spanish speakers in L2 English article acquisition: Evidence against the Fluctuation Hypothesis? *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006)*, 132–139.
- Snape, N., Umeda, M., Wiltshier, J., & Yusa, N. (2015). Teaching the complexities of English article use and choice for generics to L2 Learners. In D. Stringer, J. Garrett, B. Halloran, & S. Mossman (Eds.), *Proceedings of the 13th Generative Approaches to Second Language Acquisition*

- Conference (GASLA 2015)* (pp. 208–222). Cascadilla Proceedings Project.
- Snape, N., & Yusa, N. (2013). Explicit article instruction in definiteness, specificity, genericity and perception. In M. Whong, K.-H. Gil, & H. Marsden (Eds.), *Universal Grammar and the Second Language Classroom* (pp. 161–183). Springer. <https://doi.org/10.1007/978-94-007-6362-3>
- Thomas, M. (1989). The acquisition of English articles by first- and second-language learners. *Applied Psycholinguistics*, 10(3), 335–355. <https://doi.org/10.1017/S0142716400008663>
- Ting, H.-C. (2005). *The acquisition of articles in L2 English by L1 Chinese and L1 Spanish speakers*. University of Essex, UK.
- Trenkic, D. (2007). Variability in second language article production: Beyond the representational deficit vs. processing constraints debate. In *Second Language Research* (Vol. 23, Issue 3). <https://doi.org/10.1177/0267658307077643>
- Trenkic, D. (2008). The representation of English articles in second language grammars: Determiners or adjectives? *Bilingualism: Language and Cognition*, 11(1), 1–18. <https://doi.org/10.1017/S1366728907003185>
- Trenkic, D., & Pongpaiboj, N. (2013). Referent salience affects second language article use. *Bilingualism: Language and Cognition*, 16(1), 152–166. <https://doi.org/10.1017/S1366728912000156>
- Trenkic, D., Mirkovic, J., & Altmann, G. T. M. (2014). Real-time grammar processing by native and non-native speakers: Constructions unique to the second language. *Bilingualism*, 17(2), 237–257. <https://doi.org/10.1017/S1366728913000321>
- Tsimpli, I. M., & Dimitrakopoulou, M. (2007). The interpretability hypothesis: Evidence from wh-interrogatives in second language acquisition. *Second Language Research*, 23(2), 215–242. <https://doi.org/10.1177/0267658307076546>
- Tsimpli, I. M., & Mastropavlou, M. (2007). Feature Interpretability in L2 Acquisition and SLI: Greek Clitics and Determiners. In J. Liceras, H. Zobl, & H. Goodluck (Eds.), *The Role of Formal Features in Second Language Acquisition* (pp. 143–183). Mahwah, NJ: Lawrence Erlbaum.

- Umeda, M., Snape, N., Yusa, N., & Wiltshier, J. (2017). The long-term effect of explicit instruction on learners' knowledge on English articles. *Language Teaching Research*, 1–21.  
<https://doi.org/10.1177/1362168817739648>
- VanPatten, B. (1996). *Input Processing and Grammar Instruction: Theory and Research*. Ablex.
- VanPatten, B. (2002). Processing Instruction: An Update. *Language Learning*, 52(4), 755–803.  
<https://doi.org/10.1111/1467-9922.00203>
- VanPatten, B. (2015). Foundations of processing instruction. In *IRAL - International Review of Applied Linguistics in Language Teaching* (Vol. 53, Issue 2, pp. 91–109). <https://doi.org/10.1515/iral-2015-0005>
- VanPatten, B., & Cadierno, T. (1993). Input Processing and Second Language Acquisition: a role for instruction. *The Modern Language Journal*, 77(1), 45–57. <https://doi.org/10.2307/329557>
- VanPatten, B., & Sanz, C. (1995). From Input to Output: Processing Instruction and Communicative Tasks. In *Second Language Acquisition and Pedagogy*. Erlbaum.
- VanPatten, B., & Uludag, O. (2011). Transfer of training and processing instruction: From input to output. *System*, 39(1), 44–53. <https://doi.org/10.1016/j.system.2011.01.013>
- von Heusinger, K. (2019). Specificity. In P. Portner, K. Heusinger, & C. Maienborn (Eds.), *Semantics - Noun Phrases and Verb Phrases* (pp. 70–111). De Gruyter Mouton.
- Warden, D. A. (1976). The influence of context on children's use of identifying expressions and references. *British Journal of Psychology*, 67(1), 101–112. <https://doi.org/10.1111/j.2044-8295.1976.tb01501.x>
- White, L. (2003). Fossilization in steady state L2 grammars: Persistent problems with inflectional morphology. *Bilingualism: Language and Cognition*, 6(2), 129–141.  
<https://doi.org/10.1017/S1366728903001081>
- Wong, W., & Ito, K. (2018). The effects of processing instruction and traditional instruction on L2 online processing of the causative construction in French: an eye-tracking study. *Studies in Second Language Acquisition*, 40(2), 241–268. <https://doi.org/10.1017/S0272263117000274>



Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge University Press.

<https://doi.org/10.1017/CBO9780511519772>

Zhao, H., & MacWhinney, B. (2018). The instructed learning of form–function mappings in the

English article system. *Modern Language Journal*, 102(1), 99–119.

<https://doi.org/10.1111/modl.12449>

## Appendices

### Appendix A: Full List of Formulaic Sequences Identified in Corpus Data

**Table 22**

*Formulaic Sequences Identified in Corpus Data*

| Formulaic sequence   | Number of occurrences |
|--|-----------------------|
| next/last + temporal modifier (e.g. year, month, time)                     | 70                    |
| kind/type/sort of N (e.g. type of film, kind of car)                       | 34                    |
| for example  | 28                    |
| at/to/after etc. work  | 24                    |
| at/to/after etc. school  | 23                    |
| a/an + temporal modifier (e.g. twice a week)                               | 19                    |
| at/to the beach  | 17                    |
| at/to/after etc. university  | 16                    |
| at night   | 15                    |
| in addition  | 14                    |
| all over/around the world  | 12                    |
| at/from home   | 12                    |
| in the morning   | 11                    |
| one of the + superlative adjective + noun (e.g. one of the biggest issues) | 11                    |
| on/down/to the street(s)   | 10                    |
| at/to/after etc. college   | 9                     |
| take care  | 9                     |
| times + a/an + temporal modifier (3 times a week)                          | 9                     |
| in/to/out of bed   | 8                     |
| by + e-mail/phone etc.   | 8                     |
| for/take a long time   | 8                     |
| have a nice day/trip etc.  | 8                     |
| be/fall in love  | 8                     |
| in the afternoon   | 8                     |
| go for a walk  | 7                     |
| in order to  | 7                     |
| on holiday(s)/vacation   | 7                     |
| in the end   | 6                     |
| in the evening   | 6                     |
| on the phone   | 6                     |
| at the moment  | 5                     |
| at/on the weekend  | 5                     |
| in the long run  | 5                     |
| on Monday/Tuesday etc. afternoon/night/evening                             | 5                     |
| to the movies/theatre/store/zoo  | 5                     |

|  |   |
|--|---|
| work long hours                        | 5 |
| as a result                            | 4 |
| at the same time                       | 4 |
| be part of                             | 4 |
| by the way                             | 4 |
| in/to (the) court                      | 4 |
| day by day                             | 4 |
| from paycheck to paycheck              | 4 |
| from time to time                      | 4 |
| in fact                                | 4 |
| point of view                          | 4 |
| year after year                        | 4 |
| in/at the beginning                    | 3 |
| to/at the cinema                       | 3 |
| have/catch the flu                     | 3 |
| go the extra mile                      | 3 |
| have a good/great time                 | 3 |
| in advance                             | 3 |
| in favour                              | 3 |
| in relation to                         | 3 |
| go to jail                             | 3 |
| (to/under) the ocean                   | 3 |
| on a daily basis                       | 3 |
| on/off stage                           | 3 |
| on top (of that)                       | 3 |
| in/to the sea                          | 3 |
| in (the) summer                        | 3 |
| take part                              | 3 |
| take place                             | 3 |
| in/to town                             | 3 |
| in (the) winter                        | 3 |
| behind schedule                        | 2 |
| focus on/lose sight of the big picture | 2 |
| by car/train etc.                      | 2 |
| day after day                          | 2 |
| day and night                          | 2 |
| do a great job                         | 2 |
| free of charge                         | 2 |
| give a call                            | 2 |
| have fun                               | 2 |
| hear the end of it                     | 2 |
| in case                                | 2 |
| in effect                              | 2 |

|                            |   |
|----------------------------|---|
| in mind                    | 2 |
| in public                  | 2 |
| in the week                | 2 |
| keep in touch              | 2 |
| lose sight                 | 2 |
| the media                  | 2 |
| to/in the mountains        | 2 |
| night after night          | 2 |
| in the forest/woods        | 2 |
| in the old days            | 2 |
| once upon a time           | 2 |
| to/in the park             | 2 |
| pay attention              | 2 |
| the public                 | 2 |
| set the table              | 2 |
| step by step               | 2 |
| take into account          | 2 |
| time by time               | 2 |
| value for money            | 2 |
| see wood for the trees     | 2 |
| any time of the day        | 1 |
| as a rule                  | 1 |
| at a time                  | 1 |
| at ease                    | 1 |
| at limit                   | 1 |
| at odds                    | 1 |
| at present                 | 1 |
| above average              | 1 |
| breach of contract         | 1 |
| by chance                  | 1 |
| by contrast                | 1 |
| by virtue                  | 1 |
| climb up the career ladder | 1 |
| drop a line                | 1 |
| under the earth            | 1 |
| from entry level           | 1 |
| visit the gym              | 1 |
| have a look                | 1 |
| in a way                   | 1 |
| in charge                  | 1 |
| in comparison              | 1 |
| in conclusion              | 1 |
| in contact                 | 1 |

|                             |   |
|-----------------------------|---|
| in debt                     | 1 |
| in line                     | 1 |
| in place                    | 1 |
| in respect of               | 1 |
| in return                   | 1 |
| in tandem                   | 1 |
| in terms of                 | 1 |
| in the thirties             | 1 |
| in the water                | 1 |
| in the workplace            | 1 |
| in total                    | 1 |
| it is time                  | 1 |
| keep an eye                 | 1 |
| last holiday                | 1 |
| make a long story short     | 1 |
| make a difference           | 1 |
| make matters worse          | 1 |
| make sense                  | 1 |
| in the nature               | 1 |
| on foot                     | 1 |
| on the contrary             | 1 |
| on the same page            | 1 |
| on time                     | 1 |
| on track                    | 1 |
| out of line                 | 1 |
| over the years              | 1 |
| over time                   | 1 |
| overstep the mark           | 1 |
| play close to the vest      | 1 |
| pull a face                 | 1 |
| in the rain                 | 1 |
| right of way                | 1 |
| on the road                 | 1 |
| on safari                   | 1 |
| take a moment               | 1 |
| take control                | 1 |
| the moral high ground       | 1 |
| the way the cookie crumbles | 1 |
| the wrong + noun            | 1 |
| tie the knot                | 1 |
| to death                    | 1 |
| top to bottom               | 1 |
| on the web                  | 1 |

|                   |   |
|-------------------|---|
| what a mess       | 1 |
| what's the matter | 1 |
| with regard to    | 1 |
| without question  | 1 |

## Appendix B: Technical Details for Statistical Models

I followed Gelman and Hill's (2007) and Bates et al.'s (2015) recommendations for model selection for all the models in this dissertation.

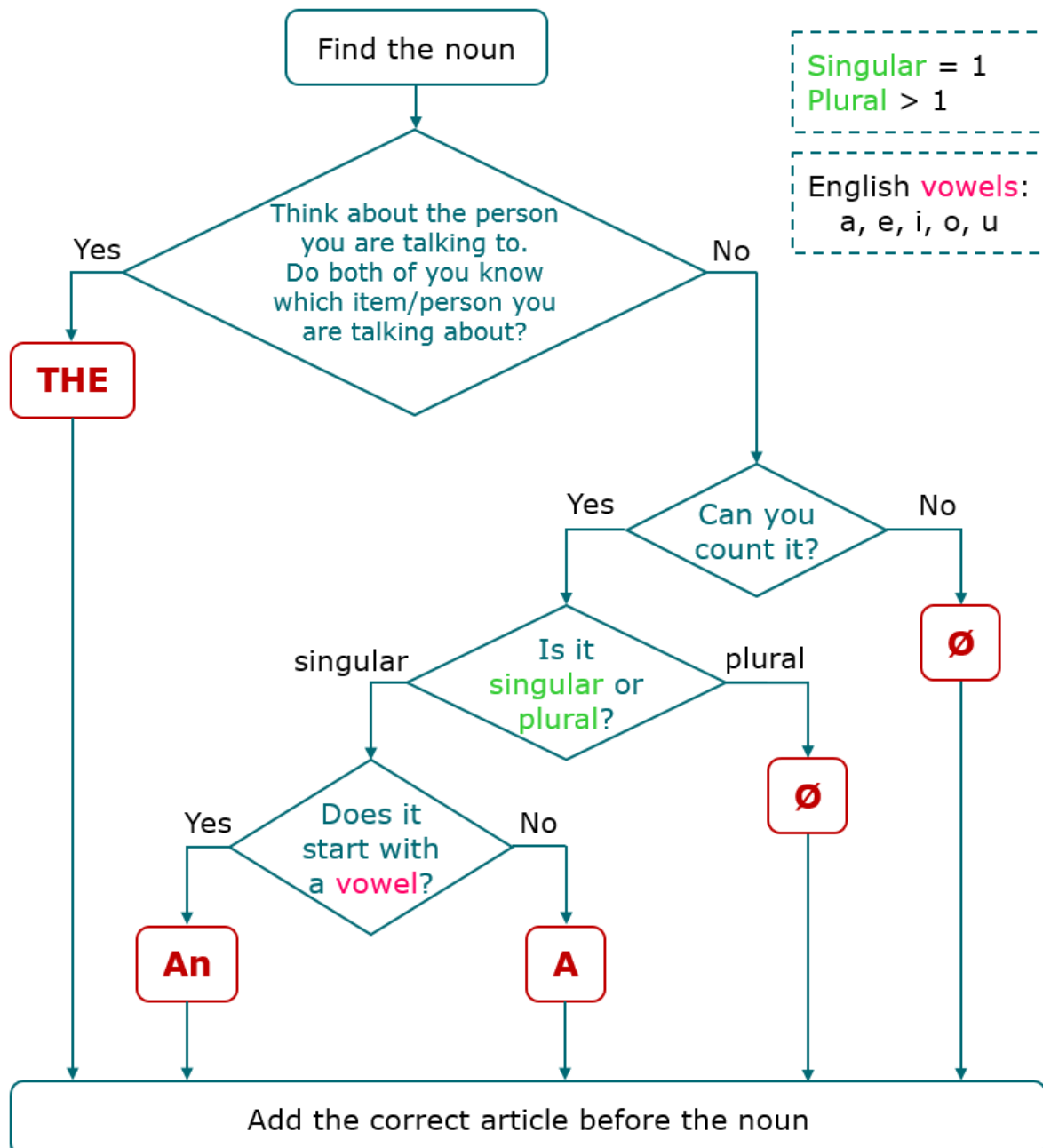
For binomial regression models, I used the BOBYQA algorithm as suggested by Bolker (2014). To avoid convergence issues, I increased the number of function evaluations from the default  $1 \times 10^4$  to  $1 \times 10^5$  and tightened the final radius of the trust region (which describes the scale of parameter uncertainty on convergence) from the default  $2 \times 10^{-7}$  to  $2 \times 10^{-9}$ .

