CLI and Transfer in a Trilingual Context: Acquisition, Development, and Maintenance of L3 German Grammatical Gender

Dissertation Prospectus

Megan Marisa Brown

List of Abbreviations

AoA	Age of Acquisition
CEM	Cumulative Enhancement Model
CITH	Cumulative Input Threshold Hypothesis
CLI	Cross-Linguistic Influence
CP	Critical Period
GM	Grammatical Mapping
L1	First Language
L1P	L1 Privilege Model
L2	Second Language
L2S	L2 Status Factor Model
L3	Third Language
LEAP-Q	Language Experience and Proficiency Questionnaire
Ln	An additional language beyond the L3 (i.e. L4, L5...)
LPM	Linguistic Proximity Model
OoA	Order of Acquisition
PLCM	Principal Language of Communication Model
SM	Scalpel Model
TLA	Third Language Acquisition
TPM	Typological Primacy Model
UG	Universal Grammar

Contents

1   Introduction & Background

It has long been established that when adults learn a new language, the mistakes that these learners make are not random, but systematic (i.e. Adémian, 1976; Corder, 1967; Nemser, 1971; Selinker, 1972), and in many cases their errors can be attributed to the influence of the first or native language (L1) grammar (Hawkins and Chan, 1997; Smith and Tsimpli, 1995). The phenomenon of one of a bilingual or multilingual's languages affecting another is known as cross-linguistic influence (CLI). CLI is a broad term used to refer to influence between any two languages in a multilingual's system at any point throughout the speaker's lifetime and learning experience. However, when this CLI occurs specifically from a previously known language onto the very earliest stages of new language learning, it is often referred to as "transfer". The transfer of L1 grammatical features at the early stages of second language (L2) learning is a topic that has been extensively discussed in the literature (e.g. Flynn, 1981; White, 2003).

In L2 acquisition research, the source of transfer to the L2 is relatively straightforward: the L1 is the only previously known language from which the L2 can transfer previously acquired grammatical features. However, in a case where a bilingual then goes on to acquire a third language (L3), the process of transfer to the L3 is much more complicated. The L3 has two potential source languages from which it could theoretically transfer linguistic information. To date, there remains a level of debate in the field regarding which language(s) are able to transfer at the initial stages of L3 acquisition, as well as the degree to which additional variables such as age of acquisition (AoA), (psycho-)typological similarity of the languages in question, language use in that speaker's community, and individual cognitive factors may play a role in this process. The characterization of these factors has been a topic of much theorization, experimentation, and debate for L3 researchers for the last two decades (Hammarberg, 2012).

In an attempt to isolate key predictors of L3 initial state transfer, L3 researchers have proposed a variety of L3 acquisition models. The primary goal of the majority of these models is to predict the source language(s) of transfer at the initial stages of the L3 acquisition. García Mayo and Rothman (2012) have organized these models into two general categories: wholesale transfer models and piecemeal transfer models.

As their name suggests, wholesale transfer models predict that one language is selected in its entirety to serve as the exclusive source language at the initial stages of L3 acquisition. The non-selected language, meanwhile, is unable to contribute features to the L3 initial grammar, even if those features would be beneficial. Four models of L3 acquisition are typically categorized as wholesale models: the L2 Status Factor Model (L2S: Bardel and Falk, 2007; Falk and Bardel, 2011; Falk, 2017; Falk and Lindqvist, 2019), the L1 Privilege Model (L1P: Hermas, 2010, 2014a, 2014b), the Typological Primacy Model (TPM: Rothman, 2011, 2013, 2015), and the Principal Language of Communication Model (PLCM: Fallah, Jabbari, and Fazilatfar, 2016; Fallah and Jabbari, 2018).

The claims and predictions of individual wholesale transfer models differ in regards to which factors play a role in determining whether the L1 or the L2 serves as the source language. Both L1P and L2S claim that order of language acquisition serves as the primary determiner for the selection of the language for initial state transfer. As its name suggests, L1P predicts that the L1 serves as the source of transfer in a manner similar

to that of L2 acquisition (Hermas, 2010, 2014a, 2014b). L2S meanwhile assumes that languages developed after the so-called "Critical Period" (CP) are neurologically different from those acquired in childhood, as claimed by Paradis (1994). As a result of the L2 and L3 being neurologically more similar, the L2 will serve as the transfer source at the L3 initial stages (Bardel and Falk, 2012). TPM meanwhile claims that the L3 learner transfers the previously known language that is the most typologically similar to the new L3, as determined first by lexical similarity, followed by phonological similarity, followed by morphology and then syntax (Rothman, 2011, 2013, 2015). It should be noted that TPM makes predictions based exclusively on actual linguistic typological similarity, and explicitly states that perceived similarity does not play a role in this process. However, Nelson, Krzysik, Lewandowska, and Wrembel (2021) have shown that psycho-typological similarity could also serve as a factor in L3 CLI and/or transfer. Finally, PLCM claims that learners transfer the language that they are most reliant on in their everyday life. In other words, the learner's primary language of communication serves as the source language for L3 transfer.

While each of these four wholesale transfer models makes different, and in many cases conflicting, predictions about individual cases of initial L3 transfer within a particular learner, L1P, L2S, TPM, and PLCM all share the same core claim that one entire previously known grammar is transferred at the initial stages of L3 development, while grammatical information from the non-transferred language is not accessible to the learner at the initial stages of language learning.

Piecemeal transfer models, meanwhile, predict that all previously known languages are available to the learner to some degree for the purposes of L3 transfer, and that rather than having one entire previously known language grammar transfer all at once at the onset of acquisition, transfer happens on a "piecemeal" basis, with features transferring as needed throughout the early stages of the acquisition process. The Cumulative Enhancement Model (CEM: Flynn, Foley, and Vinnitskaya, 2004; Berkes and Flynn, 2012) was the first model to make such a proposal, and holds that all previously known languages have the ability to potentially influence and enhance the development of subsequent language acquisition. When considering CEM, however, it is important to note that despite some summaries of CEM which have been made by other researchers (e.g. Rothman, 2010, 2011; Slabakova, 2017) CEM makes no direct claims regarding the transfer of unhelpful grammatical features, such as the case for Rothman and Cabrelli Amaro (2010), where the unhelpful feature of L2 Spanish null subject was found to transfer to L3 French.

It is on this topic of unhelpful or non-facilitative feature transfer where the Linguistic Proximity Model (LPM: Westergaard, Mitrofanova, Mykhaylyk, and Rodina, 2017) looks to further develop the piecemeal transfer ideas of CEM. Westergaard et al. and her colleagues propose that language acquisition is cumulative, and that learners have access to all previously acquired linguistic knowledge, with no part of that previous linguistic competence being blocked at any stage of the acquisition process similar to the claims of CEM. They differ from CEM's claim in proposing that in addition to the facilitative transfer which occurs based on structural similarity, similar to TPM, non-facilitative influence is also explicitly predicted as a result of learners incorrectly assuming that a linguistic property is shared between a previously acquired language and the L3. This non-facilitative transfer may occur as a result of a misanalysis of the L3 input, or the learner not having sufficient L3 input to make a more accurate assessment.

Structural similarity, however, may not be the only factor involved in property-by-property transfer. In her Scalpel Model (SM) of L3 acquisition, Slabakova (2017) proposes a set of claims which are very similar to LPM in two key ways. First, both models claim that acquisition happens on a property-by-property basis, rather than wholesale transfer. Additionally, both models explicitly predict and attempt to explain non-facilitative transfer from previously known languages. However, Slabakova argues that in addition to the linguistic factors considered by LPM, cognitive and experiential factors can also influence the transfer of a particular feature. While Slabakova notes that her list of potential extra-linguistic influences is certainly not complete, she suggests processing complexity, misleading input, and construction frequency as three of the extra-linguistic variables which have been found to influence L3 acquisition success (Slabakova, 2017). It should be noted that the influence of certain factors such as order of acquisition and social context, which have previously been emphasized in certain wholesale transfer models (Bardel and Falk, 2007; Fallah et al., 2016) could also be accounted for in SM as extra-linguistic influences.

While wholesale vs. piecemeal distinction listed above is often considered a key factor in differentiating between the many models of L3 acquisition (Puig-Mayenco, González Alonso, & Rothman, 2018), another key factor that differentiates between models of L3 is whether their predictions relate exclusively to the initial stages of L3 acquisition, or continue through later L3 development. As originally designed, several of the models listed above, such as TPM (Rothman, 2010), only make explicit predictions regarding true ab initio L3 transfer, and do not make any clear predictions related to CLI in later stages of L3 development. This is the case even for post-initial stage "beginner" learners with low proficiency in the L3 (Puig-Mayenco & Rothman, 2020).

This newly arising distinction between initial transfer and CLI in later stages of L3 development raises several concerns related to previous interpretations of findings in L3 research and implications for L3 models. For example, Ghezlou, Koosha, and Reza Lotfi's (2018) analysis of L3 English adjective placement claims to support the predictions of the L2S. However, the experiment examined participants with L3 English proficiency levels ranging from beginner to advanced. This raises the question of whether the CLI observed in this study was true initial transfer, as predicted by L2S, or some later form of CLI. While this distinction between initial transfer and later developmental CLI is incredibly difficult to tease apart in an experimental design context, it is critical when considering experimental findings and their implications for L3 acquisition models. In order to differentiate between research questions related to L3 initial state transfer, and later cases of post-initial state CLI, this paper will refer to research related to post-initial stages L3 acquisition as "L3 development".