For multinomial models, I increased the positive convergence tolerance from the default  $1 \times 10^{-8}$  up to  $3 \times 10^{-6}$  or  $3 \times 10^{-5}$  to help convergence (where necessary), always checking that this does not change the model estimates significantly.

To analyse random effects structures, I used Principal Components Analysis (PCA), which shows the number of parameters explaining 100% variance out of the total number of parameters. I used log-likelihood ratio tests to compare model fits. To assess models' predictive power, I used the C-statistic, or the area under the Receiver Operating Characteristic (ROC) curve, which ranges from 0.5 to 1 (Hosmer et al., 2013). To detect multicollinearity, I calculated VIFs, or variance inflation factors, where values below 5 indicate low collinearity, 5-10 moderate collinearity, above 10 strong collinearity, which should be avoided (James et al., 2013). I checked overdispersion using the `performance` package in R (Lüdtke et al., 2021), as suggested by Bolker (2021).

## Definite Article

**A/AN/Ø** or **THE** ?





## Appendix D: Structure of Pilot Experimental Group Activities

**Figure 44**

*Factorial Structure of Pilot Experimental Activities: Block 1*

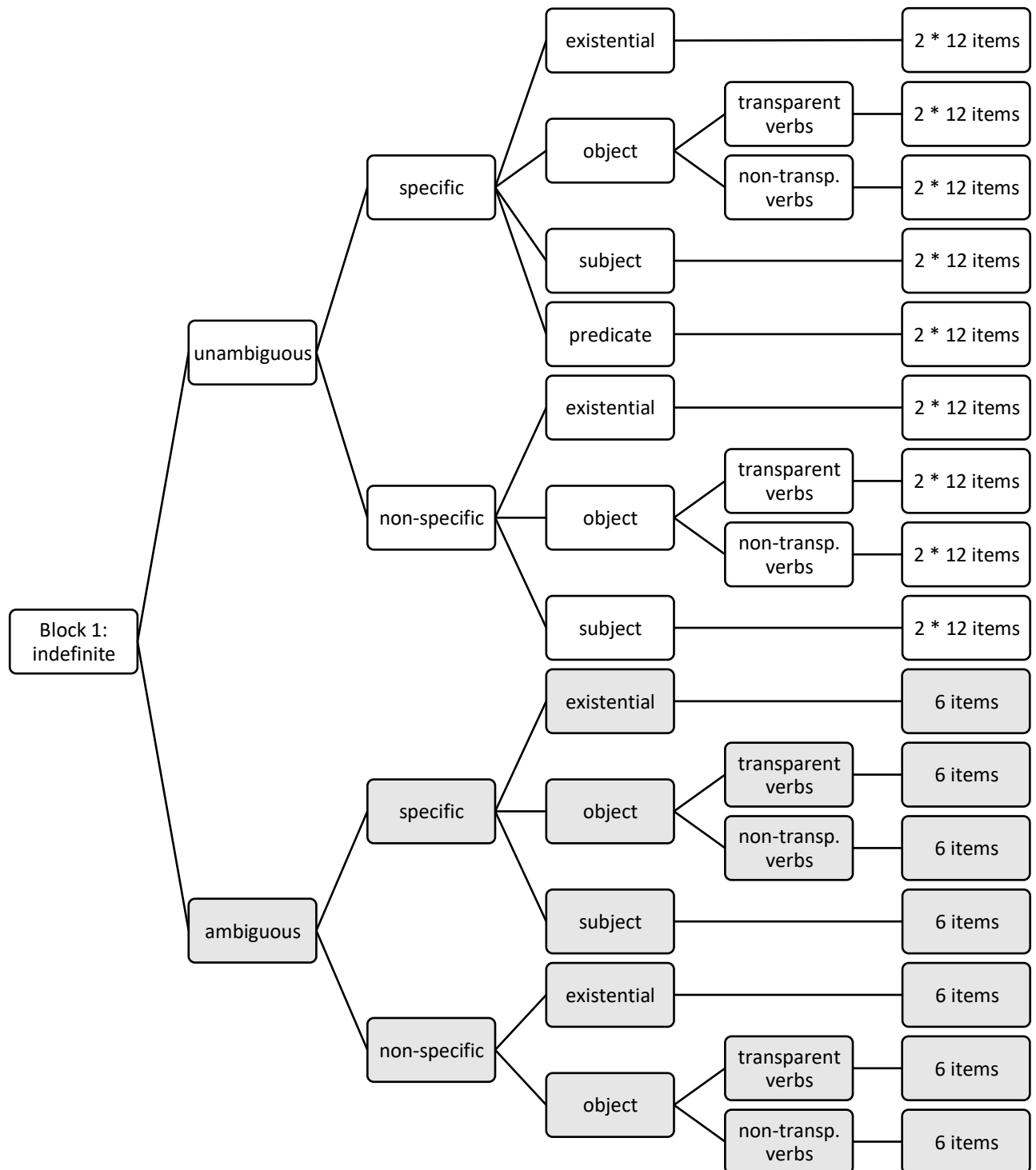


Figure 44 shows how the items were structured by factors in the indefinite block (Block 1).

Where the right-hand box states “2 \* 12 items”, these were structured into sets of 12 by noun type,

abstractness, and modifier presence, as shown in Figure 45. The number of items is also multiplied by 2 because each set was created with high-frequency and with low-frequency nouns. Boxes with “6 items” were structured by noun type and abstractness, while modifier presence was evenly distributed but not included as a factor.

**Figure 45**

*Item Structure Within 12-Item Sets*

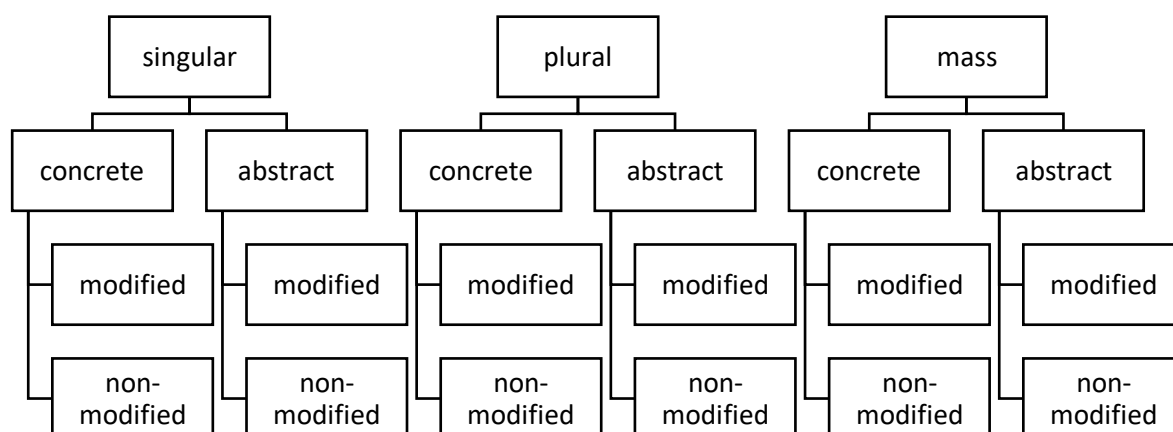
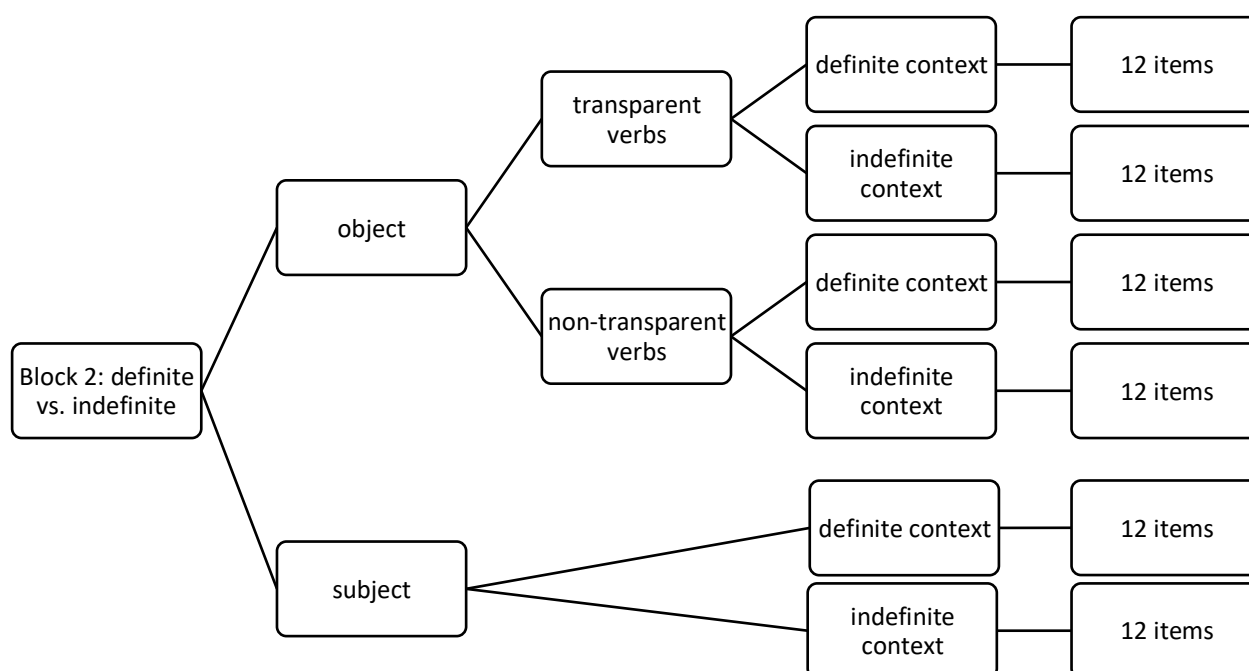


Figure 46 shows how the items were structured by factors in the definite vs. indefinite block (Block 2). Each category had 12 items structured as shown in Figure 45.

**Figure 46**

*Factorial Structure of Pilot Experimental Activities: Block 2*



**Appendix E: Example Control Group Activity**

This content has been redacted due to third-party copyright.

## Appendix F: Pilot Tests Structure and Example Items

Table 23 shows the factorial structure of test items on the GJT in the pilot study. The columns Def (definiteness), Spec (specificity), Abstr (abstractness), and Ntype (noun type) specify the category of item. The next three columns show the syntactic positions in which each item category appeared on the different versions of the test (A/B/C). Underlined items indicate that the noun was also prenominally modified. Syntactic position and prenominal modification were balanced across test versions; note that for different items different test versions contained more than one instance (e.g. object and subject positions for the first item category in Version B). In the Example column, the ungrammatical version of an item is given with an asterisk in parentheses.

**Table 23**

*Structure of Pilot GJT Test Items*

|                                  | Def      | Spec     | Abstr | Ntype | Version A        | Version B        | Version C        | Example   |
|----------------------------------|----------|----------|-------|-------|------------------|------------------|------------------|---|
| Definite vs. Indefinite contrast | definite | specific | sing  | obj   |                  | <u>obj</u> , sub | <u>sub</u>       | Where have you been? - We spent <i>the (*an) afternoon</i> in the library.  |
|                                  |          |          |       | mass  | <u>obj</u>       | sub              | obj, <u>sub</u>  | <i>The (*∅) literature</i> that you gave me was not very useful.  |
|                                  |          |          |       | pl    | obj, <u>sub</u>  | <u>obj</u>       | sub              | Can you find out <i>the (*∅) ages</i> of those children? - Yes, I can ask their parents.  |
|                                  |          |          | concr | sing  | <u>obj</u> , sub | <u>sub</u>       | obj              | Is my car ready? - No, our mechanic is still fixing <i>the (*a) car engine</i> .  |
|                                  |          |          |       | mass  | sub              | obj, <u>sub</u>  | <u>obj</u>       | Why don't we plant some flowers in our garden now? - <i>The (*∅) ground</i> will be frozen in one week. It's not a good time.   |
|                                  |          |          |       | pl    | <u>sub</u>       | obj              | <u>obj</u> , sub | I love going to the French Alps in summer. - Me too. <i>The (*∅) mountain lakes</i> that I visited last year were so beautiful. |
|                                  |          |          |       |       |                  |                  |                  |   |
|                                  |          |          |       |       |                  |                  |                  |   |

|                         | Def        | Spec         | Abstr | Ntype | Version A                   | Version B                   | Version C                    | Example  |
|-------------------------|------------|--------------|-------|-------|-----------------------------|-----------------------------|------------------------------|--|
| Indefinite a/∅ contrast | indefinite |              | abstr | sing  | obj                         | <u>obj</u> , sub            | <u>sub</u>                   | The company has created <i>a (*the) new department</i> . - Which one?  |
|                         |            |              |       | mass  | sub                         | obj, <u>sub</u>             | <u>obj</u>                   | <i>(*The) Agriculture</i> is the main source of income for this region. - I see. What do they grow here?           |
|                         |            |              |       | pl    | obj, <u>sub</u>             | <u>obj</u>                  | sub                          | So how was the talk? - Great. He explained <i>(*the) interesting concepts</i> . - Which ones? Tell me.             |
|                         |            |              | concr | sing  | <u>obj</u>                  | sub                         | obj, <u>sub</u>              | I want to sit down, but there aren't any free seats here. - Oh, look! I see <i>a (*the) free bench</i> over there. |
|                         |            |              |       | mass  | <u>obj</u> , sub            | <u>sub</u>                  | obj                          | <i>(*The) Coal</i> was the biggest source of energy in this region last year. - Oh, I didn't know that.            |
|                         |            |              |       | pl    | <u>sub</u>                  | obj                         | <u>obj</u> , sub             | What did you have for lunch? - I had some fish with <i>(*the) roast vegetables</i> .                               |
|                         |            | specific     | abstr | sing  | <u>pre</u> , ex, <u>sub</u> | <u>obj</u> , pre            | sub, obj, <u>ex</u>          | There was <i>a (*∅) terrible accident</i> .  |
|                         |            |              |       | mass  | ex, <u>obj</u>              | sub, <u>pre</u> , <u>ex</u> | pre, obj, <u>sub</u>         | This isn't <i>(*a) poetry</i> .  |
|                         |            |              |       | pl    | <u>ex</u> , obj, pre        | <u>sub</u> , ex             | sub, <u>pre</u> , <u>obj</u> | They wanted <i>(*a) different options</i> .  |
|                         |            |              | concr | sing  | sub, <u>pre</u>             | ex, obj, <u>sub</u>         | pre, <u>ex</u> , <u>obj</u>  | A <i>(*∅) gun</i> was found in his bag.  |
|                         |            |              |       | mass  | <u>ex</u> , obj, <u>sub</u> | <u>obj</u> , sub, pre       | ex, <u>pre</u>               | This is just <i>(*a) mud</i> , don't worry.  |
|                         |            |              |       | pl    | <u>obj</u> , sub, pre       | obj, <u>pre</u> , <u>ex</u> | ex, <u>sub</u>               | I saw <i>(*a) fishing boats</i> on the river.  |
|                         | indefinite | non-specific | abstr | sing  | <u>obj</u> , sub            | <u>ex</u> , obj             | <u>sub</u> , ex              | We make <i>a (*∅) copy</i> of each new document that we receive.   |
|                         |            |              |       | mass  | <u>sub</u> , ex             | <u>obj</u> , sub            | <u>ex</u> , obj              | There is always <i>(*a) magic</i> in the air in this place.  |
|                         |            |              |       | pl    | <u>ex</u> , obj             | <u>sub</u> , ex             | <u>obj</u> , sub             | If you want to build <i>(*a) relationships</i> , you need to listen to people.                                     |
|                         |            |              | concr | sing  | <u>obj</u> , sub            | <u>ex</u> , obj             | <u>sub</u> , ex              | He doesn't want to visit <i>a (*∅) farm</i> .  |
|                         |            |              |       | mass  | <u>ex</u> , obj             | <u>sub</u> , ex             | <u>obj</u> , sub             | <i>(*A) Steel</i> is often used in construction.   |
|                         |            |              |       | pl    | <u>sub</u> , ex             | <u>obj</u> , sub            | <u>ex</u> , obj              | You won't see <i>(*a) green hills</i> anywhere.  |

Table 24 provides the factorial structure of EI items, which was similar to the one for the GJT. Several example prompts are provided under the table with target nominals italicised and the number of the item which they exemplify given in parentheses.

**Table 24**

*Structure of Pilot EI Test Items*

| Def        | Spec         | Abstr | Ntype | Item No. | Version A                   | Version B                   | Version C                    |
|------------|--------------|-------|-------|----------|-----------------------------|-----------------------------|------------------------------|
| definite   | specific     |       | sing  | 1        | obj                         | <u>obj</u> , sub            | <u>sub</u>                   |
|            |              | abstr | mass  | 2        | <u>obj</u>                  | sub                         | obj, <u>sub</u>              |
|            |              |       | pl    | 3        | obj, <u>sub</u>             | <u>obj</u>                  | sub                          |
|            |              |       | sing  | 4        | <u>obj</u> , sub            | <u>sub</u>                  | obj                          |
|            |              | concr | mass  | 5        | sub                         | obj, <u>sub</u>             | <u>obj</u>                   |
|            |              |       | pl    | 6        | <u>sub</u>                  | obj                         | <u>obj</u> , sub             |
| indefinite | specific     |       | sing  | 7        | <u>pre</u> , ex, <u>sub</u> | <u>obj</u> , pre            | sub, obj, <u>ex</u>          |
|            |              | abstr | mass  | 8        | ex, <u>obj</u>              | sub, <u>pre</u> , <u>ex</u> | pre, obj, <u>sub</u>         |
|            |              |       | pl    | 9        | <u>ex</u> , obj, pre        | <u>sub</u> , ex             | sub, <u>pre</u> , <u>obj</u> |
|            |              |       | sing  | 10       | sub, <u>pre</u>             | ex, obj, <u>sub</u>         | pre, <u>ex</u> , <u>obj</u>  |
|            |              | concr | mass  | 11       | <u>ex</u> , obj, <u>sub</u> | <u>obj</u> , sub, pre       | ex, <u>pre</u>               |
|            |              |       | pl    | 12       | <u>obj</u> , sub, pre       | obj, <u>pre</u> , <u>ex</u> | ex, <u>sub</u>               |
|            | non-specific |       | sing  | 13       | <u>obj</u> , sub            | <u>ex</u> , obj             | <u>sub</u> , ex              |
|            |              | abstr | mass  | 14       | <u>sub</u> , ex             | <u>obj</u> , sub            | <u>ex</u> , obj              |
|            |              |       | pl    | 15       | <u>ex</u> , obj             | <u>sub</u> , ex             | <u>obj</u> , sub             |
|            |              |       | sing  | 16       | <u>obj</u> , sub            | <u>ex</u> , obj             | <u>sub</u> , ex              |
|            |              | concr | mass  | 17       | <u>ex</u> , obj             | <u>sub</u> , ex             | <u>obj</u> , sub             |
|            |              |       | pl    | 18       | <u>sub</u> , ex             | <u>obj</u> , sub            | <u>ex</u> , obj              |

Examples of EI prompts:

- Lucy is *a marketing manager* (10). She's making *new catalogues* (12) for our clients. *The catalogue* (4) for you is almost finished.
- There is *a concert* (7) tonight. We can get *an extra ticket* (16). Do you want to come? - Sorry, there are *important tasks* (9) I need to finish.
- *True wisdom* (14) cannot be learned. There may be *teachers* (18) who can give you *the great knowledge* (2) that you want. But they can't make you wise.

- There are *periods* (15) of crisis in every civilisation. But those which limit *human freedom* (14) always pay *a price* (13) and fail.

Table 25 shows the writing prompts given to participants in the different versions of the test.

**Table 25**

*Pilot Free Writing Prompts*

| Test version | Writing prompt   |
|--------------|--|
| A            | <p>Your friend from another country is going to visit your city/town. Give them some recommendations.</p> <p>You can write about:</p> <ul style="list-style-type: none"> <li>- places to visit</li> <li>- food &amp; drinks to try</li> <li>- weather</li> <li>- things they should or shouldn't take with them</li> </ul> <p>Write 80-100 words.</p>          |
| B            | <p>What are the positive and negative things about using smart phones?</p> <p>You can write about:</p> <ul style="list-style-type: none"> <li>- social media and real-life communication</li> <li>- technical features of smart phones</li> <li>- health problems</li> <li>- what you like or dislike about smart phones</li> </ul> <p>Write 80-100 words.</p> |
| C            | <p>Describe your school / university / work place.</p> <p>You can write about:</p> <ul style="list-style-type: none"> <li>- where it is located and what it looks like</li> <li>- people you study or work with</li> <li>- regular activities</li> <li>- what you like or dislike about it</li> </ul> <p>Write 80-100 words.</p>                               |

**Appendix G: Consent Form and Background Questionnaire for the Pilot**

# Consent and background

---

**Start of Block: Consent form**

Q1

Dear Participant,

Thank you for choosing to participate in our research project!

Before we start, please complete this form. If you have any questions, do not hesitate to contact me at [Email address redacted]

Kateryna Derkach

---

**Q2 CONSENT FORM**

**Project title:** Processing Instruction Learning Activities

**Research team:** Kateryna Derkach and Theodora Alexopoulou

If you have any questions, please contact Kateryna Derkach, [Email address redacted]

---



Q3 Please read and tick all the boxes below.

☐

I confirm that I have read and understand the information sheet dated 20/05/2019 for the above mentioned study and have had the opportunity to ask questions.

☐

I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, and without my rights being affected.

☐

I understand that any data that are collected will be used and stored anonymously, in accordance with the Data Protection Act. Results are normally presented in terms of groups of individuals. If any individual data were presented, the data would be completely anonymous, without any means of identifying the individuals involved.

☐

I understand that these data may be used in analyses, publications, and conference presentations by researchers at the University of Cambridge and their collaborators at other research institutions. I give permission for these individuals to have access to these data.

☐

I understand that personal information (such as language background, age, gender, education level, length of stay in an English-speaking country) will be collected as part of this research. Full data will only be accessible to the research team. However, anonymised data may be used in analyses, publications and conference presentations. For full details on how researchers will use your personal information, see <https://www.information-compliance.admin.cam.ac.uk/data-protection/research-participant-data>

☐

I have been provided a copy of this form to keep. (This will be sent to your e-mail when you complete this form)

☐

I agree to participate in the above mentioned study run by Kateryna Derkach, a PhD student at the Faculty of MML at the University of Cambridge.

---

Q4 Please enter your name.

---

Q5 Today's date (dd/mm/yyyy)

---

Q6 Your signature

Q7 **Name of researcher:** Kateryna Derkach

**Date:** 10/07/2019

**Signature of researcher:**

Q8 *The project has received ethical approval from the Faculty's Research Ethics Committee. If you have any questions or complaints about the ethical aspects of this study, please contact [Email address redacted]*

End of Block: Consent form

Start of Block: Personal Information

Q9 In this section we will ask you to give us some personal information and some information about your language background. Everything you share will be kept confidential.

Q55 Your contact e-mail (we will use this to send you links to practice activities)

☐ E-mail 

---

Q11 Your age (in years)

---

---

Q12 Your gender

- ☐ female
- ☐ male
- ☐ other
- ☐ prefer not to say
- 

Q14 Your country of origin (In which country were you born?)

▼ Afghanistan ... Other (please specify in the box below)

---

Q15 Please specify your country of origin.

\_\_\_\_\_

---

Q18 Do you normally live in your country of origin?

- ☐ Yes
- ☐ No
-

Q17 What is your country of residence? (Which country do you normally live in?)

▼ Afghanistan ... Other (please specify in the box below)

-----

Q19 Please specify your country of residence.

\_\_\_\_\_

-----

Q20 During which ages did you live in your country of origin? (For example: from 0 to 10 years old)

\_\_\_\_\_

-----

Q21 During which ages have you lived in your country of residence? (For example: from 11 to 17 years old)

\_\_\_\_\_

-----

Q22 What is the highest level of education that you have completed?

☐ High school

☐ Bachelor's degree or equivalent

☐ Master's degree or equivalent

☐ PhD or equivalent

☐ Other (please specify) \_\_\_\_\_

-----

Q23 Are you a school or college student at the moment?

☐ Yes

☐ No

---

Q24 Please specify at which level you are studying at the moment.

☐ High school (indicate year) \_\_\_\_\_

☐ Bachelor's degree or equivalent (indicate year)  
\_\_\_\_\_

☐ Master's degree or equivalent

☐ PhD or equivalent

☐ Other (please specify) \_\_\_\_\_

End of Block: Personal Information

---

Start of Block: Language Background

Q25 What is your native language(s)?

\_\_\_\_\_

---

Q26 At what age did you first hear your native language(s)?

\_\_\_\_\_

---

Q27 At what age did you start learning English?

\_\_\_\_\_

---

Q28 How have you learned English up to now? Tick all the options that apply and indicate the details in the boxes provided.

☐

English classes at school: a) for how many years? b) how many hours/week?

---

☐

English classes outside of school, e.g. in a language school or with a private tutor: a) for how many years? b) how many hours/week?

---

☐

Books, films, songs, other media: how many hours/week?

---

☐

Communication with English-speaking friends: how many hours/week?

---

☐

Other (please specify details)

---

---

Q29 What is your English level at your language school?

☐ Elementary (A1)

☐ Pre-intermediate (A2)

☐ Intermediate (B1)

☐ Upper-Intermediate (B2)

☐ Advanced (C1)

☐ Other (please specify) \_\_\_\_\_

---

Q30 Have you got any English language test certificates? Please tick all that apply and indicate your results (you can give approximate results if you do not remember exactly).

- ☐ IELTS \_\_\_\_\_
- ☐ TOEFL \_\_\_\_\_
- ☐ PET \_\_\_\_\_
- ☐ FCE \_\_\_\_\_
- ☐ CAE \_\_\_\_\_
- ☐ PTE \_\_\_\_\_
- ☐ Other (please specify test and result)  
\_\_\_\_\_
- 

Q32 What language(s) did the teachers use in your classes at these education levels?

- ☐ Primary school \_\_\_\_\_
- ☐ Secondary school \_\_\_\_\_
- ☐ High school \_\_\_\_\_
- ☐ College/university \_\_\_\_\_
- 

Q31 Have you learned or used any other languages, except your native language(s) and English?

- ☐ Yes
- ☐ No
-

## Q36 Language 1

- ☐ Language \_\_\_\_\_
- ☐ Level \_\_\_\_\_
- ☐ How often do/did you use it? \_\_\_\_\_
- ☐ Where and at what ages did you use it? (For example: when I lived in Spain for 6 months when I was 12) \_\_\_\_\_
- 

## Q37 Language 2

- ☐ Language \_\_\_\_\_
- ☐ Level \_\_\_\_\_
- ☐ How often do/did you use it? \_\_\_\_\_
- ☐ Where and at what ages did you use it? (For example: when I lived in Spain for 6 months when I was 12) \_\_\_\_\_
- 

## Q38 Language 3

- ☐ Language \_\_\_\_\_
- ☐ Level \_\_\_\_\_
- ☐ How often do/did you use it? \_\_\_\_\_
- ☐ Where and at what ages did you use it? (For example: when I lived in Spain for 6 months when I was 12) \_\_\_\_\_
-



Q39 If you speak any more languages, please give details in the box below.

---

---

---

---

---

Q33 Have you lived in another country for more than 3 months?

☐ Yes

☐ No

Q34 Country 1

☐ Country \_\_\_\_\_

☐ At what ages did you live there? (For example: from 15 to 17 years old)

---

☐ Which language(s) did you use there?

---

Q40 Country 2

☐ Country \_\_\_\_\_

☐ At what ages did you live there? (For example: from 15 to 17 years old)

\_\_\_\_\_

☐ Which language(s) did you use there?

\_\_\_\_\_

-----

Q41 Country 3

☐ Country \_\_\_\_\_

☐ At what ages did you live there? (For example: from 15 to 17 years old)

\_\_\_\_\_

☐ Which language(s) did you use there?

\_\_\_\_\_

-----

Q42 If you have lived in more countries, please give details in the box below.

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

-----

Q35 Is there anything else that you feel is interesting or important about your language background that you'd like us to know?

\_\_\_\_\_

\_\_\_\_\_

---

---

---

End of Block: Language Background

---

**Appendix H: Links to Example Video Lessons for Main Intervention Study**

Lesson 1.1. - <https://youtu.be/ETI1fBX846U>

Lesson 1.2 - <https://youtu.be/ay-A-JnK4CM>

Lesson 2 - <https://youtu.be/TkvCQOAQTpU>

Lesson 5.1 - <https://youtu.be/yYufa6nQxT8>

Lesson 5.2 - <https://youtu.be/kEJIPc-hhvl>

Lesson 6 - <https://youtu.be/tSIFxpD3CDc>

Lesson 10.1 - <https://youtu.be/06bnZqso6D8>

Lesson 10.2 - [https://youtu.be/MRGAZX\\_rz70](https://youtu.be/MRGAZX_rz70)

To see how the videos were integrated into the activity, see activity 5 via this link:

[https://cambridge.eu.qualtrics.com/jfe/form/SV\\_eECzmvRiqxJHSCy](https://cambridge.eu.qualtrics.com/jfe/form/SV_eECzmvRiqxJHSCy)

**Appendix I: Example Activities from the Main Intervention**

This Appendix contains example items of different categories from activities 11 and 14. The order of the presentation of items and options was randomised for participants.

**There is a big kestrel in the sky.**

You could use this sentence to talk about...

*(Choose all possible answers)*☐

...one kestrel

☐

...more than one kestrel

☐

...part of one kestrel

☐

...more than one kettle

**There are big kestrels in the sky.**

You could use this sentence to talk about...

*(Choose all possible answers)*☐

...one kestrel

☐

...more than one kestrel

☐

...part of one kestrel

☐

...one kettle

**A goad made them work harder.**

You could use this sentence to talk about...

*(Choose all possible answers)*☐

...one goad

☐

...more than one goad

☐

...part of one goad

☐

...one piece of gold

**Goads made them work harder.**

You could use this sentence to talk about...

*(Choose all possible answers)*☐

...one goad

☐

...more than one goad

☐

...part of one goad

☐

...more than one goat

**They made a serious botch.**

You could use this sentence to talk about...

*(Choose all possible answers)*☐

...one botch

☐

...more than one botch

☐

...part of one botch

☐

...one bottle

**They made serious botches.**

You could use this sentence to talk about...