Some L3 researchers have claimed that their L3 initial transfer models have implications in an L3 development context. For example, Slabakova (in press) has combined ideas from her Scalpel Model (Slabakova, 2017), the Bottleneck Hypothesis (Slabakova, 2019) and the theory of Full Transfer Potential (Westergaard, 2021a, 2021b) to propose that there is a single language mechanism for all monolingual, bilingual, and multilingual individuals, in which language differentiation uses a pathway-internal mechanism of raising and lowering the activation of words, features and structures that are currently participating in the linguistic computation. While this proposal has merit in larger theoretical discussions on the underlying neurolinguistic processing

in multilinguals, it does not make any clear, testable predictions related to specific outcomes of particular L3 learners.

Meanwhile, Cabrelli and Iverson (in press) have proposed a model which addresses a key question in L3 development that until this point had remained unexamined in the L3 literature. Assuming that non-facilitative grammatical information initially transfers to the L3, what factors might influence the process of overcoming that transfer effect? According to the Cumulative Input Threshold Hypothesis (CITH), CLI from the L2 is overcome faster than CLI from the L1, an idea which has a clear relation to the initial state predictions of the L2S model of initial state transfer and which most definitely warrants additional investigation.

Finally, Fernández-Berkes and Flynn (in press) have proposed and extension of Lust's (2012) "Grammatical Mapping" (GM) paradigm to apply to L3 acquisition. Fernández-Berkes and Flynn define their L3 GM paradigm as "the integration of linguistic constituents in the [specific language grammar] via a creative, structure-dependent acquisition process in the multilingual mind, and which is guided and constrained by Universal Grammar (UG) as well as informed by the learners' previous linguistic experience." In short, their model claims that L3 development is fundamentally guided by both UG and grammatical knowledge from previously acquired languages. GM claims that previously known languages are not prioritized with respect to other in their ability to enhance grammatical development.

Despite overall emphasis on initial and early stages of L3 acquisition in terms of formal models, there is a relatively sound level of evidence in the experimental literature for what factors may impact post-initial stages of L3 development, and it is not surprising that in many cases these are in fact the same factors which have been claimed to be key predictors of initial language transfer. For example, age of L2/L3 acquisition (Pfenninger, 2020), language dominance (Puig-Mayenco, Rothman, & Tubau, 2019), and proficiency in all known languages (Sciutti, 2020), have all been found to play a role in CLI in later stages of L3 development.

Overall, current models of L3 acquisition and development fail to make clear and consistent predictions about transfer of previously known linguistic features across the early stages of L3 learning. L3 initial stage models make no clear predictions regarding later stages of L3 development, and while current L3 development research shows promise in terms of revealing the linguistic and extralinguistic factors that may influence the accessibility of previously acquired grammatical information, this field of research is still young and additional research is greatly needed.

The proposed dissertation project examines transfer and CLI in L3 German learners throughout their first semester of L3 instruction, with a focus on these students' acquisition and use of the morpho-syntactic feature of grammatical gender. The project aims to capture both initial transfer of gender at the onset of L3 learning, if it occurs, as well as later points in the L3 development process, addressing the current debate regarding L3 wholesale vs. piecemeal initial-stage transfer (Puig-Mayenco et al., 2018), and how additional factors such as age of acquisition, proficiency in each language, and typological similarity influence later L3 development.

1.1 Grammatical Gender Acquisition

The term "gender", when used to refer to a linguistic phenomenon, can be separated into two distinct categories: lexical gender and grammatical gender. The former is a semantically motivated gender agreement which is dependent on the sex or gender of an animate referent, such as in the use of the pronouns "he" and "she" in English. The latter refers to a system of sorting of nouns into two or more arbitrary classes, categories, or groups, which do not necessarily correlate with biological gender in any way (Kramer, 2015; Ayoun, 2022). Within a grammatical gender system, there are two distinct processes that must take place: gender assignment and gender concord. Gender assignment is, as the name suggests, the assignment of a particular noun to an arbitrary category such as "masculine", "feminine", or "neuter". This category may or may not align with the biological gender of the noun it is referring to. Gender concord is the process of morpho-syntactic agreement between the gender of a given noun and other parts of the phrase such as adjectives and determiners. Grammatical gender systems are present in German, which has a three-way masculine/feminine/neuter category system, as well as in Romance languages such as Spanish or French, which have a two-way masculine/feminine category system. These gender systems of German (1), Spanish (2), and French (3) are exemplified below. It should be noted that in the case of French grammatical gender, not all grammatical gender markers present on adjectives in the written form of the language are pronounced in the spoken form. English does not have a nominal grammatical gender system.

(1) **Ein** schön-**er** Schlüssel
DET_{Masc} beautiful-$_{\text{Masc}}$ key$_{\text{Masc}}$
'a beautiful key' (German)

(2) Un-**a** clave roj-**a**
DET-$_{\text{Fem}}$ key$_{\text{Fem}}$ beautiful-$_{\text{Fem}}$
'a beautiful key' (Spanish)

(3) Un-**e** **belle** clé
DET-$_{\text{Fem}}$ beautiful$_{\text{Fem}}$ key$_{\text{Fem}}$
'a beautiful key' (French)

While children acquiring a grammatically gendered language such as Spanish or French have typically developed both gender assignment and gender concord by the age of 3 (Socarrás, 2011), the process of learning a grammatical gender system in adulthood has been shown to be much more difficult. Researchers such as (Sabourin, 2001), White, Valenzuela, Kozlowka-Macgregor, and Leung (2004) and Ahn (2015) have found evidence that knowledge of grammatical gender in the L1 can aid in the acquisition of gender in the L2. This is the case even in instances where agreement in the L2 is more complex than in the L1. For example, Ellis, Conradie, and Huddlestone (2012) found that L2 learners of German with L1 Italian produced more agreement on determiners and adjectives in their L2 German speech than those with (non-gendered) English or Afrikaans as their L1.

In cases where grammatical gender is present in the L2 but not in the L1, however, research has resulted in more mixed findings. For example, researchers have found that L2 learners can achieve scores comparable to those of native speakers on offline measures of grammatical gender ability such as written gender use or grammaticality judgments (Oliphant, 1998; Granfeldt, 2000; White et al., 2004). However, other experiments using more online cognitive measures of grammatical gender processing such as eye-tracking (i.e., Grüter,

Lew-Williams, and Fernald, 2012; Spino, 2022), ERP (i.e., Alemán Bañón, Fiorentino, and Gabriele, 2012; Foucart and Frenck-Mestre, 2010; von Grebmer zu Wolfsthurn, Pablos Robles, and Schiller, 2021)and fMRI (i.e., Miceli et al., 2002) studies show significant differences between native and non-native grammatical gender use and processing, leading to a debate in the literature as to the degree to which the gender system of highly proficient L2 learners can ever be truly native-like. For the purposes of this study, the question of interest is not whether the grammatical gender developed by L2 learners is truly native-like, but rather to what degree the L1 vs. L2 systems are capable of transferring to the L3. Given the findings from the bilingualism literature that (a) L1 grammatical gender knowledge influences L2 learning and (b) advanced L2 speakers whose L1 lacks grammatical gender are able to develop some sort of grammatical gender system, the question remains as to how the native vs. non-native gender within an L3 learner will influence L3 grammatical gender acquisition and development over the first year of L3 instruction.

1.2 Previous Work on L3 Grammatical Gender

Previous work by this author examined the role of two key factors on L3 German grammatical gender acquisition by beginner German learners: order of language acquisition, and typological similarity. The trilingual participants examined were beginner German learners with L1 English, L2 Spanish or L1 Spanish, L2 English (Brown, 2018, 2020a). The L1/L2 pair of Spanish and English was used due to a dual disparity in (a) typological similarity to the target L3, German, and (b) occurrence of the target feature: whereas English is typologically more similar to German (Eberhard, Simons, & Fennig, 2022), only Spanish has a gender system. As demonstrated in Figure 1, the key finding from this study was that participants who had a background of L1 English/L2 Spanish/L3 German were significantly better at identifying German grammatical gender errors compared to learners with L1 Spanish/L2 English/L3 German or L1 English/L2 German control participants.

One potential interpretation of these results is that they fall in line with the predictions of L2S, and that the higher scores from the L2 Spanish speakers is the result of wholesale initial transfer of L2 Spanish. However, beginner participants in this work had all completed up to one year (two semesters) of German instruction before participating in the experiment. As a result, it remains unclear from the data whether the findings were the result of initial transfer of the L2, or some early phase in L3 interlanguage development. This possibility raises questions regarding the point when this difference based on order of acquisition that was found in this study emerges. Do learners with an L2 Spanish background have this advantage from the very onset of L3 German acquisition or does it emerge at some later point in the learning process? In order to examine this question, Brown (2020b) conducted a follow-up experiment attempting to capture the true initial state of L3 German acquisition.