*(Choose all possible answers)*☐

...one botch

☐

...more than one botch

☐

...one instance of one botch

☐

...part of one watch

**She is a dowd.**

You could use this sentence to talk about...  
(Choose all possible answers)

- ☐ ...one dowd
- ☐ ...more than one dowd
- ☐ ...part of one dowd
- ☐ ...part of one crowd

**They are dowds.**

You could use this sentence to talk about...  
(Choose all possible answers)

- ☐ ...one dowd
- ☐ ...more than one dowd
- ☐ ...part of one dowd
- ☐ ...more than one crowd

**There is a chasm in every society.**

You could use this sentence to talk about...  
(Choose all possible answers)

- ☐ ...one chasm in every society
- ☐ ...more than one chasm in every society
- ☐ ...part of one chasm in every society
- ☐ ...part of charm in every society

**There are chasms in every society.**

You could use this sentence to talk about...  
(Choose all possible answers)

- ☐ ...one chasm in every society
- ☐ ...more than one chasm in every society
- ☐ ...part of one chasm in every society
- ☐ ...more than one instance of charm

**There is always a cheap gaud in her bag.**

You could use this sentence to talk about...  
(Choose all possible answers)

- ☐ ...one gaud
- ☐ ...more than one gaud
- ☐ ...part of one gaud
- ☐ ...more than one god

**There are always cheap gauds in her bag.**

You could use this sentence to talk about...  
(Choose all possible answers)

- ☐ ...one gaud
- ☐ ...more than one gaud
- ☐ ...part of one gaud
- ☐ ...part of one god

**She works for a pittance.**

You could use this sentence to talk about...  
(Choose all possible answers)

- ☐ ...one pittance
- ☐ ...more than one pittance
- ☐ ...part of one pittance
- ☐ ...part of pity

**If you go on a rampage, you may get in trouble.**

You could use this sentence to talk about...  
(Choose all possible answers)

- ☐ ...one rampage
- ☐ ...more than one rampage
- ☐ ...part of one rampage
- ☐ ...one campaign

**The school has punished truancy before.**

You could use this sentence to talk about...  
(Choose all possible answers)

- ☐ ...one case of truancy
- ☐ ...more than one case of truancy
- ☐ ...a lot of truancy
- ☐ ...one instance of cruelty

**They work for pittances.**

You could use this sentence to talk about...  
(Choose all possible answers)

- ☐ ...one pittance
- ☐ ...more than one pittance
- ☐ ...part of one pittance
- ☐ ...a little bit of pity

**Is thrush a common problem in babies?**

You could use this sentence to talk about...  
(Choose all possible answers)

- ☐ ...one instance of thrush
- ☐ ...more than one instance of thrush
- ☐ ...thrush in many babies
- ☐ ...part of trash

**They study sheep scab at this lab.**

You could use this sentence to talk about...  
(Choose all possible answers)

- ☐ ...one piece of scab
- ☐ ...more than one piece of scab
- ☐ ...part of scab
- ☐ ...more than one crab



**A small gripe cannot stop him from doing what he wants.**

You could use this sentence to talk about...  
(Choose all possible answers)

- ☐ ...one gripe
- ☐ ...more than one gripe
- ☐ ...part of one gripe
- ☐ ...one bunch of grapes

**Small gripes cannot stop him from doing what he wants.**

You could use this sentence to talk about...  
(Choose all possible answers)

- ☐ ...one gripe
- ☐ ...more than one gripe
- ☐ ...part of one gripe
- ☐ ...one small grape

*In which dialogue does the purple phrase fit best?*

- ☐ - You will make the presentation tomorrow. - I remember, I've prepared everything.
- ☐ - You will make the presentation tomorrow. - Oh, thanks for letting me know. What should I present?
- ☐ - You will make the presentation tomorrow. - I'll take three kilos, please.

*In which dialogue does the purple phrase fit best?*

- ☐ - You will make a presentation tomorrow. - I remember, I've prepared everything.
- ☐ - You will make a presentation tomorrow. - Oh, thanks for letting me know. What should I present?
- ☐ - You will make a presentation tomorrow. - I'll take three kilos, please.

*In which context does the purple phrase fit best?*

- ☐ I've looked at the different types of luggage at this shop. The big suitcases are very expensive, so I won't buy any of those.
- ☐ Usually the big suitcases are very expensive, so I'm thinking of buying something second-hand.
- ☐ I don't read newspapers because the big suitcases are very expensive.

*In which context does the purple phrase fit best?*

- ☐ I've looked at the different types of luggage at this shop. Big suitcases are very expensive, so I won't buy any of those.
- ☐ Usually big suitcases are very expensive, so I'm thinking of buying something second-hand.
- ☐ I don't read newspapers because big suitcases are very expensive.

## Appendix J: Main Intervention Tests – Structure and Example Items

Table 26 shows the factorial structure of test items on the GJT in the main intervention study. The columns Def (definiteness), Spec (specificity), Abstr (abstractness), and Ntype (noun type) specify the category of item, and each category is assigned an item number. The next three columns show the syntactic positions in which each item category appeared on the different versions of the test (A/B/C). Underlined items indicate that the noun was also prenominally modified. Syntactic position and prenominal modification were balanced across test versions; note that for different items different test versions contained more than one instance (e.g. object and subject positions for the first item category in Version B).

Tables 27, 28, 29 show the full list of test items (excluding distractor items) for test versions A, B, C, respectively. The first column in these tables shows the item category number from Table 26. The Test part column indicates in which part of the test the item appeared, as tests were split up into two parts. The target nominals are in bold.

Tables 30, 31, 32 show the factorial structure and prompts for the EI task in the main intervention study for test versions A, B, C, respectively.

Table 33 shows the writing prompts used for each test version in parts 1 and 2 of the test.

Table 26

Structure of Main Intervention GJT Items

|                                  | Def        | Spec         | Abstr | Ntype | Item No. | Version A          | Version B          | Version C          |
|----------------------------------|------------|--------------|-------|-------|----------|--------------------|--------------------|--------------------|
| Definite vs. Indefinite contrast | definite   | specific     | abstr | sing  | 1        | obj                | obj, sub           | sub                |
|                                  |            |              |       | mass  | 2        | obj                | sub                | obj, sub           |
|                                  |            |              |       | pl    | 3        | obj, sub           | obj                | sub                |
|                                  |            |              | concr | sing  | 4        | obj, sub           | sub                | obj                |
|                                  |            |              |       | mass  | 5        | sub                | obj, sub           | obj                |
|                                  |            |              |       | pl    | 6        | sub                | obj                | obj, sub           |
|                                  | indefinite |              | abstr | sing  | 7        | obj                | obj, sub           | sub                |
|                                  |            |              |       | mass  | 8        | sub                | obj, sub           | obj                |
|                                  |            |              |       | pl    | 9        | obj, sub           | obj                | sub                |
|                                  |            |              | concr | sing  | 10       | obj                | sub                | obj, sub           |
|                                  |            |              |       | mass  | 11       | obj, sub           | sub                | obj                |
|                                  |            |              |       | pl    | 12       | sub                | obj                | obj, sub           |
| Indefinite a/∅ contrast          | indefinite | specific     | abstr | sing  | 13       | ex, obj, sub, pred | ex, obj, sub, pred | ex, obj, sub, pred |
|                                  |            |              |       | mass  | 14       | ex, obj, sub, pred | ex, obj, sub, pred | ex, obj, sub, pred |
|                                  |            |              |       | pl    | 15       | ex, obj, sub, pred | ex, obj, sub, pred | ex, obj, sub, pred |
|                                  |            |              | concr | sing  | 16       | ex, obj, sub, pred | ex, obj, sub, pred | ex, obj, sub, pred |
|                                  |            |              |       | mass  | 17       | ex, obj, sub, pred | ex, obj, sub, pred | ex, obj, sub, pred |
|                                  |            |              |       | pl    | 18       | ex, obj, sub, pred | ex, obj, sub, pred | ex, obj, sub, pred |
|                                  |            | non-specific | abstr | sing  | 19       | ex, obj, sub       | ex, obj, sub       | ex, obj, sub       |
|                                  |            |              |       | mass  | 20       | ex, obj, sub       | ex, obj, sub       | ex, obj, sub       |
|                                  |            |              |       | pl    | 21       | ex, obj, sub       | ex, obj, sub       | ex, obj, sub       |
|                                  |            |              | concr | sing  | 22       | ex, obj, sub       | ex, obj, sub       | ex, obj, sub       |
|                                  |            |              |       | mass  | 23       | ex, obj, sub       | ex, obj, sub       | ex, obj, sub       |
|                                  |            |              |       | pl    | 24       | ex, obj, sub       | ex, obj, sub       | ex, obj, sub       |

Table 27

Main Intervention GJT Test A (Target Items Only)

| Item No. | ID  | Test part | Synt & mod | Grammatical  | Ungrammatical  |
|----------|-----|-----------|------------|--|--|
| 1        | A01 | 1         | obj        | What did you do yesterday? - I spent <b>the morning</b> exercising in the park and then went to work.            | What did you do yesterday? - I spent <b>a morning</b> exercising in the park and then went to work.          |
| 2        | A02 | 2         | obj        | Have you talked to <b>the new management</b> of the company?   | Have you talked to <b>new management</b> of the company?   |
| 3        | A03 | 1         | obj        | Can you find out <b>the ages</b> of those children?  | Can you find out <b>ages</b> of those children?  |
|          | A04 | 2         | sub        | So, are Linda and Bob going to act in this film? - No, they said <b>the roles</b> were not interesting for them. | So, are Linda and Bob going to act in this film? - No, they said <b>roles</b> were not interesting for them. |

| Item No. | ID  | Test part | Synt & mod | Grammatical  | Ungrammatical  |
|----------|-----|-----------|------------|--|--|
| 4        | A05 | 1         | <u>obj</u> | Is my car ready? - No, we are still fixing <b>the car engine</b> .   | Is my car ready? - No, we are still fixing <b>a car engine</b> .   |
|          | A06 | 2         | sub        | <b>The hall</b> that you're looking for is on the left of the lift.  | <b>A hall</b> that you're looking for is on the left of the lift.  |
| 5        | A07 | 2         | sub        | <b>The ground</b> in our garden is great for growing roses.  | <b>Ground</b> in our garden is great for growing roses.  |
| 6        | A08 | 1         | <u>sub</u> | I have visited a lot of mountain lakes in my life, but <b>the mountain lakes</b> that we saw in the Alps last summer were the most beautiful of all.                           | I have visited a lot of mountain lakes in my life, but <b>mountain lakes</b> that we saw in the Alps last summer were the most beautiful of all.                                   |
| 7        | A09 | 2         | obj        | So, what did Mary do with her prize money? - Oh, she gave it to <b>a fund</b> . - What fund? - I have no idea.   | So, what did Mary do with her prize money? - Oh, she gave it to <b>the fund</b> . - What fund? - I have no idea.   |
| 8        | A10 | 1         | sub        | So, do most local people work in tourism? - Well, a lot of people do, but <b>farming</b> is also very popular.   | So, do most local people work in tourism? - Well, a lot of people do, but <b>the farming</b> is also very popular.   |
| 9        | A11 | 2         | obj        | They spent such a long time solving this problem. It took <b>decades</b> !   | They spent such a long time solving this problem. It took <b>the decades</b> !   |
|          | A12 | 1         | <u>sub</u> | Normally we never cut our prices, but <b>small discounts</b> on some items have been made during the crisis.   | Normally we never cut our prices, but <b>the small discounts</b> on some items have been made during the crisis.   |
| 10       | A13 | 2         | <u>obj</u> | Let's sit down. I see <b>a free bench</b> over there.  | Let's sit down. I see <b>the free bench</b> over there.  |
| 11       | A14 | 1         | <u>obj</u> | What type of fuel do you use? - Mostly gas. - How about something new? We offer <b>high-quality eco-friendly fuel</b> , and our consultants can help you choose the right one. | What type of fuel do you use? - Mostly gas. - How about something new? We offer <b>the high-quality eco-friendly fuel</b> , and our consultants can help you choose the right one. |
|          | A15 | 2         | sub        | <b>Coal</b> was the biggest source of energy in this region last year.   | <b>The coal</b> was the biggest source of energy in this region last year.   |
| 12       | A16 | 1         | <u>sub</u> | I had a strange dream. <b>Random people</b> were asking me questions in different languages. I couldn't understand anything!   | I had a strange dream. <b>The random people</b> were asking me questions in different languages. I couldn't understand anything!   |
| 13       | A17 | 1         | ex         | There is <b>an opinion</b> that this plan is not realistic.  | There is <b>opinion</b> that this plan is not realistic.   |
|          | A18 | 1         | obj        | He is writing <b>a response</b> to a client's e-mail.  | He is writing <b>response</b> to a client's e-mail.  |
|          | A19 | 2         | <u>sub</u> | <b>A quiet holiday</b> helped him to feel better.  | <b>Quiet holiday</b> helped him to feel better.  |
|          | A20 | 1         | <u>pre</u> | This is <b>a beautiful dream</b> .   | This is <b>beautiful dream</b> .   |
| 14       | A21 | 2         | <u>ex</u>  | There is <b>honesty</b> in this book.  | There is <b>an honesty</b> in this book.   |
|          | A22 | 2         | obj        | Why did she feel <b>anger</b> ?  | Why did she feel <b>an anger</b> ?   |
|          | A23 | 1         | <u>sub</u> | <b>Strong evidence</b> was found that he had connections with local gangsters.   | <b>A strong evidence</b> was found that he had connections with local gangsters.   |

| Item No. | ID  | Test part | Synt & mod | Grammatical   | Ungrammatical  |
|----------|-----|-----------|------------|---|--|
|          | A24 | 1         | <u>pre</u> | This is <b>extreme pessimism</b> .                              | This is <b>an extreme pessimism</b> .                              |
| 15       | A25 | 1         | <u>ex</u>  | There were <b>interesting conversations</b> after the meeting.  | There were <b>an interesting conversations</b> after the meeting.  |
|          | A26 | 2         | obj        | I'm writing down <b>addresses</b> in my notebook.               | I'm writing down <b>an addresses</b> in my notebook.               |
|          | A27 | 1         | sub        | <b>Accounts</b> were blocked after three failed login attempts. | <b>An accounts</b> were blocked after three failed login attempts. |
|          | A28 | 2         | pre        | These are <b>words</b> from his letter.                         | These are <b>a words</b> from his letter.                          |
| 16       | A29 | 2         | <u>ex</u>  | There is <b>an open pack</b> on the table.                      | There is <b>open pack</b> on the table.                            |
|          | A30 | 1         | <u>obj</u> | She used <b>a white board</b> .                                 | She used <b>white board</b> .                                      |
|          | A31 | 2         | sub        | They said <b>a gun</b> was found in his bag.                    | They said <b>gun</b> was found in his bag.                         |
|          | A32 | 2         | <u>pre</u> | He is <b>a good journalist</b> .                                | He is <b>good journalist</b> .                                     |
| 17       | A33 | 1         | ex         | There is <b>sand</b> in my shoe.                                | There is <b>a sand</b> in my shoe.                                 |
|          | A34 | 2         | <u>obj</u> | They used <b>cotton wool</b> to make clouds.                    | They used <b>a cotton wool</b> to make clouds.                     |
|          | A35 | 2         | sub        | <b>Smoke</b> made it difficult to breathe.                      | <b>A smoke</b> made it difficult to breathe.                       |
|          | A36 | 2         | pre        | This is just <b>mud</b> , don't worry.                          | This is just <b>a mud</b> , don't worry.                           |
| 18       | A37 | 2         | ex         | There are <b>chairs</b> in the dining room.                     | There are <b>a chairs</b> in the dining room.                      |
|          | A38 | 1         | <u>obj</u> | I saw <b>fishing boats</b> on the river.                        | I saw <b>a fishing boats</b> on the river.                         |
|          | A39 | 1         | <u>sub</u> | <b>New drugs</b> were tested in the laboratory.                 | <b>A new drugs</b> were tested in the laboratory.                  |
|          | A40 | 1         | pre        | They are not <b>drivers</b> .                                   | They are not <b>a drivers</b> .                                    |
| 19       | A41 | 2         | ex         | There is always <b>a limit</b> on people's spending.            | There is always <b>limit</b> on people's spending.                 |
|          | A42 | 1         | <u>obj</u> | Students take <b>a difficult test</b> each month.               | Students take <b>difficult test</b> each month.                    |
|          | A43 | 1         | <u>sub</u> | <b>A new project</b> is started every week.                     | <b>New project</b> is started every week.                          |
| 20       | A44 | 1         | <u>ex</u>  | There is always <b>new software</b> out there.                  | There is always <b>a new software</b> out there.                   |
|          | A45 | 1         | <u>obj</u> | <b>Stress</b> increases blood pressure.                         | <b>A stress</b> increases a blood pressure.                        |
|          | A46 | 2         | sub        | <b>Chess</b> may be the most difficult game.                    | <b>A chess</b> may be the most difficult game.                     |
| 21       | A47 | 1         | <u>ex</u>  | There are <b>interesting events</b> every week.                 | There are <b>an interesting events</b> every week.                 |
|          | A48 | 2         | obj        | Is it difficult to build <b>relationships</b> ?                 | Is it difficult to build <b>a relationships</b> ?                  |
|          | A49 | 1         | sub        | <b>Contracts</b> are signed every day.                          | <b>A contracts</b> are signed every day.                           |
| 22       | A50 | 2         | <u>ex</u>  | There is always <b>a friendly face</b> here.                    | There is always <b>friendly face</b> here.                         |
|          | A51 | 1         | obj        | He wants to visit <b>a farm</b> .                               | He wants to visit <b>farm</b> .                                    |
|          | A52 | 2         | <u>sub</u> | <b>A new bag</b> is always expensive.                           | <b>New bag</b> is always expensive.                                |
| 23       | A53 | 1         | ex         | There is always <b>steam</b> in the shower.                     | There is always <b>a steam</b> in the shower.                      |

| Item No.  | ID  | Test part | Synt & mod | Grammatical  | Ungrammatical  |
|-----------|-----|-----------|------------|--|--|
|           | A54 | 2         | <u>obj</u> | We don't eat <b>white bread</b> .                      | We don't eat <b>a white bread</b> .                      |
|           | A55 | 2         | <u>sub</u> | <b>Steel</b> is often used in construction.            | <b>A steel</b> is often used in construction.            |
| <b>24</b> | A56 | 2         | <u>ex</u>  | There may be <b>coats</b> on sale.                     | There may be <b>a coats</b> on sale.                     |
|           | A57 | 2         | <u>obj</u> | It's not nice to point <b>fingers</b> at other people. | It's not nice to point <b>a fingers</b> at other people. |
|           | A58 | 1         | <u>sub</u> | Are <b>potato chips</b> good for you?                  | Are <b>a potato chips</b> good for you?                  |

Table 28

Main Intervention GJT Test B (Target Items Only)

| Item No. | ID  | Test part | Synt & mod | Grammatical   | Ungrammatical   |
|----------|-----|-----------|------------|---|---|
| <b>1</b> | B01 | 1         | <u>obj</u> | John is in hospital. - I know, I saw <b>the car crash</b> .   | John is in hospital. - I know, I saw <b>a car crash</b> .   |
|          | B02 | 2         | <u>sub</u> | <b>The title</b> of this book is very interesting.  | <b>A title</b> of this book is very interesting.  |
| <b>2</b> | B03 | 2         | <u>sub</u> | <b>The literature</b> that you gave me was useful.  | <b>Literature</b> that you gave me was useful.  |
| <b>3</b> | B04 | 1         | <u>obj</u> | I need to know <b>the exact ages</b> of those students.   | I need to know <b>exact ages</b> of those students.   |
| <b>4</b> | B05 | 1         | <u>sub</u> | So, how much should we pay for this lunch? - I don't know. <b>The bill</b> is gone. I just put it here a minute ago, and now I can't find it!                     | So, how much should we pay for this lunch? - I don't know. <b>A bill</b> is gone. I just put it here a minute ago, and now I can't find it!                           |
| <b>5</b> | B06 | 2         | <u>obj</u> | They had a picnic on <b>the grass</b> in their garden.  | They had a picnic on <b>grass</b> in their garden.  |
|          | B07 | 1         | <u>sub</u> | <b>The white gold</b> that they used to make this ring was produced in Germany.   | <b>White gold</b> that they used to make this ring was produced in Germany.   |
| <b>6</b> | B08 | 2         | <u>obj</u> | This is such a funny house. Just look at <b>the windows</b> !   | This is such a funny house. Just look at <b>windows</b> !   |
| <b>7</b> | B09 | 2         | <u>obj</u> | What's new at work? - Well, our company is creating <b>a new department</b> , and they are asking me to manage it. - Oh, congratulations!                         | What's new at work? - Well, our company is creating <b>the new department</b> , and they are asking me to manage it. - Oh, congratulations!                           |
|          | B10 | 1         | <u>sub</u> | They have lots of awards. For example, <b>an award</b> was given to the most creative participant.  | They have lots of awards. For example, <b>the award</b> was given to the most creative participant.   |
| <b>8</b> | B11 | 1         | <u>obj</u> | I didn't think I would become emotional about his win, but in fact I felt <b>pride</b> .  | I didn't think I would become emotional about his win, but in fact I felt <b>the pride</b> .  |
|          | B12 | 2         | <u>sub</u> | So, what advice can you give to our listeners? - Well, I think most people do not learn from advice. <b>Personal experience</b> has been the best teacher for me. | So, what advice can you give to our listeners? - Well, I think most people do not learn from advice. <b>The personal experience</b> has been the best teacher for me. |

| Item No. | ID  | Test part | Synt & mod | Grammatical  | Ungrammatical  |
|----------|-----|-----------|------------|--|--|
| 9        | B13 | 2         | <u>obj</u> | I thought this seminar would be boring, but we discussed <b>very interesting concepts</b> actually.  | I thought this seminar would be boring, but we discussed <b>the very interesting concepts</b> actually.  |
| 10       | B14 | 1         | sub        | How was your day? - Something funny happened actually. <b>A man</b> came into my office looking for his goldfish. I have no idea who he was!             | How was your day? - Something funny happened actually. <b>The man</b> came into my office looking for his goldfish. I have no idea who he was!               |
| 11       | B15 | 2         | <u>sub</u> | Can I help you? - I'm looking for some cheap polyester. - Sorry, but <b>natural fabric</b> is the only thing you'll find here.                           | Can I help you? - I'm looking for some cheap polyester. - Sorry, but <b>the natural fabric</b> is the only thing you'll find here.                           |
| 12       | B16 | 1         | obj        | Excuse me! Could you help me, please? This parking meter only takes change, and I haven't got any. - Let me see. I'm sure I have <b>coins</b> somewhere. | Excuse me! Could you help me, please? This parking meter only takes change, and I haven't got any. - Let me see. I'm sure I have <b>the coins</b> somewhere. |
| 13       | B17 | 1         | ex         | There is <b>a question</b> at the end.   | There is <b>question</b> at the end.   |
|          | B18 | 2         | <u>obj</u> | I can see <b>a big advantage</b> here.   | I can see <b>big advantage</b> here.   |
|          | B19 | 1         | sub        | <b>A group</b> of students prepared this project.  | <b>Group</b> of students prepared this project.  |
|          | B20 | 1         | <u>pre</u> | This is definitely <b>a new trend</b> .  | This is definitely <b>new trend</b> .  |
| 14       | B21 | 2         | <u>ex</u>  | There is <b>low trust</b> in this organisation.  | There is <b>a low trust</b> in this organisation.  |
|          | B22 | 1         | <u>obj</u> | We have improved <b>internal communication</b> in our team.  | We have improved <b>an internal communication</b> in our team.   |
|          | B23 | 2         | sub        | <b>Fitness</b> is helping me to lose weight.   | <b>A fitness</b> is helping me to lose weight.   |
|          | B24 | 2         | pre        | This is <b>poetry</b> , so you may like it.  | This is <b>a poetry</b> , so you may like it.  |
| 15       | B25 | 1         | ex         | There are <b>differences</b> between these books.  | There are <b>a differences</b> between these books.  |
|          | B26 | 2         | <u>obj</u> | They chose <b>different options</b> to try.  | They chose <b>a different options</b> to try.  |
|          | B27 | 2         | <u>sub</u> | <b>Short pauses</b> helped me get through the exam.  | <b>A short pauses</b> helped me get through the exam.  |
|          | B28 | 2         | <u>pre</u> | These are <b>new features</b> .  | These are <b>a new features</b> .  |
| 16       | B29 | 2         | ex         | There is <b>a club</b> in this area.   | There is <b>club</b> in this area.   |
|          | B30 | 1         | obj        | He broke <b>an arm</b> yesterday.  | He broke <b>arm</b> yesterday.   |
|          | B31 | 1         | <u>sub</u> | <b>A new computer</b> was installed in the IT room.  | <b>New computer</b> was installed in the IT room.  |
|          | B32 | 2         | pre        | Is she <b>a singer</b> ?   | Is she <b>singer</b> ?   |
| 17       | B33 | 1         | <u>ex</u>  | There is <b>strong alcohol</b> in this glass.  | There is <b>a strong alcohol</b> in this glass.  |
|          | B34 | 1         | obj        | I put <b>ice</b> in your lemonade.   | I put <b>an ice</b> in your lemonade.  |
|          | B35 | 2         | sub        | <b>Snow</b> covered the playground.  | <b>A snow</b> covered the playground.  |
|          | B36 | 1         | <u>pre</u> | This is <b>human blood</b> .   | This is <b>a human blood</b> .   |
| 18       | B37 | 2         | <u>ex</u>  | There were <b>rose bushes</b> in the garden.   | There were <b>a rose bushes</b> in the garden.   |

| Item No. | ID  | Test part | Synt & mod | Grammatical  | Ungrammatical  |
|----------|-----|-----------|------------|--|--|
|          | B38 | 2         | obj        | We bought <b>dressses</b> at this shop.                | We bought <b>a dresses</b> at this shop.             |
|          | B39 | 1         | <u>sub</u> | <b>Big boxes</b> were blocking the door.               | <b>A big boxes</b> were blocking the door.           |
|          | B40 | 1         | pre        | They are <b>builders</b> .                             | They are <b>a builders</b> .                         |
| 19       | B41 | 1         | <u>ex</u>  | There is <b>a historical centre</b> in most towns.     | There is <b>historical centre</b> in most towns.     |
|          | B42 | 2         | obj        | Do you make <b>a copy</b> of each new document?        | Do you make <b>copy</b> of each new document?        |
|          | B43 | 1         | <u>sub</u> | <b>A new name</b> is added to this list every day.     | <b>New name</b> is added to this list every day.     |
| 20       | B44 | 2         | ex         | There is always <b>magic</b> in the air.               | There is always <b>a magic</b> in the air.           |
|          | B45 | 1         | obj        | They often play <b>baseball</b> after school.          | They often play <b>a baseball</b> after school.      |
|          | B46 | 2         | <u>sub</u> | <b>Personal finance</b> is very important.             | <b>A personal finance</b> is very important.         |
| 21       | B47 | 1         | ex         | There are <b>clues</b> in each book.                   | There are <b>a clues</b> in each book.               |
|          | B48 | 2         | <u>obj</u> | Listening to <b>different accents</b> is useful.       | Listening to <b>a different accents</b> is useful.   |
|          | B49 | 2         | <u>sub</u> | Are <b>final exams</b> usually difficult?              | Are <b>a final exams</b> usually difficult?          |
| 22       | B50 | 2         | <u>ex</u>  | Usually there is <b>a new painting</b> on display.     | Usually there is <b>new painting</b> on display.     |
|          | B51 | 1         | <u>obj</u> | I often have <b>a relaxing bubble bath</b> after work. | I often have <b>relaxing bubble bath</b> after work. |
|          | B52 | 1         | sub        | <b>A beach</b> is a great place to relax.              | <b>Beach</b> is a great place to relax.              |
| 23       | B53 | 2         | <u>ex</u>  | There is always <b>brown rice</b> in the cupboard.     | There is always <b>a brown rice</b> in the cupboard. |
|          | B54 | 2         | obj        | Do they use <b>plastic</b> in their products?          | Do they use <b>a plastic</b> in their products?      |
|          | B55 | 1         | sub        | <b>Silver</b> is often used to make mirrors.           | <b>A silver</b> is often used to make mirrors.       |
| 24       | B56 | 1         | ex         | There are always <b>markers</b> on my desk.            | There are always <b>a markers</b> on my desk.        |
|          | B57 | 1         | <u>obj</u> | You won't see <b>green hills</b> anywhere.             | You won't see <b>a green hills</b> anywhere.         |
|          | B58 | 2         | sub        | <b>Buses</b> usually arrive on time.                   | <b>A buses</b> usually arrive on time.               |