This author's follow-up project examined true ab initio German learners, bilinguals with L1 English/L2 Spanish or L1 Spanish/L2 English, and L1 English/L2 Mandarin controls with no previous grammatical gender knowledge, all of whom had no previous German experience. Participants were taught a simplified version of the German grammatical gender system. After being trained on the lexical items, participants completed a forced choice task where they were asked to differentiate between grammatically correct spoken
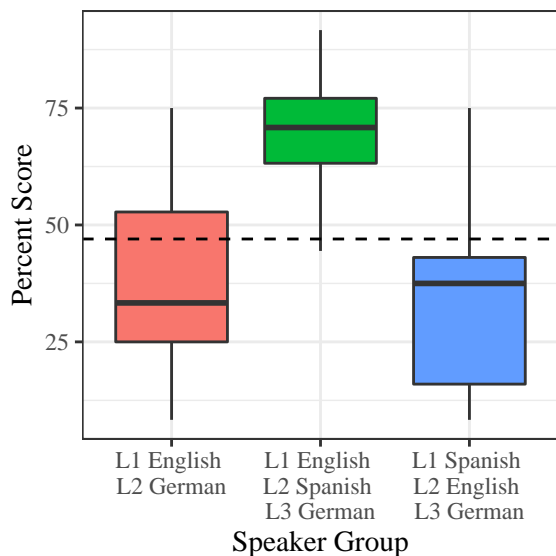
Figure 1: Results from Brown (2020a): L3 German **beginner** learners with L1 English/L2 Spanish show a **significant advantage** in identifying L3 German grammatical gender errors
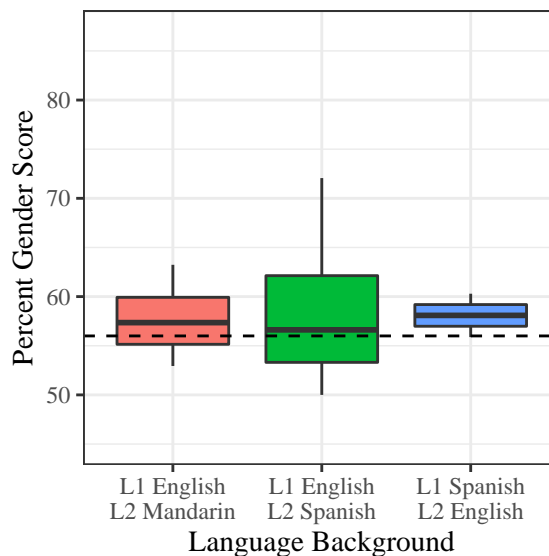

Figure 2: Results from Brown (2020b): L3 German **ab initio** learners with L1 English/L2 Spanish show no significant advantage in identifying L3 German grammatical gender errors

sentences, and those containing a grammatical gender error. Results of this experiment differed from those of this author's initial study in that no significant difference was found between the three language background groups in their ability to identify grammatical gender errors. Instead, overall all participants struggled with both the grammatical gender agreement questions, as shown in Figure 2.

Taken together, the author's previous works suggest that in the case of Spanish/English bilinguals acquiring L3 German, all participants begin the L3 acquisition process with roughly the same degree of competence in German grammatical gender. However, at some point during the first year of German acquisition, learners with L1 English/L2 Spanish/L3 German gain a significant advantage in the German grammatical gender learning process and begin to outperform learners with L1 Spanish/L2 English?L3 German and L1 English/L2 German on grammatical gender-related tasks.

It should be noted that this author's previous findings are based on particularly small samples of participants. This was unavoidable at the time due to difficulty finding participants with the target language background, limited resources for participant compensation, and, in the case of the second experiment, interruptions in data collection as a result of the COVID-19 pandemic. However, it should be noted that the findings of this author's previous work fall in line with other recent work on L3 grammatical gender acquisition. Jaensch (2012) examined L3 German gender assignment in learners with either L1 Spanish L2 English or L1 Japanese L2 English and found no advantage of L1 Spanish gender knowledge in L3 German gender assignment success. Instead, L2 English proficiency was found to be the best predictor in German gender assignment success. Długosz's (2021) examination of L1 Polish, L2 Norwegian L3 Swedish, all three of which have grammatical gender systems, found that in offline tasks similar to the ones used in this author's first

study, participants were found to exclusively transfer their L2 Norwegian grammatical gender knowledge over their L1 Polish. When considering online gender decision tasks, however, Długosz found evidence of competition between the Norwegian and Swedish gender systems at the level of gender retrieval. Finally, Ecke's (2022) examination of L1 Spanish, L2 English, L3 German speakers did in fact find evidence of CLI from L1 Spanish onto L3 German. However, this CLI from Spanish was overall non-facilitative transfer, as learners were found to reduce the complexity of the article/gender selection process by adopting lexical equivalents and over-using the most frequent forms. CLI from Spanish did not aid these participants in developing any type of native-like sensitivity to phonological/morphological gender cues.

Taken together, these findings bring further attention to the questions raised by this author's previous work. How do factors such as L2 status, age of L2/L3 acquisition, L2/L3 proficiency, and linguistic similarity influence a learner's acquisition of the unique and complex property of grammatical gender? Current models of L3 acquisition, which focus primarily on initial transfer, generally do not make clear predictions related to later L3 developmental differences (González Alonso & Rothman, 2017). As a result, these current models are unable to clearly account for the the combined findings of this author's previous work, or any other previous L3 gender acquisition findings outlined here, as all of the L3 grammatical gender research discussed above focuses on post-initial state learners. These results bring to attention the need for more complete models of L3 development, which can clearly account for occurrences of CLI not only at the initial stages of L3 acquisition, but throughout the larger L3 development process.

The proposed dissertation project aims to address these questions and gaps in the literature by more precisely examining the point at which learners with different orders of acquisition begin to diverge in their grammatical gender acquisition patterns, as well as how factors such as length of L3 exposure and L3 proficiency play a role, and how those roles may change across early phases of L3 grammatical gender development.

2 Proposed Dissertation and Tentative Chapter Outline

2.1 Objectives and Rationale

Given the findings of the author's previous work on L3 grammatical gender, outlined in §1.2, the objective of the proposed dissertation is to characterize the development of grammatical gender in adult L3 German learners, with particular emphasis on the role of transfer/CLI from previously known languages throughout the L3 acquisition process. The dissertation will address three key research questions in this area:

- Does previous grammatical gender knowledge from the L1 and/or L2 transfer at the initial stages of L3 German acquisition?
- Can previous grammatical gender knowledge influence later stages of L3 German development through cross-linguistic influence?
- What additional factors may promote or inhibit the transfer of previous grammatical gender knowledge to the L3? Factors under consideration in this study include age of L2 acquisition, proficiency in L2 and/or L3, and/or frequency of use, and (psycho-)typological similarity of the languages.

The proposed dissertation aims to address these questions by conducting two experiments of L3 German grammatical gender acquisition and development: one cross-sectional study, outlined in §3, and one longitudinal study, outlined in §4.

2.2   Outline of Proposed Chapters

The proposed format of the dissertation is as follows:

1. Introduction
   - This section will outline the significance of the study and its place in the literature, as well as provide an outline of the dissertation
2. General Background: L3 acquisition and development
   - This section will provide a context for the research being conducted in this dissertation, with an emphasis on outlining the history and state of the field of L3 acquisition in regards to L3 initial state transfer, post-initial state L3 development, and the linguistic, cognitive, and socio-cultural factors that may impact these processes.
3. General Background: Grammatical gender
   - This section will provide a review of the literature related to grammatical gender, including morpho-syntactic theories of grammatical gender, neuro-linguistic research, and research on native and non-native gender acquisition. This section will primarily focus on research related to the gender systems of German, Spanish, French, and Italian. It will also briefly discuss how these grammatical gender systems compare the gender systems of other more typologically distant languages.
4. Experiment 1: Cross-Sectional Study
   - See §3 for more details.
5. Experiment 2: Longitudinal Study
   - See §4 for more details.
6. Discussion
   6.1 Experiment 1 and 2 findings, summary and implications when considered together
   6.2 Findings in relation to current L3 acquisition models
   6.3 Implications for L3 acquisition theory and pedagogy
7. Conclusion
   7.1 Main Findings
   7.2 Future Directions

3   Experiment 1

3.1   Experiment 1 Description

The primary goal of Experiment 1 is to replicate this author's original findings in a modified version of the same experiment (Brown, 2020a). More specifically, the main task of this study will be a grammaticality

judgment task, modeled after the task from the previous work, but the major updates to the protocol will be in the improvements to the language background questionnaire and the inclusion of participants with knowledge of other Romance language knowledge besides Spanish. The original project consisted of a brief set of language background questions targeted at categorizing participants into one of the four target language background categories (L1 English/L2 Spanish/L3 German, L1 Spanish/L2 English/L3 German, L1 English/L2 German or Native German). The language background questionnaire used in the original study will be replaced by the more standardized trilingual Language Experience and Proficiency Questionnaire (LEAP-Q: Marian, Blumenfeld, and Kaushanskaya, 2007. This will allow for a more thorough analysis of factors such as age of acquisition, language dominance, and perceived language similarity.

In addition to this major update, this updated project will differ from the original study in the following ways. First, the trilingual participant pool will include not only participants with knowledge of Spanish, but also French and/or Italian, both of which also have grammatical gender. The original project exclusively targeted English/Spanish/German trilinguals and English/German bilinguals, and as a result was forced to omit an unexpectedly high number of participants with knowledge of additional Romance languages. Given that participants for this project are incredibly difficult to find, the inclusion and full analysis of additional languages could greatly widen the participant pool, as well as provide a broader insight into the transfer of grammatical gender in this context. French and Italian were included specifically because of the availability of the LexTALE vocabulary test in those languages (Brysbaert, 2013; Amenta, Badan, and Brysbaert, 2020), as well as their familiarity to this author.