Table 29

Main Intervention GJT Test C (Target Items Only)

| Item No. | ID  | Test part | Synt & mod | Grammatical  | Ungrammatical  |
|----------|-----|-----------|------------|--|--|
| 1        | C01 | 2         | <u>sub</u> | Wellington won the Battle of Waterloo. <b>The famous victory</b> was so important that it ended Napoleon's rule. | Wellington won the Battle of Waterloo. <b>A famous victory</b> was so important that it ended Napoleon's rule. |



| Item No. | ID  | Test part | Synt & mod | Grammatical   | Ungrammatical   |
|----------|-----|-----------|------------|---|---|
| 2        | C02 | 2         | obj        | Have you read <b>the research</b> that is reported in this article?   | Have you read <b>research</b> that is reported in this article?   |
|          | C03 | 1         | <u>sub</u> | <b>The new equipment</b> that we bought is much better.   | <b>New equipment</b> that we bought is much better.   |
| 3        | C04 | 1         | sub        | So, are Tom and Katie going to act in this film?- No, they said <b>the roles</b> were not interesting for them.               | So, are Tom and Katie going to act in this film?- No, they said <b>roles</b> were not interesting for them.                     |
| 4        | C05 | 1         | obj        | I dropped my phone and cracked <b>the screen</b> .  | I dropped my phone and cracked <b>a screen</b> .  |
| 5        | C06 | 2         | <u>obj</u> | There is good and bad stuff in your partner, just like in any other person. You just need to focus on <b>the good stuff</b> . | There is good and bad stuff in your partner, just like in any other person. You just need to focus on <b>good stuff</b> .       |
| 6        | C07 | 1         | <u>obj</u> | Have you seen <b>the beautiful gardens</b> that our tour guide described?   | Have you seen <b>beautiful gardens</b> that our tour guide described?   |
|          | C08 | 2         | sub        | Most parts of the plane were fine after the accident. Only <b>the wings</b> were broken.                                      | Most parts of the plane were fine after the accident. Only <b>wings</b> were broken.  |
| 7        | C09 | 1         | <u>sub</u> | What happened? - <b>A large deposit</b> was made into our account.  | What happened? - <b>The large deposit</b> was made into our account.  |
| 8        | C10 | 2         | <u>obj</u> | There was a teacher in my school who gave me <b>individual attention</b> .  | There was a teacher in my school who gave me <b>the individual attention</b> .  |
| 9        | C11 | 2         | sub        | Usually they hire managers from inside the company, but <b>offers</b> to a few outside candidates have been made this year.   | Usually they hire managers from inside the company, but <b>the offers</b> to a few outside candidates have been made this year. |
| 10       | C12 | 1         | obj        | I bought <b>an album</b> of some local music in that souvenir shop.   | I bought <b>the album</b> of some local music in that souvenir shop.  |
|          | C13 | 2         | <u>sub</u> | Have you heard? <b>A completely new element</b> has been created!   | Have you heard? <b>The completely new element</b> has been created!   |
| 11       | C14 | 1         | obj        | Hi, I'm looking for some shower gel. Have you got any? - Sorry, we only have <b>soap</b> at the moment.                       | Hi, I'm looking for some shower gel. Have you got any? - Sorry, we only have <b>the soap</b> at the moment.                     |

| Item No. | ID  | Test part | Synt & mod | Grammatical   | Ungrammatical   |
|----------|-----|-----------|------------|---|---|
| 12       | C15 | 1         | <u>obj</u> | Excuse me, are there any potatoes on the menu? - I'm not sure. They have <b>roast vegetables</b> , but I don't know which ones.               | Excuse me, are there any potatoes on the menu? - I'm not sure. They have <b>the roast vegetables</b> , but I don't know which ones.               |
|          | C16 | 2         | sub        | Tell me about a scary moment in your life. - I was walking one night when suddenly <b>dogs</b> came out of nowhere. I didn't know what to do! | Tell me about a scary moment in your life. - I was walking one night when suddenly <b>the dogs</b> came out of nowhere. I didn't know what to do! |
| 13       | C17 | 2         | <u>ex</u>  | There was <b>a terrible accident</b> .  | There was terrible accident.  |
|          | C18 | 2         | obj        | It took <b>a century</b> to build this cathedral.   | It took <b>century</b> to build this cathedral.   |
|          | C19 | 1         | <u>sub</u> | <b>A PowerPoint presentation</b> was prepared by our sales team.  | <b>PowerPoint presentation</b> was prepared by our sales team.  |
|          | C20 | 2         | pre        | This is <b>a company</b> with high standards of quality.  | This is <b>company</b> with high standards of quality.  |
| 14       | C21 | 1         | ex         | There is <b>support</b> if you need it.   | There is <b>a support</b> if you need it.   |
|          | C22 | 2         | obj        | I am offering <b>help</b> .   | I am offering <b>a help</b> .   |
|          | C23 | 1         | <u>sub</u> | <b>Good health</b> is the most valuable thing I have.   | <b>A good health</b> is the most valuable thing I have.   |
|          | C24 | 2         | pre        | Is this <b>happiness</b> ?  | Is this <b>a happiness</b> ?  |
| 15       | C25 | 1         | <u>ex</u>  | There are <b>serious problems</b> at work.  | There are <b>a serious problems</b> at work.  |
|          | C26 | 1         | <u>obj</u> | They took <b>long breaks</b> yesterday.   | They took <b>a long breaks</b> yesterday.   |
|          | C27 | 2         | sub        | <b>Lies</b> only made his situation worse.  | <b>A lies</b> only made his situation worse.  |
|          | C28 | 1         | <u>pre</u> | These are <b>interesting interviews</b> .   | These are <b>an interesting interviews</b> .  |
| 16       | C29 | 1         | <u>ex</u>  | There is <b>a nice table</b> in the room.   | There is <b>nice table</b> in the room.   |
|          | C30 | 2         | obj        | She bought <b>an apple</b> .  | She bought <b>apple</b> .   |
|          | C31 | 2         | sub        | <b>A ball</b> is flying over the fence.   | <b>Ball</b> is flying over the fence.   |
|          | C32 | 1         | <u>pre</u> | She is <b>a great actress</b> .   | She is <b>great actress</b> .   |
| 17       | C33 | 2         | ex         | There is <b>milk</b> in the fridge.   | There is <b>a milk</b> in the fridge.   |
|          | C34 | 1         | <u>obj</u> | This risotto tastes great! What's your secret? - First, I toasted the rice, and then I added <b>white wine</b> .                              | This risotto tastes great! What's your secret? - First, I toasted the rice, and then I added <b>a white wine</b> .                                |
|          | C35 | 1         | <u>sub</u> | <b>Standard bronze</b> was used for this sculpture.   | <b>A standard bronze</b> was used for this sculpture.   |
|          | C36 | 1         | pre        | This is <b>butter</b> .   | This is <b>a butter</b> .   |

| Item No. | ID  | Test part | Synt & mod | Grammatical  | Ungrammatical  |
|----------|-----|-----------|------------|--|--|
| 18       | C37 | 2         | ex         | There are <b>cars</b> in the car park.                                   | There are <b>a cars</b> in the car park.                               |
|          | C38 | 1         | <u>obj</u> | I saw <b>big houses</b> .  | I saw <b>a big houses</b> .  |
|          | C39 | 2         | sub        | <b>Cups</b> are in the cupboard.   | <b>A cups</b> are in the cupboard.                                     |
|          | C40 | 2         | <u>pre</u> | They are <b>good doctors</b> .   | They are <b>a good doctors</b> .                                       |
| 19       | C41 | 1         | ex         | There is always <b>a chance</b> that you will win.                       | There is always <b>chance</b> that you will win.                       |
|          | C42 | 2         | <u>obj</u> | I hope you find <b>an interesting topic</b> to discuss.                  | I hope you find <b>interesting topic</b> to discuss.                   |
|          | C43 | 2         | sub        | If you are allergic to something, <b>a recipe</b> can always be changed. | If you are allergic to something, <b>recipe</b> can always be changed. |
| 20       | C44 | 2         | <u>ex</u>  | There is usually <b>heavy traffic</b> on this highway.                   | There is usually <b>a heavy traffic</b> on this highway.               |
|          | C45 | 2         | obj        | They need <b>privacy</b> .   | They need <b>a privacy</b> .   |
|          | C46 | 1         | sub        | <b>Optimism</b> makes my life easier.                                    | <b>An optimism</b> makes my life easier.                               |
| 21       | C47 | 2         | <u>ex</u>  | There are always <b>clever ideas</b> at these conferences.               | There are always <b>a clever ideas</b> at these conferences.           |
|          | C48 | 1         | <u>obj</u> | I don't like <b>long lectures</b> .                                      | I don't like <b>a long lectures</b> .                                  |
|          | C49 | 1         | sub        | <b>Motives</b> are often difficult to understand.                        | <b>A motives</b> are often difficult to understand.                    |
| 22       | C50 | 1         | ex         | There is <b>a bank</b> in each district.                                 | There is <b>bank</b> in each district.                                 |
|          | C51 | 2         | obj        | Do you need <b>an umbrella</b> ?   | Do you need <b>umbrella</b> ?  |
|          | C52 | 1         | <u>sub</u> | If <b>a phone case</b> doesn't fit, just return it.                      | If <b>phone case</b> doesn't fit, just return it.                      |
| 23       | C53 | 1         | ex         | There is always <b>wood</b> in our fireplace.                            | There is always <b>a wood</b> in our fireplace.                        |
|          | C54 | 1         | <u>obj</u> | I will buy <b>table salt</b> .   | I will buy <b>a table salt</b> .                                       |
|          | C55 | 2         | <u>sub</u> | <b>Intense heat</b> is always harmful for plants.                        | <b>An intense heat</b> is always harmful for plants.                   |
| 24       | C56 | 2         | <u>ex</u>  | There may be <b>fried eggs</b> for breakfast.                            | There may be <b>a fried eggs</b> for breakfast.                        |
|          | C57 | 1         | obj        | He often reads <b>newspapers</b> in the morning.                         | He often reads <b>a newspapers</b> in the morning.                     |
|          | C58 | 2         | <u>sub</u> | <b>Diamond earrings</b> are very expensive.                              | <b>A diamond earrings</b> are very expensive.                          |

Table 30

## Main Intervention EI Test A Structure and Items

| Def        | Spec     | Abstr | Ntype | Synt & mod  | Test part | Prompt  |
|------------|----------|-------|-------|-------------|-----------|---|
| definite   | specific | abstr | sing  | obj         | 1         | I started to understand <b>the process</b> of making films much better after I did an internship in Hollywood.                      |
|            |          |       | mass  | <u>obj</u>  | 2         | If you cannot use <b>the theoretical knowledge</b> that you got at university, you may not find a good job.                         |
|            |          | concr | pl    | sub         | 2         | I'm sorry, but <b>the aims</b> of this project are just not possible to achieve if the deadline is next week.                       |
|            |          |       | sing  | sub         | 2         | I'm happy to say that <b>the catalogue</b> that you ordered last week is almost finished and will be printed tomorrow.              |
|            |          |       | mass  | sub         | 1         | I wasn't very happy with breakfast this morning, because <b>the coffee</b> wasn't good and there were no fresh fruit or vegetables. |
|            |          |       | pl    | <u>obj</u>  | 1         | Check <b>the kitchen cupboard</b> if you need anything for that cake you were planning to make for Caroline's birthday.             |
| indefinite | specific | abstr | sing  | <u>ex</u>   | 1         | There is <b>a great concert</b> tonight, which I thought might be interesting for you if you like modern classical music.           |
|            |          |       | sing  | obj         | 2         | We booked <b>a tour</b> to New Zealand after my brother told me that the Lord of the Rings trilogy was filmed there.                |
|            |          |       | mass  | sub         | 1         | I can tell you that <b>popularity</b> has changed him a lot, because I knew him before he became a star.                            |
|            |          |       | mass  | <u>obj</u>  | 2         | We had <b>great enthusiasm</b> when the project started, but we were very demotivated by the financial problems.                    |
|            |          |       | mass  | <u>pred</u> | 1         | I know this is <b>modern architecture</b> , but I hate it, because it doesn't go well with all the other buildings.                 |
|            |          |       | pl    | ex          | 1         | When I was sixteen, there were <b>periods</b> when I was so angry at my parents I didn't even want to see them.                     |
|            |          |       | pl    | <u>sub</u>  | 2         | I'm glad that <b>good marks</b> helped me get into one of the best universities, where I met all my future business partners.       |
|            |          | concr | sing  | sub         | 2         | When I was younger, <b>a manager</b> gave me some great advice that has really helped me progress in my career.                     |
|            |          |       | sing  | <u>obj</u>  | 1         | I just had <b>a chocolate biscuit</b> and some tea this morning, because I couldn't find anything proper to eat.                    |
|            |          |       | sing  | pred        | 2         | Lucy is <b>a journalist</b> who has worked for the BBC for years and has a lot of experience in reporting.                          |
|            |          |       | mass  | ex          | 2         | There is <b>popcorn</b> in the kitchen if you decide to invite your friends over and watch something on Netflix.                    |
|            |          |       | mass  | <u>sub</u>  | 1         | Helen said that <b>real silk</b> was too expensive at that shop that you recommended, so we decided to order online.                |
|            |          |       | pl    | obj         | 1         | They bought <b>smartphones</b> directly from China, which they are now selling to local people cheaper than any other tech store.   |
|            |          |       | pl    | <u>ex</u>   | 2         | Can you explain why there are <b>dirty clothes</b> on the floor after I asked you three times to tidy up before dinner?             |
| no n       |          | abstr | sing  | <u>sub</u>  | 2         | Remember that <b>clear goals</b> make up fifty percent of your success, so don't start anything without planning first.             |

| Def | Spec | Abstr | Ntype | Synt & mod | Test part | Prompt  |
|-----|------|-------|-------|------------|-----------|---|
|     |      |       | sing  | ex         | 1         | If there is a <b>deficit</b> , we will have to cut down on our costs, for example move to a smaller office.                       |
|     |      |       | mass  | obj        | 2         | He always shows <b>respect</b> to everyone, so it was a big shock for me that he got into such a scandal.                         |
|     |      |       | mass  | <u>ex</u>  | 1         | There is always <b>great wisdom</b> in folklore, which has been passed on from generation to generation for centuries.            |
|     |      |       | pl    | <u>obj</u> | 1         | If you set <b>unrealistic targets</b> , you will find yourself very demotivated when you cannot reach them.                       |
|     |      |       | pl    | sub        | 2         | Of course, <b>prices</b> are important for those customers who have a limited budget, but we can't make more discounts.           |
|     |      | concr | sing  | obj        | 1         | We don't have a <b>cleaner</b> to tidy up your mess every day, which is only getting worse during school holidays.                |
|     |      |       | sing  | <u>ex</u>  | 2         | There is probably a <b>French bakery</b> on every corner if this place is so popular with tourists.                               |
|     |      |       | mass  | sub        | 1         | You should know that <b>dust</b> is a big problem for us, because our windows are looking over a very busy street.                |
|     |      |       | mass  | <u>obj</u> | 2         | In many places people don't have access to <b>clean water</b> , which makes their lives more difficult than we can imagine.       |
|     |      |       | pl    | <u>sub</u> | 1         | Sadly, <b>poor countries</b> cannot spend their resources on protecting the environment, because they have more important issues. |
|     |      |       | pl    | ex         | 2         | There are <b>atoms</b> of carbon in all living things, as I'm sure you remember from your school chemistry lessons.               |

Table 31

## Main Intervention EI Test B Structure and Items

| Def        | Spec     | Abstr | Ntype | Synt & mod  | Test part | Prompt  |
|------------|----------|-------|-------|-------------|-----------|---|
| definite   | specific | abstr | sing  | sub         | 2         | I understand that <b>the situation</b> is difficult, but I don't know when I will be able to finish this work.                |
|            |          |       | mass  | obj         | 2         | You must lower <b>the concentration</b> of gas in these tubes before we can start testing the new equipment.                  |
|            |          | concr | pl    | <u>sub</u>  | 1         | I thought <b>the installation instructions</b> for this device looked very simple, but now I know that's not the case.        |
|            |          |       | sing  | <u>obj</u>  | 1         | When I visited Prague, I loved walking around <b>the old town</b> and taking photos in the old narrow streets.                |
|            |          |       | mass  | obj         | 1         | To be honest, I don't like <b>the material</b> that this coat is made of, because it looks cheap, so I'm not going to buy it. |
|            |          |       | pl    | <u>obj</u>  | 2         | What did you think of <b>the leading actors</b> from that film which we saw last Saturday at Susan's birthday party?          |
| indefinite | specific | abstr | sing  | <u>obj</u>  | 2         | We are forgetting <b>an important factor</b> , which may change our decision completely if we take it seriously.              |
|            |          |       | sing  | sub         | 1         | After yesterday's meeting, <b>a plan</b> was prepared and presented to the management, who agreed to support it.              |
|            |          |       | mass  | <u>sub</u>  | 1         | Unfortunately, <b>sudden acceleration</b> made us stop the experiment, because we had to cool down the machines.              |
|            |          |       | mass  | ex          | 2         | There is <b>tennis</b> on TV later, which I would really like to watch, because it's the Wimbledon men's final.               |
|            |          |       | pl    | obj         | 2         | We are doing <b>experiments</b> later today, after which we will finally have the answers to some important questions.        |
|            |          |       | pl    | <u>ex</u>   | 1         | There are <b>serious errors</b> in your test, which you should definitely discuss with your teacher before the exam.          |
|            |          | concr | pl    | pred        | 1         | I know that these are just <b>fantasies</b> , which can never come true, but I still cannot stop thinking about them.         |
|            |          |       | sing  | <u>sub</u>  | 2         | In fact, <b>a small team</b> of specialists is already working on this problem, which should help us find a solution soon.    |
|            |          |       | sing  | ex          | 1         | There is <b>a bottle</b> of white wine in the fridge if you'd like to offer some to our guests over dinner.                   |
|            |          |       | mass  | obj         | 2         | I paid <b>cash</b> at the supermarket yesterday, because I couldn't remember the new pin for my credit card.                  |
|            |          |       | mass  | <u>ex</u>   | 2         | There is <b>olive oil</b> on the top left shelf above the sink, which you can use for cooking or put in salads.               |
|            |          |       | mass  | <u>pred</u> | 1         | This jacket is <b>real leather</b> , so it's not surprising that it's more expensive than the jackets we saw at the market.   |
|            |          |       | pl    | <u>obj</u>  | 1         | My mother gave us <b>silver teaspoons</b> for our wedding anniversary, which go nicely with our silver tea set.               |
|            |          |       | pl    | sub         | 2         | Did you know that <b>shops</b> in this area sell cheese from local farmers, who have used the same recipe for centuries.      |
|            |          |       | no n- | abstr       | sing      | <u>ex</u>   |

| Def | Spec | Abstr | Ntype | Synt & mod | Test part | Prompt  |
|-----|------|-------|-------|------------|-----------|---|
|     |      |       | sing  | obj        | 2         | I can't think of <b>a theory</b> that would explain these results, so I think I'll need to read more literature.              |
|     |      |       | mass  | sub        | 1         | Of course, <b>evolution</b> can explain how life has developed, but we still don't fully understand how it began.             |
|     |      |       | mass  | <u>obj</u> | 2         | Laws must protect <b>human freedom</b> , because politicians and governments often try to take it away.                       |
|     |      |       | pl    | <u>sub</u> | 2         | These days <b>holiday cruises</b> are much cheaper, because there are lots of companies offering them for any kind of budget. |
|     |      |       | pl    | ex         | 1         | There are <b>rules</b> for everyone, so you cannot come in late every morning just because you are a big boss now.            |
|     |      | concr | sing  | <u>obj</u> | 1         | I can probably get <b>an extra ticket</b> for tomorrow's show if you would like to take your new friend with us.              |
|     |      |       | sing  | sub        | 2         | Did you know that <b>a professor</b> cannot be fired unless their university finds a serious reason for it?                   |
|     |      |       | mass  | <u>sub</u> | 1         | Usually <b>strawberry jam</b> on toast and a cup of tea is enough for me, because I'm not very hungry for breakfast.          |
|     |      |       | mass  | ex         | 2         | There is <b>electricity</b> almost everywhere in Europe and America, which is not true for many African countries.            |
|     |      |       | pl    | obj        | 1         | People think they can trust <b>scientists</b> , but in fact there are so many things that they cannot be sure about.          |
|     |      |       | pl    | <u>ex</u>  | 2         | I'm sure there are <b>rich investors</b> out there, who would be very interested in a business idea like this.                |

Table 32

## Main Intervention EI Test C Structure and Items

| Def        | Spec     | Abstr | Ntype | Synt & mod | Test part | Prompt   |
|------------|----------|-------|-------|------------|-----------|--|
|            |          | abstr | sing  | <u>sub</u> | 2         | Your report says <b>the chemical reaction</b> in the first experiment started as normal but then suddenly stopped.             |
|            |          |       | mass  | obj        | 1         | He liked <b>the advice</b> that he got during the in-company training on time management, but he doesn't use it.               |
|            |          |       | pl    | sub        | 2         | It's strange that <b>the properties</b> of water, which have been studied for many years, are still not fully understood.      |
| definite   | specific | concr | sing  | obj        | 1         | Have you seen <b>the menu</b> of that new Mexican restaurant in the city centre which your brother recommended?                |
|            |          |       | mass  | <u>obj</u> | 2         | I've tried <b>the milk chocolate</b> that my sister brought me from Switzerland, but I didn't think it was special in any way. |
|            |          |       | pl    | <u>sub</u> | 1         | I think <b>the new buildings</b> in the city centre are just horrible, because they don't match the style of the old town.     |
| indefinite | specific | abstr | sing  | ex         | 1         | There is <b>a solution</b> that I think will work in our case but only if we can get the support of our partners.              |
|            |          |       | sing  | <u>sub</u> | 2         | Detectives arrived quickly, but <b>a huge mess</b> was left in the room, which made it very difficult to find anything.        |

| Def | Spec | Abstr        | Ntype | Synt & mod  | Test part | Prompt  |
|-----|------|--------------|-------|-------------|-----------|---|
|     |      |              | sing  | pred        | 1         | This is <b>an environment</b> where you won't be afraid to make mistakes or ask questions even if they seem silly.              |
|     |      |              | mass  | obj         | 2         | We finally have <b>access</b> to the Internet, so we can watch that new series on Netflix which your sister recommended.        |
|     |      |              | mass  | <u>ex</u>   | 1         | There was <b>great luxury</b> at the president's villa, which made a lot of people in his country very angry.                   |
|     |      |              | pl    | sub         | 2         | It's true that <b>standards</b> are very high in this field, but this doesn't mean that I will give up without even trying.     |
|     |      |              | pl    | <u>obj</u>  | 1         | I have <b>important tasks</b> to work on, and two of them must be finished by tomorrow if I don't want to lose my job.          |
|     |      | concr        | sing  | <u>ex</u>   | 2         | There is <b>a nice coffee shop</b> across the street, which was opened only a month ago by an Italian family.                   |
|     |      |              | sing  | obj         | 1         | I got <b>a belt</b> for my brother's birthday because I know he needs it, but he never has the time to go shopping.             |
|     |      |              | mass  | <u>obj</u>  | 2         | Volunteers are collecting <b>plastic rubbish</b> in the park today, which I thought was a great idea, so I'm joining, too.      |
|     |      |              | mass  | sub         | 1         | He said that <b>cotton</b> was used to make all the T-shirts, but he couldn't remember which material was used for the dresses. |
|     |      |              | pl    | <u>sub</u>  | 1         | Fortunately, <b>marketing specialists</b> are already working on a new promotion plan, which should increase sales.             |
|     |      |              | pl    | ex          | 2         | There are <b>factories</b> in every town in this region, which is why the air quality is so bad around here.                    |
|     |      |              | pl    | <u>pred</u> | 2         | These are <b>new components</b> , which haven't been tested yet, but I can guarantee that they will work perfectly.             |
|     |      | non-specific | sing  | <u>obj</u>  | 1         | This travel agency can plan <b>a wonderful vacation</b> for you even if you don't have a lot of money to spend.                 |
|     |      |              | sing  | sub         | 2         | Generally, <b>an essay</b> should be marked by two examiners, who don't know the identity of the student who wrote it.          |
|     |      |              | mass  | <u>sub</u>  | 2         | As you know, <b>high quality</b> is our priority, which means our products must be carefully tested before we sell them.        |
|     |      |              | mass  | ex          | 1         | There is always <b>confidence</b> in his voice, because he knows he is the most qualified expert in the field.                  |
|     |      |              | pl    | obj         | 1         | I usually hate <b>puzzles</b> , but this one was so interesting that I decided to try and solve it together with my friends.    |
|     |      |              | pl    | <u>ex</u>   | 2         | There are <b>basic elements</b> in every story, which you should analyse if you want to write a book or film review.            |
|     |      | concr        | sing  | <u>sub</u>  | 1         | I think <b>a responsible person</b> will never promise to do anything if they know they cannot do it.                           |
|     |      |              | sing  | ex          | 2         | There is always <b>a pound</b> in my pocket when I go to a supermarket, but this new jacket has no pockets.                     |
|     |      |              | mass  | obj         | 2         | I always use <b>honey</b> instead of sugar, because I believe it's better for your health and it also tastes great.             |
|     |      |              | mass  | <u>ex</u>   | 1         | There is usually <b>French toast</b> for breakfast, which is my favourite, but you can also get a bowl of cereal or yoghurt.    |



| Def | Spec | Abstr | Ntype | Synt &<br>mod | Test<br>part | Prompt  |
|-----|------|-------|-------|---------------|--------------|---|
|     |      |       | pl    | sub           | 1            | If we agree that <b>teachers</b> are very important in our lives, why do we forget to thank them for their hard work? |
|     |      |       | pl    | <u>obj</u>    | 2            | I would love to work with <b>film directors</b> who have original ideas and who are not afraid to try new things.     |

Table 33

## Main Intervention Writing Prompts

|                        | Version A   | Version B   | Version C  |
|------------------------|---|---|--|
| Test part 1: Narrative | Watch this cartoon about Simon's Cat (Hambush):<br><a href="https://youtu.be/TUBK4qUGo_gw">https://youtu.be/TUBK4qUGo_gw</a>  | Watch this cartoon about Simon's Cat (Laundry room):<br><a href="https://youtu.be/L8qOb8mdT_U">https://youtu.be/L8qOb8mdT_U</a>   | Watch this cartoon about Simon's Cat (Hot water):<br><a href="https://youtu.be/ApN73TUVMEU">https://youtu.be/ApN73TUVMEU</a>   |
|                        | Your friend hasn't seen it. Tell him/her what happened in the cartoon.<br>Make sure to use these words: <i>plate, bread, fridge, ham, broom</i> .<br>Write about 100 - 150 words. | Your friend hasn't seen it. Tell him/her what happened in the cartoon.<br>Make sure to use these words: <i>spider, laundry, broom, washing machine, rubbish</i> .<br>Write about 100 - 150 words. | Your friend hasn't seen it. Tell him/her what happened in the cartoon.<br>Make sure to use these words: <i>shower, toilet paper, water, toothbrush, faucet</i> .<br>Write about 100 - 150 words. |
| Test part 2: Essay     | Discuss this question:<br><i>Are social media (for example, Facebook, Instagram etc.) a good source of news? Why (not)?</i><br>Write 100-150 words.                               | Discuss this question:<br><i>What are the laws about smoking and alcohol in your country? Do you think they are fair? Why (not)?</i><br>Write 100-150 words.                                      | Discuss this question:<br><i>What are the COVID restrictions or rules in your country? Do you think they are fair? Why (not)?</i><br>Write 100-150 words.  |

## Appendix K: Main Intervention Information Sheet, Consent Form, Background Questionnaire, and C-test

# Info, consent, survey and c-test

---

### Start of Block: Information Sheet

#### Cambridge University Research Project in Language Learning

Would you like to participate in a research project by the University of Cambridge and make your own contribution to the study of second language learning?

As a bonus, you could **get personalised feedback on your writing from a Cambridge researcher**, as well as a **Certificate of Participation from the University of Cambridge!**

I am a PhD student at the University of Cambridge studying how people learn English and exploring teaching methods and techniques that can help learners improve their English. Please find more information about this project below and take your time to decide if you would like to participate. I hope you agree to support our project, which can help other English learners like you to improve their learning outcomes.

#### Why have I been chosen?

The participants of this study are young adults who are learning English as a foreign language. In other words, you can participate if English is NOT your native language. You also need to have completed your **16th birthday** to participate in our study.

#### Do I have to take part?

Participation is entirely voluntary, so you do not have to participate in this project. If you agree to participate, you can change your mind at any time and drop out of the project without any penalty.

#### What will happen to me if I take part?

The entire study is carried out online through accessing a website at Cambridge University.

You will need to complete **15 learning activities** each taking about **10-15 minutes**. You can do the learning activities individually on any device at any convenient time over a period of **3 weeks**.

Before you start and after you finish, you will need to take a **language test online (two 45-minute sessions)**.

There will be another language test 3 months after you complete your training.

The activities and tests include multiple choice questions and short writing tasks.

You will receive links to all the training sessions and tests via e-mail after you have read this information sheet and have consented to participate.

#### What will I get?

You will receive a certificate of participation from the principal investigator of the project at the University of Cambridge if you complete all the practice activities and tests. After the final test you will be able to request

personal feedback on your writing from the researcher at Cambridge.

**How do I start?**

To start participating in this project, click → at the bottom of this page to fill in the consent form.

**Are there possible disadvantages and/or risks in taking part?**

No.

**Will my taking part in this project be kept confidential?**

All the information we collect about you will be kept strictly confidential. All data will be identified only by a code, with personal details kept on a secure server with access only by the immediate research team.

**How long will my data be retained?**

In line with University policy, data will be kept till the completion of the project + 10 years. After this, it may be destroyed or sent (in anonymised form) to the University's digital Institutional Repository.

**What will happen to the results of the research project?**

The project is scheduled to be completed in 2021. Results will be presented at conferences and written up in journals. Results are normally presented in terms of groups of individuals. If any individual data are presented, the data will be completely anonymous, without any means of identifying the individuals involved. Anonymised data will be deposited to the Apollo---University of Cambridge Repository, <https://www.repository.cam.ac.uk/>, for future research and study.

**Ethical review of the study**

The project has received ethical approval from the Research Ethics Committee of the Faculty of Modern and Medieval Languages and Linguistics at the University of Cambridge. Thank you very much for your participation! If you have any questions about the project, the process of participation, technical issues or any other aspect of this study, please contact me via e-mail.

Yours sincerely,

*Kateryna Derkach*

[Email address redacted]

25/01/2021

End of Block: Information Sheet

---

Start of Block: Consent form

**CONSENT FORM**

**Project title:** Processing Instruction Learning Activities

**Research team:** Kateryna Derkach and Theodora Alexopoulou

If you have any questions, please contact Kateryna Derkach, [Email address redacted]

Please read and tick all the boxes below.

☐

I confirm that I have read and understand the information sheet dated 25/01/2021 for the above mentioned study and have had the opportunity to ask questions.

☐

I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, and without my rights being affected.

☐

I understand that any data that are collected will be used and stored anonymously, in accordance with the Data Protection Act. Results are normally presented in terms of groups of individuals. If any individual data were presented, the data would be completely anonymous, without any means of identifying the individuals involved.

☐

I understand that these data may be used in analyses, publications, and conference presentations by researchers at the University of Cambridge and their collaborators at other research institutions. I give permission for these individuals to have access to these data.

☐

I understand that personal information (such as language background, age, gender, education level, length of stay in an English-speaking country) will be collected as part of this research. Full data will only be accessible to the research team. However, anonymised data may be used in analyses, publications and conference presentations. Anonymised data will be deposited to the Apollo---University of Cambridge Repository, <https://www.repository.cam.ac.uk/>, for future research and study. For full details on how researchers will use your personal information, see <https://www.information-compliance.admin.cam.ac.uk/data-protection/research-participant-data>

☐

I have been provided a copy of this form to keep. (This will be sent to your e-mail when you complete this form)

☐

I agree to participate in the above mentioned study run by Kateryna Derkach, a PhD student at the Faculty of MMLL at the University of Cambridge.

---

☐

I confirm that I am over 18 years old.

---

☐

I confirm that I am over 16 years old.