Second, this updated study will include participants with intermediate and advanced levels of German proficiency in addition to the early/beginner learners targeted in the previous study. German proficiency, as measured by a LexTALE vocabulary test (Lemhöfer, 2012), length of formal instruction, and reported German use will be considered as variables to predict variation in grammatical gender development and use at later stages of the German acquisition process.

Finally, in terms of the grammaticality judgment task itself, the overall number of questions will be lowered in order to prevent participant drop-out.

3.2   Procedure

3.2.1   Experimental Tasks

The procedure consists of an online, survey-style task performed independently by participants on their own computers. This task will be created using the Qualtrics online survey tool (Qualtrics, 2021) . In the task, participants will be presented with a total of 30 experimental questions and 30 control/distractor questions in a randomized order. All experimental sentences will involve a singular noun in the nominative case, along with its definite article. Within the 30 experimental questions, 15 will be completely grammatically correct. The other 15 experimental questions will consist of a grammatical gender mismatch between the noun and the article. An example of a correct and an incorrect sentence are given in (4) and (5), respectively, where (5) demonstrates a grammatical gender mismatch between a neuter definite determiner and a masculine noun.

Additionally, all sentences will be in the present tense, and all nouns used will be taken from a variety of introductory-level German textbooks. Control/distractor sentences will either be grammatically correct or will consist of grammatical errors unrelated to gender, such as number or case errors.

(4) Die   Krawatte ist grün
DET$_{FEM}$ tie$_{FEM}$ is green
'The tie is green'

(5) ***Das** **Baum** ist groß
DET$_{\mathbf{Neu}}$ key$_{\mathbf{Masc}}$ is tall
'The tree is tall'

Within the set of experimental questions, types of grammatical errors will be evenly distributed. For example, there will be the same number of masculine nouns incorrectly marked with a feminine determiner as masculine nouns incorrectly marked with a neuter determiner. Additionally, sentences will be maximally balanced with reference to the similarity between the grammatical gender of the subject noun of the German sentence and the grammatical gender of the corresponding translation noun in the Romance languages under consideration. For example, the word "skirt" is feminine is Spanish (la falda) and masculine in German (der Rock). So in the analysis of an experimental question containing the noun "Rock" it would be noted that that particular noun has an opposite gender in the participants' other gendered language, Spanish.

After completing the grammaticality judgment task, participants will be asked to complete a test of German grammatical gender. Participants will be presented with a bare noun in German and asked to assign a grammatically gendered determiner to each noun. All German nouns included in the study will be tested in this section. This will allow the author to differentiate between errors resulting from a participant's lack of knowledge of a particular noun or its gender assignment, and errors that are truly the result of a participant not having command of the larger German grammatical gender system.

3.2.2 Participants

Participants in this study will fall into one of three German proficiency groups: Beginner, Intermediate or Advanced; as well as into one of four language background categories: L1 English/L2 Romance/L3 German, L1 Romance/L2 English/L3 German, L1 English L2 German, and L1 German/L2 English. The project aims to have a total of 15 participants in each combination of proficiency level and language background (subgroup), for a total of 150 participants, as outlined in Table 1.

German proficiency for this categorization will be based on participants' score on the German LexTALE vocabulary test. Language background group will be determined based participant's reported age of acquisition of each of their known languages provided in the LEAP-Q. All bilingual and trilingual participants must have acquired their L2 after the age of 8 qualifying them as late child bilinguals or late/adult bilinguals based on Montrul's (2002) categorization.

It should be noted that while the target number of participants in this project is large, the procedure of this study involves participants simply completing a survey-style questionnaire on their own computer. The output of the questionnaire has been designed to be fast and easy to grade, and the only major tasks related to data collection are to advertise the study and provide the link to as many potential language programs as

	German Proficiency			
	Beginner	**Intermediate**	**Advanced/Native**	**Total**
L1 English/L2 Romance/L3 German	15	15	15	45
L1 Romance/L2 English/L3 German	15	15	15	45
L1 English/L2 German	15	15	15	45
L1 German/L2 English	N/A	N/A	15	15
Total	45	45	60	**150**

Table 1: Participant numbers by language background and German proficiency.

possible. Additionally, due to the brief nature of the study, participants will not be paid for their participation in this experiment, resulting in no monetary limitations on participant numbers. For this reason, this author believes that the participant numbers proposed here are ambitious, but possible.

3.3 Planned Analysis

Statistical analysis in this project will examine the dependent variable of response accuracy (i.e. participant's ability to correctly identify a grammatical gender error). Mixed model logistic regression will be used to examine how the factors of (1) age of L2 acquisition, (2) L2 proficiency, and (3) perceived language similarity as reported in the LEAP-Q, among other factors may influence overall success in L3 grammatical gender acquisition.

3.4 Potential Results

The first key result anticipated in this experiment is whether the findings from the beginner participants in this study mirror those from this author's original study (Brown, 2020a), with L1 English/L2 Romance/L3 German participants showing an advantage in identifying German grammatical gender errors compared to the participants with L1 Romance/L2 English/L3 German or L1 English/L2 German. If it is the case that the results of the previous work can be replicated, then the second key result will be whether the participants with L1 Romance/L2 English/L3 German show evidence of grammatical gender transfer or CLI at later stages of L3 German development. If it is the case that the L1 English/L2 Romance/L3 German participants show a significant advantage in identifying grammatical gender errors even at the lowest levels of L3 German proficiency, while the L1 Romance/L2 English/L3 German participants perform on par with the L1 English/L2 German controls, then this would be evidence for some form of wholesale L2 transfer at the onset of L3 acquisition, coupled with some form of inaccessibility of the L1 grammar for transfer even at later stages in L3 development. An alternative possible finding could be that the beginner level participants perform in line with the results of the previous work, but that at the intermediate proficiency level L1 Romance/L2 English/L3 German participants begin to show a similar advantage similar to that of the L1 English/L2 Romance/L3 German participants, suggesting that some form of CLI or transfer from the L1 has occurred at a later point in the L3 development process, which could fall in line with more piecemeal models of L3 acquisition as well as models of L3 development. When results are available, the resulting dissertation will include a full analysis of the results' implications for all other listed L3 acquisition models.

4 Experiment 2

4.1 Experiment 2 Description

While Experiment 1 aims to replicate this author's previous findings, Experiment 2 is designed to more deeply examine the L3 gender acquisition process within an individual learner, and the way that the role of previously known languages may change throughout the acquisition process. This will be done by using a longitudinal approach, tracking the same learners throughout their first semester of German instruction. By comparing results from the same participants at different points in the acquisition process, this method is able to control for certain confounding individual cognitive variables known to impact adult grammatical gender acquisition such as real-time language processing speed (Grüter et al., 2012), declarative memory and procedural memory (Kempe, Brooks, & Kharkhurin, 2010). Additionally, Experiment 2 will include a more comprehensive battery of experimental measures including response times and eye-tracking, which will allow for the examination of differences in CLI/transfer in regards to online vs. offline gender processing, a factor which has recently been shown to play a role in L3 gender specifically (Ecke, 2022).

It should be noted that this experiment is funded by *Language Learning*'s Dissertation Grant, as well as the Boston University Linguistics Department's Graduate Research Grant. In addition to this, an application has been submitted for the National Science Foundation's Linguistics Dissertation Improvement Grant. If received, this grant would be used towards the compensation of additional participants in the study, as well as for the inclusion of an undergraduate research assistant (RA) in the project to aid in data collection and processing. §4.2.2 presents the target participant numbers for this experiment both in the case where the author does not relieve the NSF grant and in the case where this study does receive this funding.

4.2 Procedure

4.2.1 Experimental Tasks

Participation in the experiment will take place via a virtual video conference meeting with the researcher or the RA, and participation will consist of three separate meetings over the course of the semester. The first meeting will take place within the first two weeks of the participant's beginning instruction in German. The second meeting will take place at the midpoint of the same semester of instruction, so that the participant has received roughly 6-8 weeks of German instruction. The third meeting will take place at the end of the semester, totaling roughly 15 weeks of German instruction.

Each individual meeting will last between an hour and an hour and a half, and will consist of three experimental tasks: a forced choice listening task, a self-paced reading task, and a LexTALE vocabulary test. Participants will complete all three of these tasks in German and, in the L3 learner participants, will also complete these three in their previously known Romance language (Spanishm French, or Italian). This is done in order to assess L1/L2 proficiency and ensure that L3 learner participants have fully acquired a Romance grammatical gender system in their L1 or L2. Participants will also complete the LexTALE vocabulary task in English as a measure of English proficiency. The last step of each meeting will be the

Order	Language	Task
1		Forced Choice Listening Task
2	German	Self-Paced Reading Task
3		LexTALE Vocabulary Task
4		Forced Choice Listening Task
5	Romance	Self-Paced Reading Task
6		LexTALE Vocabulary Task
7	English	LexTALE Vocabulary Task
8		LEAP-Q

Table 2: Order of tasks in Experiment 2. This procedure will be the same at all three meetings with each participants. Bilingual control participants will not complete the tasks involving Romance languages.

completion of a modified version of the LEAP-Q (Marian et al., 2007). All participants will complete the tasks in German first, in order to minimize the potential for co-activation of other known languages, followed by the Romance language tasks, followed by the English LexTALE and LEAP-Q. Participants will complete all of these tasks in that order at all three meetings throughout the semester. Table 2 lists all tasks in the Experiment 2 protocol in the order that they will be presented to the participant.

The forced choice listening task will be the first task that participants complete in German and, in the case of the L3 learners, also in Spanish, French or Italian. In order to maximize activation of the target language, as well as to ensure that participants do in fact know the target lexical items, all participants will start the forced choice listening task with a short lexical training session. A total of 12 nouns will be used, which have been balanced for German grammatical gender as well as the grammatical gender of their Spanish, French and Italian translation equivalents. All nouns in this set were taken from introductory German textbooks and refer to inanimate, physical objects.

In the training section of the listening task, participants will be presented with two images, one on either side of their computer screen. After having a few seconds to look at the pictures, the participant will hear a sentence in the target language, such as the one shown in (7), and will be asked to choose the image that the sentence is referring to. Once they have responded, they will receive feedback in the form of a green check mark or red X. Each of the 12 nouns will be displayed at least twice.

(6) Listening Task Training Section example (no gender error)

Shauensie! Der   Schlüssel.
Look! DET_{Masc} key_{Masc}!

(7) Listening Task Training Section example (gender error on determiner)

Shauensie! **Die**   Schlüssel.
Look! $DET_{\textbf{Fem}}$ key_{Masc}!

After completing the training section, participants will complete the experimental section of the listening task. The experimental section will have a similar format as the training section with a few key differences. First, at the start of each question, participants will be presented with four images (rather than two as in the training task). Second, after responding to a question, participants will not receive feedback on their

response. In the key grammatical gender questions of this task, the four images will consist of the target image, another image of a noun with the same grammatical gender category as the target image, and two images of nouns with different grammatical gender categories. The sentence that participants listen to will either be completely grammatically correct in terms of grammatical gender agreement, or will contain a grammatical gender error on the determiner. It should also be noted that when this test is presented in French, only cases where grammatical gender agreement is articulated in the spoken language will be included.

If it is the case that the participant has successfully acquired grammatical gender in the target language, then responses should be as follows: in cases that do *not* contain a grammatical gender error, once the participant has heard the grammatical gender cues on the determiner, they will fixate on the images that represent nouns with the corresponding grammatical gender. This concentration is measurable using virtual eye tracking, where participants will fixate visually on data with the matching grammatical gender features. Additionally, participants who use grammatical gender cues will have faster response times. In the cases that *do* contain a grammatical gender error, participants are expected to show visual fixation on the non-target stimuli that match the (incorrect) grammatical gender marker, as well as longer response times as they process the mismatched data.

Eye tracking in particular will be a critical measure in this task, as previous eye-tracking research on grammatical gender has shown that for native speakers, gender plays a critical role in predictive online comprehension of the phrase as a whole (Cholewa, Neitzel, Bürsgens, & Günther, 2019). In other words, speakers use cues from the feature content of the articles and adjectives that precede the noun in order to predict what the noun will be. This task will examine the degree to which this study's trilingual participants are able to use these same predictive cues in each of their gendered languages, and whether that ability changes as the L3 develops. This task is anticipated to take about 10-15 minutes per language.

It should be noted that while this author does not have previous experience using virtual eye-tracking software, she does have previous experience running in-person eye-tracking experiments and using eye-tracking software from her time working on Dr. Marnie Reed's project: "Differential focus-marking in L2 processing, perception, and production of 2017 English intonational contrast: An eye-tracking study' in 2017. She has since taken the time over the last semester to familiarize herself with the Gorilla Experiment Builder and its eye tracking software (Gorilla, n.d.) in preparation for this experiment.

In terms of the Gorilla Experiment Builder and online data collection context, examination of eye-tracking data collected in an online context has shown that while this methodology entails higher variance, a lower sampling rate (Gagné & Franzen, 2021), and increased experimental time, it shows no significant differences in spatial accuracy compared to in-lab recordings for adult data (Semmelmann & Weigelt, 2018). In terms of the Gorilla software specifically, Gorilla's (Anwyl-Irvine, Massonnié, Flitton, Kirkham, & Evershed, 2020) software includes an implementation of the WebGazer.js library (Papoutsaki et al., 2016) which functions as the eye-tracking software for the experiment builder. While to this author's knowledge no previous work has used this software in the context of a grammatical gender experiment, Papoutsaki, Laskey, and Huang

(2017) conducted a replication study of two previous eye-tracking experiments in psychology an found that the WebGazer software was able to recreate the general results and trends of both previous works. In addition, other researchers in linguistics and psychology have successfully taken advantages of this virtual eye tracking software in their research (e.g. Calabrich, Oppenheim, and Jones, 2021; Tufft and Gobel, 2021; Wessel, 2021).

After completing the forced choice listening task, participants will complete the self-paced reading task in the target language. In this task, participants will be asked to read a sentence on their computer screen and then answer a question based on that sentence. However, the sentence will be presented to the participant one word at a time. In order to see the next word in the sentence, participants will need to press the space bar. Typically, in this paradigm, cases involving grammatical errors cause the participant to pause, either at that word or the word following it, leading to a longer response time at that point (Aaronson and Scarborough, 1976; Mitchell and Green, 1978).

All target sentences will consist of a precursor phrase, followed by a noun phrase consisting of a grammatical gender mismatch on the determiner, as exemplified in (9). So, if it is the case that a participant in this study recognizes a grammatical gender error in the stimulus, they will show a longer response time after revealing the noun that exposes the gender mismatch. Additionally, the task will consist of a control condition with grammatical versions of the same sentences, which allow us to set a baseline for response time. This task will use the same set of beginner-level nouns, balanced for grammatical gender that were used in the forced choice listening task. This task is anticipated to take about 10-15 minutes per language.

(8) Reading Task Example Sentence (no gender error)

Shauensie! Der   Schlüssel ist sehr groSS!
Look!   DET$_{\text{Masc}}$ key$_{\text{Masc}}$ is   very large!

(9) Reading Task Example Sentence (gender error on determiner)

Shauensie! **Die**   Schlüssel ist sehr groSS!
Look!   DET$_{\textbf{Fem}}$ key$_{\text{Masc}}$ is   very large!

Next, participants will complete a brief LexTALE (Lemhöfer and Broersma, 2012; Brysbaert, 2013; Izura, Cuetos, and Brysbaert, 2014) visual lexical decision task in the target language, which will be used as a measure of language proficiency. All participants considered in this study will test at a B1 (intermediate) level or higher on all LexTALE tasks.

After completing all tasks in all languages, the final task in each meeting of the experiment will be the modified LEAP-Q (Marian et al., 2007), which will briefly ask participants to describe their self-assessment of their proficiency in each of their languages as well as other language background information such as their frequency of use of each language, age of acquisition, and their perception of how similarity their languages are to each other.

Ideally, the set of tasks in each language would be done on three separate days to avoid co-activation of the participants' other languages during the data collection process. However, given that this experiment already

asks participants to return to the (virtual) lab on three different occasions in order to track development over the course of a year, to then separate each data collection point into three separate meetings would lead to a total of 9 visits per participant. Not only would this lead to significantly more complex scheduling requirements on the part of the researcher, but it also significantly increases the chances that participants will drop out of the study. This is the rationale for the researcher's decision to collect the data for all three of a participant's languages in one meeting, despite the risk of language co-activation.

4.2.2 Participants

Participants for this study will be recruited from introductory German classes and universities across the United States and Canada during the Fall 2022 academic semester. All data will be collected by early December of 2022. The population under consideration in this study consists of two different groups of adult German learners: L3 learner target participants and L2 learner controls, none of whom will have participated in Experiment 1. The L3 learner participants will have at least "intermediate" knowledge of both English and a Romance language as determined by LexTALE vocabulary test scores. The Romance languages included in this study are Spanish, French, and Italian. Participants with knowledge of other Romance languages such as Portuguese will be excluded from the study simply because there is not an equivalent LexTALE vocabulary test currently available in other Romance languages that could be used as an equivalent test of proficiency. The L2 learner control participants will be native English speakers learning German as an L2, with know previous knowledge of any other languages.

The L3 learner participants will then be further split into two categories: L1 English/L2 Romance and L1 Romance/L2 English, and all participants in either group will have acquired their L2 after the age of 4, so as to include early and late sequential bilingual individuals but exclude simultaneous bilinguals (Montrul, 2002).

Given the overall length of each experiment meeting, coupled with the fact that participants are required to return and repeat the experiment for a total of three experimental sessions, participants in Experiment 2 will receive monetary compensation for their time in the form of an Amazon gift card. Participants will receive a $25 gift card after completing their first meeting of the experiment at the beginning of the semester, a second $25 gift card after completing the second meeting halfway through the semester, and a $50 gift card after completing the third and final meeting of the experiment.

As noted in §4.1, funding for participant payment has been provided by the journal *Language Learning* and the Boston University Linguistics Department. Additionally, this author has applied for a National Science Foundation Dissertation Improvement Grant, which if received would more than double the number of participants included in the study. Results of the grant application are expected by June 2022. Given that overall funding amounts for this project remain unclear at this time, three different plans related to participant numbers and compensation have been have been proposed below. This ensures that a sufficient number of participants can be included in the study regardless of funding situation. All scenarios also account for the cost of virtual eye tracking software and allowance for some degree of participant drop-out.

Participant Numbers Plan A

In the case that this project is granted the requested NSF funding, the three grants together will cover the cost of compensation for 48 participants (16 in each of the three language groups) completing all three meetings of the experiment. A power analysis assuming a "medium" effect size of .29 (based on the effect size of Brown (2020a)) calculated a power level of 0.81 for this number of participants.