---

Please enter your first/given/forename.

---

Please enter your last/family/surname.

---

Today's date (dd/mm/yyyy)

---

Your signature

**Name of researcher:** Kateryna Derkach

**Date:** 25/01/2021

**Signature of researcher:**

*The project has received ethical approval from the Faculty's Research Ethics Committee. If you have any questions or complaints about the ethical aspects of this study, please contact [Email address redacted]*

End of Block: Consent form

Start of Block: Personal Information

In this section we will ask you to give us some personal information and some information about your language background. Everything you share will be kept confidential.

Your contact e-mail (we will use this to send you links to our activities)

☐ E-mail 

---

---

Please confirm your e-mail

---

---

Your age (in years)

---

---

Your gender

- ☐ female
- ☐ male
- ☐ other
- ☐ prefer not to say

---

What is the highest level of education that you have completed?

- ☐ High school
- ☐ Bachelor's degree or equivalent
- ☐ Master's degree or equivalent
- ☐ PhD or equivalent
- ☐ Other (please specify) \_\_\_\_\_
-

Are you a school or college student at the moment?

☐ Yes

☐ No

---

Please specify at which level you are studying at the moment.

☐ High school (indicate year) \_\_\_\_\_

☐ Bachelor's degree or equivalent (indicate year)  
\_\_\_\_\_

☐ Master's degree or equivalent

☐ PhD or equivalent

☐ Other (please specify) \_\_\_\_\_

End of Block: Personal Information

---

Start of Block: Language Background

What is your native language(s)?

\_\_\_\_\_

---

At what age did you start learning English? (*for example, at 6 years old*)

\_\_\_\_\_

---



How have you learned English up to now? Tick all the options that apply and indicate the details in the boxes provided.

☐

English classes at school: a) for how many years? b) how many hours/week?

---

☐

English classes outside of school, for example, in a language school or with a private tutor: a) for how many years? b) how many hours/week?

---

☐

Books, films, songs, other media: how many hours/week?

---

☐

Communication with English-speaking friends: how many hours/week?

---

☐

Other (please specify details) \_\_\_\_\_

---

What is your English level at the moment?

☐

Elementary (A1)

☐

Pre-intermediate (A2)

☐

Intermediate (B1)

☐

Upper-Intermediate (B2)

☐

Advanced (C1)

☐

Other (please specify) \_\_\_\_\_

---

Are you taking any English classes at the moment?

☐

Yes

☐

No

---

Please give more details.

☐ What type of classes are you taking? (*for example, lessons at school, language school course, private tutor etc.*) \_\_\_\_\_

☐ How long have you been taking these classes? (*for example, for 3 months*)  
\_\_\_\_\_

☐ How many hours a week do you usually spend in these classes?  
\_\_\_\_\_

-----

Have you got any English language test certificates? Please tick all that apply and indicate your results (you can give approximate results if you do not remember exactly).

☐ IELTS \_\_\_\_\_

☐ TOEFL \_\_\_\_\_

☐ PET \_\_\_\_\_

☐ FCE \_\_\_\_\_

☐ CAE \_\_\_\_\_

☐ PTE \_\_\_\_\_

☐ Other (please specify test and result)  
\_\_\_\_\_

-----

Have you learned or used any other languages, except your native language(s) and English?

☐ Yes

☐ No

-----

## Language 1

☐ Language \_\_\_\_\_

☐ Level \_\_\_\_\_

☐ How and at what ages did you learn/use it? *(For example: when I lived in Spain for 6 months when I was 12)* \_\_\_\_\_

---

## Language 2

☐ Language \_\_\_\_\_

☐ Level \_\_\_\_\_

☐ How and at what ages did you learn/use it? *(For example: when I lived in Spain for 6 months when I was 12)* \_\_\_\_\_

---

## Language 3

☐ Language \_\_\_\_\_

☐ Level \_\_\_\_\_

☐ How and at what ages did you learn/use it? *(For example: when I lived in Spain for 6 months when I was 12)* \_\_\_\_\_

---

If you speak any more languages, please give details in the box below.

---

---

---

Have you lived in an English-speaking country for more than 3 months?

☐ Yes

☐ No

---

Country 1

☐ Country \_\_\_\_\_

☐ At what ages did you live there? *(For example: from 15 to 17 years old)*  
\_\_\_\_\_

Country 2

☐ Country \_\_\_\_\_

☐ At what ages did you live there? *(For example: from 15 to 17 years old)*  
\_\_\_\_\_

Country 3

☐ Country \_\_\_\_\_

☐ At what ages did you live there? *(For example: from 15 to 17 years old)*  
\_\_\_\_\_

If you have lived in more English-speaking countries, please give details in the box below.

---

---

---

Is there anything else that you feel is interesting or important about your language background that you'd like us to know?

---

---

---

End of Block: Language Background

## Start of Block: C-test

The last step in your registration is a **short placement test**, so we can put you in the right group of participants.

You will have **10 minutes** to complete this test.

You will see two short texts with blanks. You need to fill in **one word in each blank**. The first letter is given to you.

**Example:**

*The girl was walking d\_\_\_ the street when she stepped on some ice and fell.*

**Answer:** down

When you are ready, click → to start the timer.

*(If you close this page, you can come back using the same link from your e-mail on the same device)*

## Text 1

Steven loved almost everything about his grandma. There was only one thing he hated. She always knitted sweaters for (1) h\_\_\_. Steven understood that she did it to be (2) n\_\_\_. However, all the sweaters were very ugly. Steven (3) v\_\_\_ her once a week. She had a new (4) s\_\_\_ for him each time.

- ☐ 1 \_\_\_\_\_
- ☐ 2 \_\_\_\_\_
- ☐ 3 \_\_\_\_\_
- ☐ 4 \_\_\_\_\_

Steven lived in a (5) s\_\_\_ apartment. There was no room for him to (6) k\_\_\_ all the sweaters. He had to give all of them (7) a\_\_\_. "Grandma will never find out," he thought.

- ☐ 5 \_\_\_\_\_
- ☐ 6 \_\_\_\_\_
- ☐ 7 \_\_\_\_\_

One (8) d\_\_\_\_, Steven's grandma visited him by surprise. She asked to (9) s\_\_\_\_ his sweaters. "Someone stole all of them!" he (10) s\_\_\_\_. "They were too nice." She (11) m\_\_\_\_ him ten more by the next month.

☐ 8 \_\_\_\_\_

☐ 9 \_\_\_\_\_

☐ 10 \_\_\_\_\_

☐ 11 \_\_\_\_\_

## Text 2

Nonverbal communication includes facial expressions, gestures, the distance between speakers, eye contact, voice intonations, touch, and many other minor details which can provide speakers with valuable details about each other. For example, (12) s\_\_\_\_ between people can say a lot about the level of intimacy between them: usually, the (13) s\_\_\_\_ the distance between speakers, the more friendly or (14) i\_\_\_\_ they are, and vice versa.

☐ 12 \_\_\_\_\_

☐ 13 \_\_\_\_\_

☐ 14 \_\_\_\_\_

Or if a person (15) a\_\_\_\_ eye contact, it might mean that he or she is hiding something, feels (16) u\_\_\_\_ around you, and so on.

☐ 15 \_\_\_\_\_

☐ 16 \_\_\_\_\_

Body (17) l\_\_\_ has several important functions. For instance, a person's (18) g\_\_\_ can repeat the message he or she is (19) m\_\_\_ orally; a little child explaining how birds (20) f\_\_\_ and waving his or her arms like (21) w\_\_\_ is a decent example of this function.

- ☐ 17 \_\_\_\_\_
- ☐ 18 \_\_\_\_\_
- ☐ 19 \_\_\_\_\_
- ☐ 20 \_\_\_\_\_
- ☐ 21 \_\_\_\_\_

Another function, substitution, occurs when (22) v\_\_\_ messages can be expressed by nonverbal means (like shrugging). (23) l\_\_\_ addition, gestures can be used for accenting, like when (24) r\_\_\_ one's index finger when speaking about (25) s\_\_\_ important.

- ☐ 22 \_\_\_\_\_
- ☐ 23 \_\_\_\_\_
- ☐ 24 \_\_\_\_\_
- ☐ 25 \_\_\_\_\_

At the same time, it is important to remember that sometimes body language may (26) d\_\_\_ depending on culture. For example, in some eastern countries, (27) l\_\_\_ straight in the eyes of a conversationalist is considered (28) r\_\_\_.

- ☐ 26 \_\_\_\_\_
- ☐ 27 \_\_\_\_\_
- ☐ 28 \_\_\_\_\_

Men in some Arabic countries may walk around the street (29) h\_\_\_ hands, or may kiss each other on the (30) c\_\_\_ when greeting, but this is the (31) i\_\_\_ of friendship, not romance or intimacy.

- ☐ 29 \_\_\_\_\_
- ☐ 30 \_\_\_\_\_
- ☐ 31 \_\_\_\_\_

## Appendix L: Main Intervention GJT Indefinite Mass and Plural Model Results

**Table 34**

*Main Intervention GJT Indefinite Mass and Plural Model Results*

| <i>Predictors</i>   | <b>Accuracy Rate</b>  |                   |               |
|---|---|-------------------|---------------|
|   | <i>Log-Odds</i>   | <i>std. Error</i> | <i>95% CI</i> |
| (Intercept)   | 1.63 ***  | 0.18              | 1.28   1.98   |
| group: Experimental   | 0.22  | 0.20              | -0.18   0.62  |
| test: immediate post  | 0.48 **   | 0.15              | 0.18   0.78   |
| mod: modified   | -0.01   | 0.16              | -0.32   0.31  |
| Ntype: mass   | -0.36 *   | 0.16              | -0.67   -0.05 |
| item variant: IC  | -0.73 ***   | 0.19              | -1.09   -0.36 |
| error type: over_the  | -0.11   | 0.27              | -0.64   0.43  |
| c-test  | 0.52 ***  | 0.09              | 0.35   0.70   |
| group: Experimental * test: immediate post                                  | 0.30  | 0.22              | -0.13   0.73  |
| group: Experimental * mod: modified   | -0.62 **  | 0.21              | -1.03   -0.21 |
| test: immediate post * mod: modified  | -0.40   | 0.21              | -0.82   0.02  |
| Ntype: mass * item variant: IC  | -1.29 ***   | 0.22              | -1.71   -0.86 |
| Ntype: mass * error type: over_the  | 0.48  | 0.39              | -0.29   1.25  |
| item variant: IC * error type: over_the                                     | -2.14 ***   | 0.38              | -2.89   -1.39 |
| group: Experimental * test: immediate post * mod: modified                  | 0.43  | 0.31              | -0.18   1.03  |
| Ntype: mass * item variant: IC * error type: over_the                       | 0.67  | 0.55              | -0.40   1.74  |
| <b>Random Effects</b>   |   |                   |               |
| <i>By item ID</i>   |   |                   |               |
| SD (Intercept)  | 0.43  |                   |               |
| SD (c-test)   | 0.31  |                   |               |
| <i>By participant ID</i>  |   |                   |               |
| SD (Intercept)  | 0.48  |                   |               |
| SD (Ntype: mass)  | 0.28  |                   |               |
| SD (item variant: IC)   | 0.86  |                   |               |
| N <sub>ppt_id</sub>   | 70  |                   |               |
| N <sub>item_id</sub>  | 200   |                   |               |
| Observations  | 4627  |                   |               |
| Marginal R <sup>2</sup> / Conditional R <sup>2</sup>                        | 0.329 / 0.329   |                   |               |
| Log-likelihood ratio test comparing to the null model (only random effects) | $\chi^2(15) = 315.63, p < 10^{-15}$   |                   |               |
| C-statistic   | 0.86 (strong predictive power)  |                   |               |
| VIFs  | < 5 for most terms, except two correlation terms, which have moderate values of 5 and 5.5 |                   |               |
| Overdispersion ratio  | 0.849 ( $\chi^2 = 3911.381, p = 1$ )  |                   |               |

Note. \*  $p < 0.05$  \*\*  $p < 0.01$  \*\*\*  $p < 0.001$



## Appendix M: Main Intervention GJT Indefinite Singular Model Results

**Table 35**

*Main Intervention GJT Indefinite Singular Model Results*

| <i>Predictors</i>   | <b>Accuracy Rate</b>  |                   |               |
|---|---|-------------------|---------------|
|   | <i>Log-Odds</i>   | <i>std. Error</i> | <i>95% CI</i> |
| (Intercept)   | 2.22 ***  | 0.38              | 1.47   2.98   |
| group: Experimental   | -0.68   | 0.40              | -1.46   0.11  |
| test: immediate post  | -0.26   | 0.40              | -1.05   0.53  |
| item variant: IC  | -2.24 ***   | 0.42              | -3.07   -1.41 |
| L1-type: [-art]   | -1.03 **  | 0.36              | -1.75   -0.32 |
| c-test  | 0.51 ***  | 0.08              | 0.35   0.66   |
| abstr: abstract   | -0.38 *   | 0.18              | -0.73   -0.02 |
| error type: omit  | 0.53 *  | 0.24              | 0.05   1.00   |
| group: Experimental * test: immediate post                                      | 0.23  | 0.54              | -0.83   1.29  |
| group: Experimental * item variant: IC  | 0.67  | 0.55              | -0.40   1.75  |
| test: immediate post * item variant: IC   | 0.28  | 0.51              | -0.71   1.28  |
| group: Experimental * L1-type: [-art]   | 1.40 **   | 0.52              | 0.39   2.41   |
| test: immediate post * L1-type: [-art]  | 0.61  | 0.51              | -0.38   1.60  |
| item variant: IC * L1-type: [-art]  | -0.25   | 0.51              | -1.24   0.75  |
| group: Experimental * test: immediate post * item variant: IC                   | 0.05  | 0.70              | -1.32   1.42  |
| group: Experimental * test: immediate post * L1-type: [-art]                    | -0.97   | 0.69              | -2.33   0.38  |
| group: Experimental * item variant: IC * L1-type: [-art]                        | -0.93   | 0.71              | -2.33   0.47  |
| test: immediate post * item variant: IC * L1-type: [-art]                       | -0.65   | 0.65              | -1.92   0.63  |
| group: Experimental * test: immediate post * item variant: IC * L1-type: [-art] | 1.35  | 0.90              | -0.40   3.11  |
| <b>Random Effects</b>   |   |                   |               |
| <i>By item ID</i>   |   |                   |               |
| SD (Intercept)  | 0.71  |                   |               |
| <i>By participant ID</i>  |   |                   |               |
| SD (Intercept)  | 0.29  |                   |               |
| SD (item variant: IC)   | 0.65  |                   |               |
| N <sub>ppt_id</sub>   | 70  |                   |               |
| N <sub>item_id</sub>  | 100   |                   |               |
| Observations  | 2323  |                   |               |
| Marginal R <sup>2</sup> / Conditional R <sup>2</sup>                            | 0.314 / 0.406   |                   |               |
| Log-likelihood ratio test comparing to the null model (only random effects)     | $\chi^2(18) = 143.68, p < 10^{-15}$   |                   |               |
| C-statistic   | 0.86 (strong predictive power)  |                   |               |
| VIFs  | High in a model with correlations but low (= 1) in a model without correlations |                   |               |
| Overdispersion ratio  | 0.848 ( $\chi^2 = 1951.722, p = 1$ )  |                   |               |

Note. \*  $p < 0.05$  \*\*  $p < 0.01$  \*\*\*  $p < 0.001$