Participant Numbers Plan B

If it is the case that this project is not granted the requested NSF funding, the BU Departmental funding and *Language Learning* grants alone can cover the cost of 21 participants (7 in each of the three language groups) completing all three meetings of the experiment. A power analysis assuming the same medium (.29) effect size and a power level of 0.8 found that a minimum number of participants needed in this scenario would be 48. This author will apply for additional research funding grants from the BU Linguistics department in an attempt to cover this $1200 difference in participant payment funding.

Participant Numbers Plan C

In the case that no additional funding for this project can be acquired, then this issue of participant numbers and compensation will be resolved by eliminating the mid-semester point of data collection in the procedure for all participants. This would mean that each participant completes the experiment at two time points instead of three. Additionally, participant compensation would be cut so that participants receive $25 for their first meeting and $25 for their second meeting for a total of $50 per participant. A power analysis in this scenario, assuming the same effect size and a power level of 0.8 found that a minimum number of participants needed in this scenario would be 39. While these changes to the experimental protocol are not ideal, this number of participants at this compensation rate is affordable using exclusively funds that are currently available to the researcher.

4.3 Planned Analysis

In terms of eye-tracking analysis, the key variable under consideration is fixation duration. Fixations are defined as eye movements that stabilize the retina over a stationary object of interest (Duchowski, 2017). Fixation duration on each of the four presented images will be measured at three different points within each of the listening task questions described in §4.2.1. The first point will be when the four images are first presented to the participant, before the spoken sentences is provided. At this point, participants will have no grammatical cues as to what image they may be asked to select, and fixation should be relatively equally distributed between each of the four images. The second point will be once the participant has been presented with the definite article of the provided sentence. Previous grammatical gender research using eye-tracking suggests that if it is the case that the participant have developed some form of grammatical gender system in German, then as soon as the article is presented, the participant will favor the picture(s) that align with linguistic predictions based on gender processing (Cholewa et al., 2019). The third time point will be after the noun is presented, at which point the participant should fixate on the correct target

image. While raw fixation duration in data milliseconds will be collected by Gorilla, the key data point for the analysis will instead be the *ratio* of fixations between the four presented images, an approach which has precedent in the grammatical gender eye-tracking literature (e.g.Hopp, 2013). If it is the case that a particular participant has acquired grammatical gender in German, then at point 1, fixation duration should be roughly 25%/25%/25%/25%, meaning that participants are looking at each of the four images approximately the same amount of time. At point 2, fixation duration should be roughly 50%/50%/0%/0% meaning that participants are focusing only on the two images whose German grammatical gender assignment matches that of the provided determiner. At point 3, roughly 100% of fixations should be on the picture that matches the provided noun.

Statistical analysis in this project will examine three dependent variables: response accuracy (for both the listening and reading tasks), response time (for both listening and reading), and, in the case of the eye-tracking data, fixation duration on target and grammatical-gender distractor stimuli (listening task only). Participants who have successfully acquired grammatical gender will demonstrate high response accuracy, fast response times when provided with a grammatically correct statement, delayed response times when exposed to stimuli containing grammatical gender errors, and a long fixation duration on nouns of the relevant gender class once gender information has been revealed in a listening task sentence.

4.4 Potential Results

The key potential findings of this project relate to whether participants have developed these grammatical gender abilities at each point in the data collection process, and how the additional factors of age of acquisition, proficiency, and use of all known languages correlate with acquisition success. In particular, key conclusions of this study will relate to the predictions of current models of L3 acquisition. As emphasized in Section §1, the sheer number of L3 models proposed to date makes it impractical to spell out the exact predictions of each model in this circumstance. However, given the findings in the researcher's previous work, it is important to note the exact predictions of three models in particular: L2S, SM, and LPM.

The predictions of these models in this experiment are similar to those described in §3.3. If it is the case that participants with Romance as an L2 outperform all other participants in the grammatical gender tasks from the very onset of the experiment, including the very first point of data collection when participants have minimal German exposure and could be argued to be in the initial stages of L3 acquisition, then this would support L2S and its claims of wholesale transfer of the L2 at the initial stages of L3 acquisition. If it is the case that participants with L2 Romance outperform other participants only at the the second and/or third collection meetings, then this would support the SM and its claims that (1) transfer occurs in a piecemeal manner throughout L2 development (rather than simply at the onset), and (2) that factors such as age of L2 acquisition and/or L2/L3 proficiency could potentially result in an L2 Romance advantage similar to that found in the author's original study. This expected delay in transfer is based on SM's claim that L1/L2 features are transferred to the L3 when needed based on L3 input. It is unlikely that these early learners will have transferred their gender knowledge by the time of the first experimental meeting given the participant's minimal exposure German at that point in the study. Finally, if it is the case that all L3 learner participants

outperform the L2 German learner controls by the end of the study, but the two L3 learner groups do not differ from each other, this would fall in line with LPM. This is based on the fact that LPM makes a similar claim to SM that L1/L2 features are transferred in a piecemeal fashion to the L3 when needed based on L3 input. However, LPM does not specify age of acquisition, order of acquisition, or native/non-native status as factors that might influence whether or not a particular features transfers to the L3. Therefore, all L3 learners should have the ability to transfer their previous L3 gender knowledge from their Romance language, an advantage which is not available to the L2 control participants.

Additionally, the use of both online and offline measures in this project will have important implications for the field of L3. As it stands, the wide variety of often conflicting results across studies in this field can potentially be attributed to differences in experimental methodology. In particular, the difference between theoretical linguistic and psycholingusitic/cognitive approaches to L3 research warrant additional examination in this field (García-Mayo, 2012). If it is the case that participants in this study show significant differences in their online vs. offline grammatical gender processing, then this should emphasize the importance of considering such processing differences in model and future procedures for L3 research.

While at this point in experimental design and planning, the potential findings are somewhat broad, the implications of these findings are crucial to the field of L3 acquisition and development. The results of this study will help to create more accurate, inclusive, and predictive models of L3 acquisition and development.

5   General Plan of Work

The full plan of of work from the point of the defense of this Dissertation Prospectus through the filing of the dissertation is provided in Appendix A. Key points for an on-time dissertation defense and filing include (1) Experiment 1 data collection is finished by the end of August 2022, (2) Experiment 2 data collection is completed by December 2022, so that the spring semester of 2023 is reserved for analysis, writing, and the defense.

6   Potential Results and Broader Implications

While the theoretical implications of the two experiments are outlined in §3.4 and §4.4, it is important to note that this study also has key implications for foreign language pedagogy. This project brings to attention the unique learning experiences of multilingual learners, and their learning process differs from the L2 learning process. This could be used to help understand and design pedagogical approaches for this unique case of L3 learners.

References

Aaronson, D., & Scarborough, H. (1976). Performance theories for sentence coding: Some quantitative evidence. *Journal of Experimental Psychology: Human Perception and Performance*, *2*(1), 56–70. doi: 10.1037/0096-1523.2.1.56

Adémian, C. (1976). On the nature of interlanguage systems. *Language Learning*, *26*, 297-320. doi: 10.1111/j.1467-1770.1976.tb00279.x

Ahn, H. (2015). *Second language acquisition of Korean case by learners with different first languages* (Unpublished doctoral dissertation). University of Washington.

Alemán Bañón, J., Fiorentino, R., & Gabriele, A. (2012). The processing of number and gender agreement in Spanish: An event-related potential investigation of the effects of structural distance. *Brain Research*, *1456*, 49-69.

Amenta, S., Badan, L., & Brysbaert, M. (2020). LexITA: A quick and reliable assessment tool for Italian L2 receptive vocabulary size. *Applied linguistics*, *42*(2), 292–314. doi: 10.1093/applin/amaa020

Anwyl-Irvine, A., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*, 388-497. doi: 10.3758/s13428-019-01237-x

Ayoun, D. (2022). Introduction: Gender, *geslag*, *Geschlecht*, γ´ενος, ਲਿੰਗ, 性別. In D. Ayoun (Ed.), *The acquisition of gender: Crosslinguistic perspectives* (Vol. 63, p. 1-9). Amsterdam / Philadelphia: John Benjamins Publishing Company. doi: 10.1075/sibil.63.01ayo

Bardel, C., & Falk, Y. (2007). The role of the second language in third language acquisition: The case of Germanic syntax. *Second Language Acquisition*, *23*, 459-484. doi: 10.1177/0267658307080557

Bardel, C., & Falk, Y. (2012). The L2 status factor and the declarative/procedural distinction. In J. Cabrelli Amaro, S. Flynn, & J. Rothman (Eds.), *Third language acquisition in adulthood* (p. 61-78). Amsterdam: John Benjamins. doi: 10.1353/lan.2014.0011

Berkes, E., & Flynn, S. (2012). Further evidence in support of the Cumulative-Enhancement Model: CP structure development. In J. Cabrelli Amaro, S. F. Flynn, & J. Rothman (Eds.), *Third language acquisition in adulthood* (p. 143-164). Amsterdam: John Benjamins. doi: 10.1353/lan.2014.0011

Brown, M. M. (2018). *Theories of L3 morphosyntax in the acquisition of grammatical gender in L3 German by Spanish-English bilinguals* (Unpublished master's thesis). Boston University, Boston, MA.

Brown, M. M. (2020a). Grammatical gender acquisition in sequential trilinguals: Influence of a gendered L1 vs. L2. In P. Farrell (Ed.), *Proceedings of the Linguistic Society of America* (Vol. 5, pp. 331–344). doi: 10.3765/plsa.v5i1.4712

Brown, M. M. (2020b). *L3 acquisition of grammatical gender in an artificial language environment.* Boston, MA. Retrieved from https://static1.squarespace.com/static/5962a9593e00be4fb1e5bd99/t/5edfb77242dfe00537344fe7/1591719794990/Brown_QP1_Final.pdf (PhD Program Qualifying Paper)

Brysbaert, M. (2013). LexTALE_FR: A fast, free, and efficient test to measure language proficiency in French. *Psychologica Belgica*, *53*(23-37).

Cabrelli, J., & Iverson, M. (in press). Why do learners overcome non-facilitative transfer faster from an L2 than an L1? the Cumulative Input Threshold Hypothesis. *Journal of Multilingualism*.

Calabrich, S. L., Oppenheim, G. M., & Jones, M. W. (2021). Episodic memory cues in acquisition of novel

visual-phonological associations: a webcam-based eye-tracking study. In *Proceedings of the cognitive science society conference, 2021.*

Cholewa, J., Neitzel, I., Bürsgens, A., & Günther, T. (2019). Online-processing of grammatical gender in noun-phrase decoding: An eye-tracking study with monolingual german 3rd and 4th graders. *Frontiers in Psychology*, *10*, 2586. Retrieved from https://www.frontiersin.org/article/10.3389/fpsyg.2019.02586 doi: 10.3389/fpsyg.2019.02586

Corder, S. P. (1967). The significance of learner's errors. *International Review of Applied Linguistics in Language Teaching*, *5*(4), 161. doi: 10.1515/iral.1967.5.1-4.161

Długosz, K. (2021). Transfer effects in gender assignment and gender concord in L3 Swedish: A case study. *Moderna Språk*, *115*(4), 1-26.

Duchowski, A. T. (2017). *Eye tracking methodology: Theory and practice* (Third Edition ed.). Springer.

Eberhard, D. M., Simons, G. F., & Fennig, C. D. (Eds.). (2022). *Ethnologue: Languages of the world* (25th ed.). Dallas, Texas: SIL International.

Ecke, P. (2022). Grammatical gender and article use in beginning learners of German. In D. Ayoun (Ed.), *The acquisition of gender: Crosslinguistic perspectives* (Vol. 63, p. 127-156). Amsterdam / Philadelphia: John Benjamins Publishing Company. doi: 10.1075/sibil.63.06eck

Ellis, C., Conradie, S., & Huddlestone, K. (2012). The acquisition of grammatical gender in L2 German by learners with Afrikaans, English or Italian as their L1. *Stellenbosch Papers in Linguistics*, *41*, 17-27. doi: 10.5774/41-0-131

Falk, Y. (2017). On pronouns that drop (out of German). In T. Angelovska & A. Hahn (Eds.), *L3 syntactic transfer: Models, new developments and implications* (p. 127-142). Amsterdam: John Benjamins. doi: 10.1075/bpa.5

Falk, Y., & Bardel, C. (2011). Object pronouns in German L3 syntax: Evidence for the L2 status factor. *Second Language Research*, *27*, 59–82. doi: 10.1177/0267658310386647

Falk, Y., & Lindqvist, C. (2019). L1 and L2 role assignment in L3 learning. Is there a pattern? *International Journal of Multilingualism*, *16*(4), 411-424. doi: 10.1080/14790718.2018.1444044

Fallah, N., & Jabbari, A. A. (2018). L3 acquisition of English attributive adjectives dominant language of communication matters for syntactic cross-linguistic influence. *Linguistic Approaches to Bilingualism*, *8*, 193-216. doi: 10.1075/lab.16003.fal

Fallah, N., Jabbari, A. A., & Fazilatfar, A. M. (2016). Source(s) of syntactic cross-linguistic influence (CLI): The case of L3 acquisition of English possessives by Mazandarani-Persian bilinguals. *Second Language Research*, *32*(2), 225-245. doi: 10.1177/0267658315618009

Fernández-Berkes, E., & Flynn, S. (in press). Grammatical mapping in L3 acquisition: A theory of development. In M. Brown, S. Flynn, & E. Fernández-Berkes (Eds.), *L3 development after the initial state.* John Benjamins Publishing Company.

Flynn, S. (1981). The effects of first language branching direction on the acquisition of a second language. In W. Harbert & J. Herschensohn (Eds.), *Working papers* (p. 50-62). Department of Modern Languages and Linguistics, Cornell University.

Flynn, S., Foley, C., & Vinnitskaya, I. (2004). The Cumulative-Enhancement Model for language acquisition: Comparing adults' and children's patterns of development in first, second and third language acquisition

of relative clauses. *International Journal of Multilingualism*, *1*, 3-16. doi: 10.1080/14790710408668175

Foucart, A., & Frenck-Mestre, C. (2010). Grammatical gender processing in L2: Electrophysiological evidence of the effect of L1-L2 syntactic similarity. *Bilingualism: Language and Cognition*, *14*(4), 379-399. doi: 10.1017/S136672891000012X

Gagné, N., & Franzen, L. (2021). How to run behavioural experiments online: Best practice suggestions for cognitive psychology and neuroscience. *PsyArXiv*, *67*. doi: 10.31234/osf.io/nt67j

García-Mayo, M. (2012). Cognitive approaches to L3 acquistioin. *International Journal of English Studies*, *12*(1), 129-146.

García Mayo, M., & Rothman, J. (2012). L3 morphosyntax in the generative tradition: The initial stages and beyond. In J. Cabrelli Amaro, S. Flynn, & J. Rothman (Eds.), *Third language acquisition in adulthood* (p. 9-23). Philidelphia, PA: John Benjamins. doi: https://doi.org/10.1075/sibil.46.04pil

Ghezlou, M., Koosha, M., & Reza Lotfi, A. (2018). Acquisition of adjective placement by l3 learners of english: Evidence for the l2 status factor. *International Journal of Applied Linguistics & English Literature*, *7*(2). doi: 10.7575/aiac.ijalel.v.7n.2p.175

González Alonso, J., & Rothman, J. (2017). Coming of age in L3 initial stages transfer models: deriving developmental predictions and looking towards the future. *International Journal of Bilingualism*, *21*(6), 683-697. doi: 10.1177/1367006916649265

Gorilla. (n.d.). *Gorilla Experiment Builder.* Retrieved from https://gorilla.sc

Granfeldt, J. (2000). The acquisition of the determiner phrase in bilingual and second language French. *Bilingualism: Language and Cognition*, *3*(3), 263-280. doi: 10.1017/S1366728900000377

Grüter, T., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research*, *28*(2), 191–215.

Hammarberg, B. (2012). Third language acquisition. In P. Robinson (Ed.), *The Routledge Encyclopedia of Second Language Acquisition* (p. 644-648). Hoboken: Taylor and Francis.

Hawkins, R., & Chan, C. (1997). The partial availability of Universal Grammar in second language acquisition: the 'failed functional features hypothesis'. *Second Language Research*, *13*(3), 187-226. doi: 10.1191/026765897671476153

Hermas, A. (2010). Language acquisition as computational resetting: verb movement in L3 initial state. *International Journal of Multilingualism*, *7*(4), 343-362. doi: 10.1080/14790718.2010.487941

Hermas, A. (2014a). Multilingual transfer: L1 morphosyntax in L3 English. *International Journal of Language Studies*, *8*(2), 1-24.

Hermas, A. (2014b). Restrictive relatives in L3 English: L1 transfer and ultimate attainment convergence. *International Journal of Multilingualism*, *34*(3), 361-387.

Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, *29*(1), 33-56. doi: 10.1177/0267658312461803

Izura, C., Cuetos, F., & Brysbaert, M. (2014). Lextale-Esp: A test to rapidly and efficiently assess the Spanish vocabulary size. *Psicologica: International Journal of Methodology and Experimental Psychology*, *35*(1), 49-66. Retrieved from https://www.redalyc.org/articulo.oa?id=16930557004

Jaensch, C. (2012). Acquisition of L3 German: Do some learners have it easier? In J. Cabrelli Amaro, S. Flynn, & J. Rothman (Eds.), *Third language acquisition in adulthood.* Amsterdam: John Benjamins.

doi: 10.1353/lan.2014.0011.

Kempe, V., Brooks, P., & Kharkhurin, A. (2010). Cognitive predictors of generalizationof Russian grammatical gender categories. *Language Learning*, *60*(1), 127-153.

Kramer, R. (2015). *The morphosyntax of gender* (First ed.). Oxford University Press.

Lemhöfer, K. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*(2), 325-343. doi: 10.3758/s13428-011-0146-0

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*(2), 325-343. doi: 10.3758/s13428-011-0146-0

Lust, B. (2012). Tracking universals requires Grammatical Mapping. In K. Grohmann & A. Shelkovaya (Eds.), *Linguists of tomorrow: Selected papers from the 1st cyprus postgraduate student conference in theoretical and applied linguistics* (p. 105-140). Cambridge Scholars Publishing.

Marian, V., Blumenfeld, H., & Kaushanskaya, M. (2007). Language Experience and Proficiency Questionnaire (LEAP-Q). *Journal of Speech Language and Hearing Research*, *50*(4), 940-967.

Miceli, G., Turriziani, P., Caltagirone, C., Capasso, R., Tomaiuolo, F., & Caramazza, A. (2002). The neural correlates of grammatical gender: an fMRI investigation. *Journal of Cognitive Neuroscience*, *14*(4), 618-28. doi: 10.1162/08989290260045855

Mitchell, D., & Green, D. (1978). The effects of context and content on immediate processing in reading. *Quarterly Journal of Experimental Psychology*, *30*(4), 609-636. doi: 10.1080/14640747808400689

Montrul, S. (2002). Incomplete acquisition and attrition of Spanish tense/aspect distinctions in adult bilinguals. *Bilingualism: Language and Cognition*, *5*(1), 39–68. doi: 10.1017/S1366728902000135

Nelson, C., Krzysik, I., Lewandowska, H., & Wrembel, M. (2021). Multilingual learners' perceptions of cross-linguistic distances: a proposal for a visual psychotypological measure. *Language Awareness*, *30*(2), 176–194. doi: 10.1080/09658416.2021.1897132

Nemser, W. (1971). Approximative systems of foreign language learners. *International Review of Applied Linguistics in Language Teaching (IRAL)*, *9*, 115-23. doi: 10.1515/iral.1971.9.2.115

Oliphant, K. (1998). Acquisition of grammatical gender in Italian as a foreign language. *Canadian Modern Language Review-Revue Canadienne Des Langues Vivantes*, *54*(2), 239-262. doi: 10.3138/cmlr.54.2.239

Papoutsaki, A., Laskey, J., & Huang, J. (2017). Searchgazer: Webcam eye tracking for remote studies of web search.
doi: 10.1145/3020165.3020170

Papoutsaki, A., Sangkloy, P., Laskey, J., Daskalova, N., Huang, J., & Hays, J. (2016). Webgazer: Scalable webcam eye tracking using user interactions. In *Proceedings of the twenty fifth international joint conference on artificial intelligence—ijcai 2016* (p. 3839-3845). Arlington, VA. Retrieved from https://par.nsf.gov/biblio/10024076-%20webgazer-scalable-webcam-eye-tracking-using-user-interactions

Paradis, M. (1994). Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism. In N. Ellis (Ed.), *Implicit and explicit learning of second languages* (p. 393-419). London: Academic Press.

Pfenninger, S. (2020). Age meets multilingualism: Influence of starting age on L3 acquisition across different learner populations. In C. Bardel & L. Sánchez (Eds.), *Third language acquisition: Age, proficiency*

and multilingualism (p. 167-207). Berlin: Language Science Press. doi: 10.5281/zenodo.4138745

Puig-Mayenco, E., González Alonso, J., & Rothman, J. (2018). A systematic review of transfer studies in third language acquisition. *Second Language Research*, *34*(1). doi: 10.1177/0267658318809147

Puig-Mayenco, E., & Rothman, J. (2020). Low proficiency does not mean ab initio: A methodological footnote for linguistic transfer studies. *Language Acquisition*, *27*(2), 217-226. doi: 10.1080/10489223 .2019.1677668

Puig-Mayenco, E., Rothman, J., & Tubau, S. (2019). Language dominance in the previously acquired languages modulates rate of third language (L3) development over time: a longitudinal study. *International Journal of Bilingual Education and Bilingualism*, 1-24. doi: 10.1080/13670050.2020.1792408

Qualtrics. (2021). *Qualtrics research suite.* http://www.qualtrics.com. Last retrieved October 10, 2019.: Provo, Utah, USA.

Rothman, J. (2010). On the typological economy of syntactic transfer: Word order and relative clause high/ low attachment preference in L3 Brazilian Portuguese. *International Review of Applied Linguistics in Language Teaching (IRAL)*, *48*, 245–273. doi: 10.1515/iral.2010.011

Rothman, J. (2011). L3 syntactic transfer selectivity and typological determinacy: The Typological Primacy Model. *Second Language Research*, *27*, 107–127. doi: 10.1177/0267658310386439

Rothman, J. (2013). Cognitive economy, non-redundancy and typological primacy in L3 acquisition: Evidence from initial stages of L3 Romance. In S. Baauw, F. Dirjkoningen, & M. Pinto (Eds.), *Romance languages and linguistic theory* (pp. 217–247). Amsterdam: John Benjamins. doi: 10.1177/ 0267658310386439

Rothman, J. (2015). Linguistic and cognitive motivations for the Typological Primacy Model (TPM) of third language (L3) transfer: Timing of acquisition and proficiency considered. *Bilingualism: Language and Cognition*, *18*(2), 179–190. doi: 10.1017/S136672891300059X

Rothman, J., & Cabrelli Amaro, J. (2010). What variables condition syntactic transfer? A look at the L3 initial state. *Second Language Research*, *26*(2), 189-218. doi: 10.1177/0267658309349410

Sabourin, L. (2001). L1 effects on the processing of grammatical gender in L2. *EUROSLA Yearbook*, *1*, 159-169. doi: 10.1075/eurosla.1.13sab

Sciutti, S. (2020). The acquisition of clitic pronouns in complex infinitival clauses by German-speaking learners of Italian as an L3: The role of proficiency in target and background language(s). In C. Bardel & L. Sáncheznchez (Eds.), *Third language acquisition: Age, proficiency and multilingualism* (p. 111-144). Berlin: Language Science Press. doi: 10.5281/zenodo.4138741

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching (IRAL)*, *10*(4), 209-31. doi: 10.1515/iral.1972.10.1-4.209

Semmelmann, K., & Weigelt, S. (2018). Online webcam-based eye tracking in cognitive science: A first look. *Behavior Research Methods*, *50*, 451-465. doi: 10.3758/s13428-017-0913-7

Slabakova, R. (2017). The scalpel model of third language acquisition. *International Journal of Bilingualism*, *21*(6), 651-665. doi: 10.1177/1367006916655413

Slabakova, R. (2019). The bottleneck hypothesis updated. In T. Ionin & M. Rispoli (Eds.), *Three streams of generative language acquisition research* (pp. 319–345). Amsterdam: John Benjamins.

Slabakova, R. (in press). A shared linguistic system of multilingual representations. In M. Brown, S. Flynn,

& E. Fernández-Berkes (Eds.), *L3 development after the initial state.* Amsterdam: John Benjamins Publishing Company.

Smith, N. V., & Tsimpli, I.-M. (1995). *The mind of a savant : Language learning and modularity.* Oxford, UK ; Cambridge, USA: Blackwell.

Socarrás, G. (2011). *First language acquisition in Spanish a minimalist approach to nominal agreement.* London: Continuum.

Spino, L. (2022). Investigating grammatical gender agreement in Spanish: A methodological exploration of eye tracking. In D. Ayoun (Ed.), *The acquisition of gender : Crosslinguistic perspectives* (Vol. 63, p. 183-208). Amsterdam / Philadelphia: John Benjamins Publishing Company. doi: 10.1075/sibil.63.08spi

Tufft, M. R., & Gobel, M. S. (2021). Gender and perceived cooperation modulate visual attention in a joint spatial cueing task. *Visual Cognition*, *30*(1-2), 6-27. doi: 10.1080/13506285.2021.1976892

von Grebmer zu Wolfsthurn, S., Pablos Robles, L., & Schiller, N. (2021). Cross-linguistic interference in late language learners: An ERP study. *Brain and language*, *221*, 104993–104993. doi: 10.1016/ j.bandl.2021.104993

Wessel, M. (2021). *Exploration of webcam-based eye tracking by comparing gaze behaviour of dutch sign language users during interpretation of sign language.* (Bachelor Artificial Intelligence: Utrecht University)

Westergaard, M. (2021a). L3 acquisition and crosslinguistic influence as co-activation. response to commentaries on the keynote "microvariation in multilingual situations: The importance of property-by-property acquisition.". *Second Language Research*, *37*(3), 501– 518. doi: 10.1177/02676583211007897

Westergaard, M. (2021b). Microvariation in multilingual situations: The importance of property-by-property acquisition. *Second Language Research*, *37*(3), 379–407. doi: 10.1177/0267658319884116

Westergaard, M., Mitrofanova, N., Mykhaylyk, R., & Rodina, Y. (2017). Crosslinguistic influence in the acquisition of a third language: The Linguistic Proximity Model. *International Journal of Bilingualism*, *21*(6), 666-682. doi: 10.1177/1367006916648859

White, L. (2003). *Second language acquisition and universal grammar.* Cambridge: Cambridge University Press.

White, L., Valenzuela, E., Kozlowka-Macgregor, M., & Leung, Y.-K. I. (2004). Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics*, *25*(1), 105-133. doi: 10.1017/S0142716404001067

A   Appendix: Timeline for General Plan of Work

Month	Goals
2022	
April	Defend Dissertation Prospectus Begin data collection for Experiment 1
May	Complete Procedure sections of Chapters 3 and 4 Continue Experiment 1 data collection
June	Finish Experiment 1 data collection and begin analysis Submit Chapter 3: Grammatical gender acquisition to readers
July	Submit Chapter 2: Models of L3 acquisition to readers Revise Chapter 3 based on feedback
August	Contact German departments to advertise Experiment 2 Hire RA (If applicable) Begin Experiment 2 collection Revise Chapters 2 and 3 based on feedback
September	Continue Experiment 2 data collection Submit Chapter 4: Experiment 1 to readers
October	Continue Experiment 2 data collection Revisions based on feedback
November	Continue Experiment 2 data collection Revisions based on feedback
December	Complete Experiment 2 data collection Revisions based on feedback
2023	
January	Experiment 2 statistical analysis
February	Submit Chapters 5 and 6 (Experiment 2 and Overall results) to readers
March	Submit introduction, conclusion sections to readers Final revisions and defense preparations
April	Dissertation Defense
May	Revisions and Final Dissertation Submission Begin revisions and writing for journal publication